



The Austrian Ludwig Wittgenstein Society
in cooperation with Complutense University of Madrid, University of Warsaw, Technical University Dortmund
and with the support of the Federal Government of Lower Austria, section Science and Research, and the
municipalities of Kirchberg, Otterthal, and Trattenbach is pleased to present the

46th International Wittgenstein Symposium

August 10 – 16, 2025

FEMINIST PHILOSOPHY

LANGUAGE, KNOWLEDGE AND POLITICS

Organised by

Isabel G. Gamero (Madrid)

Amadeusz Just (Warsaw)

Jasmin Trächtler (Dortmund)

BEITRÄGE CONTRIBUTIONS

46. Internationales Wittgenstein Symposium
Kirchberg am Wechsel 2025

46

WISSENSCHAFT · FORSCHUNG
NIEDERÖSTERREICH



FEMINIST PHILOSOPHY

LANGUAGE, KNOWLEDGE AND POLITICS

Beiträge der Österreichischen Ludwig Wittgenstein Gesellschaft

Contributions of the Austrian Ludwig Wittgenstein Society

Band XXXI
Volume XXXI

FEMINIST PHILOSOPHY

Beiträge des 46. Internationalen Wittgenstein Symposiums
10. – 16. August 2025
Kirchberg am Wechsel

Band XXXI

Herausgeber

Isabel G. Gamero (Madrid)
Amadeusz Just (Warsaw)
Jasmin Trächtler (Dortmund)

in Zusammenarbeit mit Joseph Wang-Kathrein

Produziert mit Unterstützung der Abteilung
Wissenschaft und Forschung (K3)
des Amtes der NÖ Landesregierung

WISSENSCHAFT • FORSCHUNG
NIEDERÖSTERREICH 

Kirchberg am Wechsel, 2025

Österreichische Ludwig Wittgenstein Gesellschaft

FEMINIST PHILOSOPHY

Contributions of the 46th International Wittgenstein Symposium
August 10 – 16, 2025
Kirchberg am Wechsel

Volume XXXI

Editors

Isabel G. Gamero (Madrid)
Amadeusz Just (Warsaw)
Jasmin Trächtler (Dortmund)

in Zusammenarbeit mit Joseph Wang-Kathrein

Produced with the support of the
Department for Science and Research (K3)
of the Province of Lower Austria

WISSENSCHAFT • FORSCHUNG
NIEDERÖSTERREICH 

Kirchberg am Wechsel, 2025
Austrian Ludwig Wittgenstein Society

Verleger | Publisher

Österreichische Ludwig Wittgenstein Gesellschaft

Austrian Ludwig Wittgenstein Society

Markt 63, A-2880 Kirchberg am Wechsel

Österreich/Austria

www.alws.at

Vorstand der ÖLWG | Executive Committee of the ALWS

Christian Kanzian (Präsident der ÖLWG)

Elisabeth Nemeth (Vizepräsidentin der ÖLWG)

Anja Weiberg (Vizepräsidentin der ÖLWG)

Joseph Wang-Kathrein (Geschäftsführer der ÖLWG)

Bettina Bussmann

Marian David

Esther Heinrich-Ramharter

Josef Mitterer

Volker Munz

Alois Pichler

Mira Magdalena Sickinger

Jasmin Trächtler

Consulting Board der ÖLWG | Consulting Board of the ALWS

Peter Kampits

Klaus Puhl

Friedrich Stadler

Paul Weingartner

ISSN 1022-3398 | ISBN 978-3-9505512-3-5

Refereed Periodical

All Rights Reserved

Copyright 2025 by the Authors

Copyright will remain with the author, rights to use with the society. No part of the material may be reproduced or utilised in any form or by any means, electronically or mechanically, including photocopying, recording, informational storage, and retrieval systems without permission from the society.

Die Redaktion dankt allen Gutachtern für ihre äußerst hilfreichen Rezensionen.

The editors are grateful to all referees for their most helpful reviews.

Die Beiträge und das Abstracts-Heft wurden mit Unterstützung einer von Joseph Wang-Kathrein, Universität Innsbruck, erarbeiteten Datenbank erstellt. Kontakt: joseph.wang@uibk.ac.at

The *Contributions* and the booklet of *Abstracts* were compiled with the support of a database developed by Joseph Wang-Kathrein, University of Innsbruck. Contact: joseph.wang@uibk.ac.at

Inhalt / Contents

In Shadow: Stereotyping and Normative Frameworks	
Rena Alcalay	10
When Feelings Don't Fit: Asexuality, Allonormativity, and Affective Injustice	
Kira aus dem Bruch & Sven Walter	22
The Material Theory of Values in Science	
Kelli Barr	34
Susanne K. Langer, The Vienna Circle, and <i>The Practice of Philosophy</i>	
Mark Nader Basafa	45
Marginalized Worldviews in Legal Institutions: Towards a Legal Hinge Epistemology	
Melina Bentes	54
Why is Wittgenstein's anthropological turn, in fact, not a post-positivist sociological manifesto? David Bloor as his own enemy	
Mehmet Onur Canpolat	63
What Hilbert really wanted to say - Frege's clarification of metamathematics	
Quirin Colares Oberrauch	73
Language and Methodological Limits in Wittgenstein and Hayek	
Jeffrey Patrick Colgan & David Ween	83
Political obligation and feminist theory. A proposal.	
Nicole Darat	94
AI ethics and the diversity of forms of life	
Michiel De Proost & Jonathan Adams	104
Seeing the World Aright: Wittgenstein, Murdoch and the Ethical Engagement with Literature	
Marcello Di Massa	115
Realism, Wittgenstein, and the Formative Years of the Croatian Analytic Tradition	
Dušan Dožudić	124
Vom Denken zum Handeln: Verbindungen vom Einschleichen der Kunst nach Wittgenstein zu feministischer Kunst	
Lilli Förster	136
A Moral Inquiry into Unintentional Speech	
Christina Fritz	148
A (quasi)-Mathematical Theory of Public Language	
Giovanni Galli & Niccolò Covoni	156
From Meaning-Facts to Form of Life: Dissolving Kripke's Puzzle with Wittgenstein	
Ruosen Gao	166

Acknowledgment as Feminist Praxis: Reframing Skepticism and Empathy through Wittgenstein, Cavell, and Aesthetic Media	174
Freya Katharina Gerz	
Wittgenstein on concepts and forms	183
Nicola Gianola	
Doing Justice to the Swansea School: An Appropriate Wittgensteinian Approach to the Philosophy of Religion	192
Jakub Gomułka & Jan Wawrzyniak	
Frauen im Wiener Kreis: Zur epistemischen Ungerechtigkeit und Kanonkritik in der analytischen Philosophie	202
Patricia Grill	
Philosophy is deceitful above all things. On non-dualising speech, the duck-rabbit and JT LeRoy	214
Thomas Hainscho	
“A Separation from the Original Soil”: Wittgenstein on Anthropological Explanation in the "Remarks on Frazer's <i>Golden Bough</i>"	225
Kyle Hardy	
The Window and the Soul: Wittgenstein, Woolf, and the Ethics of Recognition	233
Iain Haslam	
Nonbinary Identity and Social Meaning: A Reply to Kadji Amin	243
Matthew Hoffman	
The Limits of Certainty: Distinguishing Moral Certainties from Empirical Certainties	252
Minoo Hojjat	
Discovering Ethics in Wittgenstein's Picture Theory	261
Herbert Hrachovec	
Anscombe and Diamond on Wittgenstein's Notion of Sense in the <i>Tractatus</i>	269
Andrej Jandrić	
Embodied Hermeneutic Injustice	278
Miriam Jerade & Celeste Vecino	
Gender in Psychoanalysis. A brief note referring to Wittgenstein	288
Ulrike Kadi	
Anscombe's Social Epistemology	299
André Kfourì	
The origin of language in Cavell, Mulhall, and Wittgenstein	308
So Kimoto	
“Doing things with names” Recognising and Resisting the Harms of Deadnaming	318
Anna Klieber	

L. Susan Stebbing: The Art of Thinking Clearly	328
Silke Körber	
Power and the ethics of innovating: defining deliberative injustice for hermeneutic technology assessment	338
Benedict Lane	
A Note on Moore's 'Combined Reading' of the <i>Tractatus</i>	350
David Lindeman	
On Perceptual Change	360
Camila Lobo	
When Clocks Stop Working – Chrononormativity and Rhetorical-Epistemic Oppression	369
Flora Löffelmann	
Inhabiting the Dilemma: Wittgenstein, Feminism, and the Question of the Subject	380
Saori Makino	
Deep Disagreements: The Importance of Examples	389
Ryan Manhire	
Simone Weil and Wittgenstein on Character	399
Nimrod Matan	
Wittgenstein vs. the Conceptual Engineers	407
Jacob McDowell	
Family resemblances, semantic drift, and polysemy	417
Henrique Mendes	
<i>White Fraternities: A Tale of Two Smotherings</i>	426
William Moix	
What If Nothing Were the Case? On Empty Space in Wittgenstein's <i>Tractatus</i>	436
Filippo Mosca	
Epistemic and Affective Advantages of Being Oppressed	446
Carmen Mossner	
On Context in Wittgenstein's Philosophical Investigations	456
Gonzalo Nuñez	
Two Ways of Taking the Ordinary Seriously	466
Eylem Özaltun	
The default use of gender-neutral pronoun they and our epistemic (and non-epistemic) lives	474
Jeanne Pang	
Strange Props: Reflections on the Status of Hinges in and beyond <i>On Certainty</i>	481
Samuel Pedziwiatr	

Realist Intuitions in Standpoint Epistemology: Strategies for Reclaiming "Truth" Victoria Pérez Monterroso	493
Was Wittgenstein Aiming at an Identity Theory of Truth? Jimmy Plourde	501
Die Frage nach der Lebensform – Wittgensteins Begriff der Lebensform und die Begriffe Lebensmuster, Lebensgepflogenheit, Lebensschablone, Lebensteppich auf dem Prüfstand und ein religionsphilosophischer Exkurs Moritz René Pretzsch	510
Feminist anxieties and practical genealogy Ronya Ramrath	522
The method of grammatical analysis in Philippa Foot's early moral philosophy Toni Rašić	531
Giving body to hinge epistemology Janyne Sattler	540
The Threat of Solipsism: Understanding the Net with Iris Murdoch and Susan Stebbing. Francesca Scapinello	549
Scepticism and „Moorean facts“ in <i>On Certainty</i> – a deflationist reading Lauritz Seeberg	558
A Feminist Approach to Verbal Humour Mira Magdalena Sickinger	567
Revealing World-Pictures: Cordova, Wittgenstein, Culture, and Alterity Daniel Smith	577
On Irigaray's and Wittgenstein's Critiques of Formal Logic Kim Solin	587
Standpoint Epistemology and the Problem of Reflection. Carla Spielmann	598
Towards a Feminist Philosophical Inquiry on Motherhood Krissah Marga Taganas	608
On Wittgenstein on Translation Tue Trinh	619
SSK and Feminist Epistemology as Adverse Allies - Meaning Finitim as a Resource for Debates on Hermeneutical Oppression Sophie Juliane Veigl & Sonja Riegler	628
The Present State of Hermeneutic Injustice Regarding Sexual Harassment and Sexual Violence Marcia Villanueva	638

The 'Arbitrariness' of Ethics: Pursuing Wittgenstein's Suggested Analogy to Logic and Mathematics	
Vincent Vincke	649
Power – A Surveyable Representation	
Robert Vinten	660
Coerced Illocutions: 'Don't Put Words in My Mouth'	
Eirini Vryza	670
AI Models as Conceptual Agents	
Lena Wang	681
Notes on MS 168 – "... a bag full of raisins"	
Peter K. Westergaard	691
Waismann Says only Fools Rush In: The Waismannian Mechanism in Conceptual Engineering	
Jennifer Whyte	701
Why feminist epistemology needs feminist aesthetics. The feeling of unease as an example	
Katrin Wille	709
Criticism as Third-Order Change—A Feminist Examination of Wittgensteinian Practices	
Yara Windmüller	718
Beyond a question that has no name: Perspectival sense-making and hermeneutical Injustice	
Yinmei Wu	728
Wittgenstein und J. H. Newman über die Rolle von „Bildern“ im religiösen Glauben.	
Isacco Zampini	739
A Relativist Ameliorative Approach to Gender Terms	
Dan Zeman	750
Planetary and inter-scalar entanglements of feminist new materialisms, exemplified by Astrida Neimanis's Bodies of Water: <i>Posthuman Feminist Phenomenology</i>	
Wer Ziemian	758

In Shadow: Stereotyping and Normative Frameworks

Rena Alcalay (Munich, Germany)

Abstract

This paper explores Wittgenstein's metaphors of SHADOW and HINGE to develop an account of noematic correctness—a normative structure underlying concept application and judgment, particularly under conditions of uncertainty. Drawing from *Philosophical Investigations* and *On Certainty*, it argues that certain linguistic and conceptual judgments are not reducible to semantic truth or descriptive fit, but are structured by pre-intentional, context-sensitive rules often followed without explicit articulation. The metaphor of SHADOW captures the ambiguity involved in interpreting generalizations when no clear rule determines their application. HINGE, by contrast, refers to ungrounded certainties—propositions or norms exempt from doubt—that form the background conditions for justification. Together, these metaphors help explain how stereotyping flexibly adapts to novel contexts.

Building on critiques of non-normative accounts of semantic and noematic correctness (e.g., Glüer & Wikforss 2009), the paper defends the view that heuristic judgments exhibit *sui generis* normative structure. Heuristics, particularly those involved in stereotyping, show that conflicting intentions are not arbitrary or mistaken but often reflect HINGE-like commitments that guide how rules are followed and adapted. This reinterpretation offers a response to Kripke's skeptical dilemma by rejecting the regress of intentional pro-attitudes, instead locating normativity in the pre-intentional yet rationally effective structure of judgment. In doing so, the paper contributes to a Wittgensteinian framework for understanding intuitive judgment-making and sheds new light on concerns in feminist epistemology, particularly how individuals apply stereotypes, navigate ambiguity, and intuitively assess norm conformity without relying solely on explicit reasoning.

When does one have the thought: the possible movements of a machine are already there in it in some mysterious way?—Well, when one is doing philosophy. And what leads us into thinking that? The kind of way in which we talk about machines. We say, for example, that a machine *has* (possesses) such-and-such possibilities of movement; we speak of the ideally rigid machine which *can* only move in such-and-such a way. — What is this *possibility* of movement? It is not the *movement*, but it does not seem to be the mere physical conditions for moving either—as, that there is a play between socket and pin, the pin not fitting too tight in the socket. For while this is the empirical condition for movement, one could also imagine it to be otherwise. The possibility of a movement is, rather, supposed to be like a shadow of the movement itself. But do you know of such a shadow? And by a shadow I do not mean some picture of the movement—for such picture would not have to be a picture of just *this*

movement. But the possibility of this movement must be the possibility of just this movement. (See how high the seas of language run here!) (*PI* 1968: §194)

In Part I of the *Philosophical Investigations*, Wittgenstein introduces the metaphor of the _{SHADOW} to highlight the complexities of rule-following and the implicit assumptions underlying our linguistic and cognitive practices (*PI* 1968: §194). He is, of course, speaking to the unrealistic nature of philosophizing. However, this metaphor, when taken with his reflections on HINGES in *On Certainty* can also provide a compelling framework for defending content normativity, or the idea that mental content can involve correct concept categorization (*OC* 1969). This paper argues that content normativity can be grounded in noematic correctness—a term coined by Glüer and Wikforss (2009) to denote correct concept categorization. More broadly, this line of critical reflection aims to place the anachronistic concept of stereotyping into a Wittgensteinian interpretive framework.

Indeed, Glüer and Wikforss (2009) deny noematic correctness is normative. However, by shifting focus from semantic facts to the mental event of heuristics, I propose that heuristics exhibits *sui generis* qualities of noematic normativity. In making this shift, we can turn around the skeptic's doubt of linguistic intentions by showing that conflicting intentions in rule application are not arbitrary but are rational responses to novel contexts. This perspective sheds light on the heuristic of stereotypical thinking, where conceptual and linguistic rules guide quick, efficient judgments via rule-guided heuristics and where rule-guided heuristics can contribute to morally problematic biases—a concern often described as the problem of moral encroachment. Through this analysis, I aim to highlight how Wittgenstein's metaphors of _{SHADOW} and _{HINGE} offer valuable insights into how individuals adhere to conceptual rules and intuitively judge whether an utterance aligns with the correct norm, even without explicit reasoning. I will show that the concept of _{SHADOW} in the *PI* can be used to signify the role of rule-guided heuristics under conditions of uncertainty.

Considering the interaction between a speaker and an interlocutor, Wittgenstein notes how the interlocutor must guess at what the speaker intends.

But do you really explain to the other person what you yourself understand? Don't you get him to guess the essential thing? You give him examples, --but he has to guess their drift, to guess your intention." – Every explanation which I can give myself I give to him too. – "He guesses what I intend" would mean: various interpretations of my explanation come to his mind, and he lights on one of them. So, in this case he could ask; and I could and should answer him. (1968: §210)

Kripke draws from this and other passages in *PI* in his sketch of the now-famous skeptical dilemma: that it is not logically impossible for a speaker to misrepresent their prior uses of a concept, symbol, or word, though one can be mistaken about one's previous linguistic intentions (2000: 9–13). In such cases, interpretation begins to resemble a regress, where any application of a rule seems to presuppose another rule—one that the speaker must, in some way, intend to follow. As Glüer and Wikforss argue, this structure suggests that rule-following always depends on an impossible intentional pro-attitude toward the rule being applied (2009: 58).

The regress of intentional pro-attitudes appears in *PI*, where Wittgenstein writes, "How can he know how he is to continue a pattern by himself—whatever instruction you give him?" (1968: §211). This is not an isolated insight—it forms part of a broader pattern in his later philosophy: that rule-following is never a purely mechanical process (*PI* 1968: §§185–242). For instance, Wittgenstein uses the image of a machine, or a machine diagram, as a metaphor for symbolic action—but one that only appears determinate when retrospectively interpreted (*PI* 1968: §193). He complicates this further by showing how we often attempt to use clearly defined references to predict the behavior of less certain or ambiguous cases (*PI* 1968: §194). These passages emphasize our tendency to compare the uncertain to the definite, often without justification. Ongoing debates explore whether inference itself is rule-guided or whether Wittgenstein ultimately abandoned the idea that language

is governed by rules altogether (see Miller 2015; Glüer & Wikforss 2010). These tensions underscore the fragility of rule-following when divorced from shared practices and intentional structures.

While _{SHADOW} illustrates the complexity of predicting uncertain behavior based on definite references, HINGE complements this by addressing the ungrounded assumptions underlying our rule-following practices. For instance, in *OC*, Wittgenstein reflects on the proposition, “I have never been to Asia Minor,” asking, “where does this knowledge come from?” (1969: §419). If it comes from my memory, can I be wrong? He proffers that while false memories are possible, a complete distrust of memory would cause all other judgments to collapse.

Wittgenstein worries that the word _{KNOW} is being stretched too far in these contexts. Instead, he suggests that some beliefs function as HINGES—such as “*my memory does not deceive me*”—which we hold *beyond all reasonable doubt*. HINGES are not indubitable in a strict sense, but they provide the stability necessary for making other judgments.

HINGE metaphorizes the role that some propositions play within the system of beliefs: “That is to say, the questions that we raise, and our doubts depend on the fact that some propositions are exempt from doubt, are as it were like hinges on which those turn” (*OC* 1969: §341). HINGE plays the role of ungrounded assumptions required for justifying other beliefs. Some beliefs are exempt from doubt, and are justification for other, dubitable propositions. As Anna Boncompagni writes, hinges are propositions that remain “fixed in virtue of the movement around it,” like an “axis of rotation” (2019: 80). Although hinges are ungrounded assumptions, they still constitute our rational practices (Coliva 2015). Examples of HINGES from Wittgenstein’s *On Certainty* include ‘I am a human being’ (*OC* 1969: §4), ‘there is an external world’ (*OC* 1969: §20), and ‘the earth has existed a very long time’ (*OC* 1969: §233). Hinges are ones that “no human being could give up” because of the foundational role they play in the system of beliefs (Moyal-Sharrock 2000: 60). They are often tacit, unspoken, and sometimes unnoticed.

Wittgenstein’s concern with rule-guided judgments extends to the study of heuristics, albeit anachronistically, because we tend to employ heuristics with

a sense of it being correct. Yet, often, there is ambiguity in these judgments. For example, consider a simple case of generic ambiguity (Leslie 2007; Haslanger 2012, 2014). Imagine two Entomologists, Joachim and Holly, engaged in a conversation. Let M represent a judgment.

(M) Mosquitoes carry West Nile virus.

Holly and Joachim must have common ground to infer Joachim's intention that M is not meant to assert a universal claim but rather as a species-level generalization about the potential of mosquitoes to carry the virus. It has long been debated whether it is the speaker's intention that determines the meaning of the utterance or is structured through shared norms available to both the speaker and the hearer. (For speaker-oriented intentionalist views see, e.g., Davidson 1986; Kaplan 1989; Neale 1992, 2016; Predelli 1998; Recanati 2005; Perry 2009; Korta and Perry 2010; Montminy 2010; Stokke 2010; Åkerman 2010; King 2013, 2014; Speaks 2016, 2017. For hearer-oriented anti-intentionalist views see, e.g., Kaplan 1978; McGinn 1981; Wettstein 1984; Dummett 1986; Travis 1989; Reimer 1991; Green 2001; Saul 2002; Gauker 2008; Mount 2008; Lepore and Stone 2015). Shared understanding relies on implicit linguistic rules and pragmatic inferences. Signs in language are only meaningful if they can be contextually judged according to their use (*PI* 1968: §243-61). However, Wittgenstein points out that the connection between the intended rule and semantic correctness is not always intentional (cf. Crispin Wright 2007: 486-7).

Since both Entomologists have a shared understanding of their scientific domain, the ambiguity would be quickly displaced. So, let us change the context: Joachim and Holly are high school seniors in an entomology class. In this context, we can imagine Holly and Joachim are unable to rely on implicit linguistic rules to make sense of Joachim's intention. Holly might think Joachim intends a majority generalization—which states that an attribute is generalizable on the basis that a statistically significant number in the group have the attribute—in which case M would be false. Whereas M is correct if Joachim intends the *striking property* rule, which states that an attribute, α , is generalizable to a group, Γ , only if at least one member, x , of Γ has α (cf. Leslie 2007). (However, this condition must be interpreted more broadly: for α to be generalizable, Γ must also serve as a reasonably good predictor of α , and Γ

should be disposed, in the right circumstance, to have α . That is, mere existence is insufficient for robust generalization; the category must also be a *reasonably good predictor* of the property and be *disposed* to have it under the right conditions.) If Joachim is unable to explain the rule for *striking property* to Holly because it requires advanced understanding, a debate could ensue over the correctness of *M*. And even if Joachim can explain *that* rule, there may come a point where he cannot explain anymore. He may just say, ‘It just does,’ and fall back on intuition (PI 1968: §211). Indeed, Wittgenstein remarks, “intuition an unnecessary shuffle,” reflecting, I believe, his frustration with the continual detours or workarounds involved in trying to articulate an understanding of intuition (PI 1968: §213).

This part of Wittgenstein’s reflections is relevant to the intuitive judgment of heuristical correctness. How can I know to be certain—sometimes without reason—that *this* rule adheres to the correct judgment in this context? Consider the German word *morgen*. In one context, it means *tomorrow*—as in *Die Sonne wird morgen aufgehen* (“The sun will rise tomorrow”). Here, *morgen* situates an event in time. But shift the context slightly—*Guten Morgen!*—and suddenly, *morgen* no longer refers to the future. It now means *morning*, a time of day rather than a day to come. The word remains the same, but the judgment guiding its use changes entirely. Meaning is not just fixed to words but to the judgments that frame and guide them, judgments that depend on context-sensitive rules we often follow without articulating.

Some theorists argue that noemata merely describe the structure of consciousness without imposing any normative constraints on judgment. On this view, noematic correctness, like semantic correctness, is considered a non-normative concept (Glüer and Wikforss 2009). According to the non-normative account of semantic correctness, the conditions under which a concept *C* correctly applies to an object are simply the conditions under which judgments involving *C* are true. Correctness, in this sense, tracks truth but without implying that one ought to apply the concept in any particular way. It is a matter of descriptive fit, not normative guidance.

If correctness were only a matter of descriptive fit, why would we ever feel that some judgments are wrong, even when they match observable patterns? One influential approach to this question has been through the lens of moral

encroachment, which examines how moral considerations can influence the epistemic status of beliefs (see Bolinger 2020 for an overview). For instance, the stereotype in the U.S. that “Mexicans are illegal” may appear to align with certain statistical reports. Still, it remains deeply problematic precisely because it collapses legal status and personhood in ways we would reject in other domains. Scholars such as Basu (2019a, 2019b, 2019c, 2020) argue that moral stakes can and should bear on whether beliefs count as justified or count as knowledge.

But we need not appeal to moral stakes alone. There are epistemic reasons why certain judgments feel incorrect. This is where the notion of noematic correctness becomes important. Judgments are not just beliefs that track empirical regularities but also intentional acts that apply mental content in contextually structured ways. The content of a belief—how it is given or meant in experience—can fail to meet epistemic appropriate standards. In such cases, a judgment may feel wrong not because it violates a moral norm but because it misfires in how the concept is applied—in the rule it commits to. Noematic correctness, then, captures this felt epistemic misfit: it reflects the structured way in which concepts ought to be applied to phenomena—not on the basis of statistical or semantic criteria, but primarily through the commitments one makes in following a rational or logical rule.

Therefore, noematic correctness should be considered a normative structure for applying concepts in judgment. The metaphors of _{SHADOW} and _{HINGE} bridge the epistemic uncertainty that some propositions anchor our ability to make judgments at all—even under ambiguity. This epistemic ambiguity—where statistical or semantic fit no longer secures judgment—corresponds to what Wittgenstein characterizes as the _{SHADOW}: the space where meaning is underdetermined, and interpretation requires guesswork. It is precisely in such contexts that noematic correctness does its work. It provides a structured but non-empirical guide for concept application, much like Wittgenstein’s _{HINGE} propositions—those commitments that are not justified but that make justification possible. Just as “my memory does not deceive me” undergirds meaningful empirical claims, so too does “this rule applies here” function as a

^{HINGE} in the realm of judgment. In this way, the ^{SHADOW} reveals the terrain where ^{HINGE}-like certainties, grounded in noematic structure, become essential to navigating conceptual uncertainty.

This view of noematic correctness also sheds light on a key tension in Kripke's skeptical challenge—namely, the problem of conflicting intentions in rule application. But on the view I have outlined, conflicting intentions do not entail error. Rather, they are integral to how normative rules are rationally adapted to novel contexts. A speaker's current judgment may diverge from their previous linguistic intention while still legitimately following a rule. This is especially evident in cases like stereotyping, where fast, concept-driven generalizations emerge in fluid contexts. Such generalizations are shaped by normative heuristics, even when the speaker lacks a fully formed intention about how the rule ought to apply. In these moments, the speaker is not always mistaken; they might be engaging in a context-sensitive modulation of rule-use that reflects the flexibility of norm-guided cognition.

Crucially, this reinterpretation undercuts the force of the skeptical regress. If we recognize that some elements of our rule-following—particularly those embedded in doxastic commitments—function as hinges, then the chain of presupposed pro-attitudes need not extend infinitely. That is, we need not maintain an explicit or intentional pro-attitude toward every rule we follow. Some rules operate at the level of ^{HINGES}. On this view, the ^{HINGE} is the recognition that one's present linguistic intention may, at times, conflict with past intentions. This does not reflect irrationality, but rather the fact that not all motivations to follow rules are intentional. Some are structured by noematic forms of commitment that are pre-intentional, yet still normatively effective. The rule functions, in such cases, as an ungrounded but operative ^{HINGE}—not something one adopts via a pro-attitude, but something one stands upon to adopt any attitude at all.

To say that a rule functions as a ^{HINGE} is not to reduce it to implicit rule-following either. It is, rather, to acknowledge that some of our most fundamental norm-guided acts, including intuitive judgments, are enabled by orientations that are not themselves up for rational evaluation. They are exempt yet structurally necessary. In this way, the ^{HINGE} serves as both a

boundary of justification and a condition for content determinacy, allowing rule-following to proceed even in the face of ambiguity, adaptation, and conflicting intentions.

References

- Åkerman, Jonas (2010) "Communication and indexical reference", *Philosophical Studies* 149, 355–366.
- Basu, Rima (2019a) "Radical moral encroachment: The moral stakes of racist beliefs", *Philosophical Issues*. <https://doi.org/10.1111/phils.12137>
- Basu, Rima (2019b) "What we epistemically owe to each other", *Philosophical Studies* 176, 915–931.
- Basu, Rima (2019c) "The wrongs of racist beliefs", *Philosophical Studies* 176, 2497–2515.
- Basu, Rima (2020) "The spectre of normative conflict", in: Erin Beeghly; Alex Madva (eds.) *An Introduction to Implicit Bias: Knowledge and the Social Mind*, London: Routledge. <https://doi.org/10.4324/9781315107615>
- Blum, Lawrence (2004) "Stereotypes and stereotyping: A moral analysis", *Philosophical Papers* 33, 251–289. <https://doi.org/10.1080/05568640409485143>
- Bolinger, Renée Jorgensen (2020) "Varieties of moral encroachment", *Philosophical Perspectives* 34, 5–26.
- Boncompagni, Andrea (2018) "Wittgenstein on meaning, use, and linguistic commitment", in: David G. Stern (ed.) *Wittgenstein in the 1930s*, Cambridge: Cambridge University Press, 77–92. <https://doi.org/10.1017/9781108349260.005>
- Boncompagni, Andrea (2019) "Hinges, prejudices, and radical doubters", *Wittgenstein-Studien* 10, 165–181. <https://doi.org/10.1515/witt-2019-0010>
- Coliva, Annalisa (2015) *Extended Rationality: A Hinge Epistemology*, Dordrecht: Springer.
- Davidson, Donald (2005) "A nice derangement of epitaphs", in: *Truth, Language, and History*, Oxford: Clarendon Press, 89–107. (Originally published 1986)
- Dummett, Michael (1986) "A nice derangement of epitaphs: Some comments on Davidson and Hacking", in: Ernest Lepore (ed.) *Truth and Interpretation: Perspectives on the Philosophy of Donald Davidson*, Oxford: Blackwell, 459–476.
- Gauker, Christopher (2008) "Zero tolerance for pragmatics", *Synthese* 165, 359–371.

- Glüer, Kathrin; Wikforss, Åsa (2009) "Against content normativity", *Mind* 118, 31–70. <https://doi.org/10.1093/mind/fzn154>
- Glüer, Kathrin; Wikforss, Åsa (2010) "Es braucht die Regel nicht: Wittgenstein on rules and meaning", in: Daniel Whiting (ed.) *The Later Wittgenstein on Language*, Basingstoke: Palgrave Macmillan, 148–166.
- Green, Kent (2001) "Davidson's derangement: Of the conceptual priority of language", *Dialectica* 55, 239–258.
- Haslanger, Sally (2012) *Resisting Reality: Social Construction and Social Critique*, New York: Oxford University Press.
- Haslanger, Sally (2014) "The normal, the natural and the good: Generics and ideology", *Politica & Società* 3, 451–452. <https://doi.org/10.4476/78098>
- Kaplan, David (1989) "Afterthoughts", in: Joseph Almog; John Perry; Howard Wettstein (eds.) *Themes from Kaplan*, Oxford: Oxford University Press, 565–614.
- Kaplan, David (1990) "Dthat", in: Paolo Yourgrau (ed.) *Demonstratives*, Oxford: Oxford University Press, 11–33. (Originally published 1978)
- King, Jeffrey C. (2013) "Supplementives, the coordination account, and conflicting intentions", *Philosophical Perspectives* 27, 288–311.
- King, Jeffrey C. (2014) "Speaker intentions in context", *Noûs* 48, 219–237.
- Korta, Kepa; Perry, John (2010) "Intentions to refer", in: Luis Baptista; Edgar Rast (eds.) *Meaning and Context*, Bern: Peter Lang, 161–186.
- Kripke, Saul A. (2000) *Wittgenstein on Rules and Private Language: An Elementary Exposition*, 10th printing, Cambridge, MA: Harvard University Press.
- Lepore, Ernest; Stone, Matthew (2015) *Imagination and Convention: Distinguishing Grammar and Inference in Language*, Oxford: Oxford University Press.
- Leslie, Sarah-Jane (2007) "Generics and the structure of the mind", *Philosophical Perspectives* 21, 375–403. <https://doi.org/10.1111/j.1520-8583.2007.00138.x>
- McGinn, Colin (1981) "The mechanism of reference", *Synthese* 49, 157–186.
- Miller, Alexander (2015) "Blind rule-following and the 'antinomy of pure reason'", *The Philosophical Quarterly* 65, 396–416. <https://doi.org/10.1093/pq/pqv023>
- Montminy, Martin (2010) "Context and communication: A defence of intentionalism", *Journal of Pragmatics* 42, 2910–2918.

- Mount, Alison (2008) "Intentions, gestures, and salience in ordinary and deferred demonstrative reference", *Mind & Language* 23, 145–164.
- Moyal-Sharrock, Danièle (2005) "Unravelling certainty", in: *Readings of Wittgenstein's On Certainty*, London: Palgrave Macmillan, 76–99.
- Neale, Stephen (1992) "Paul Grice and the philosophy of language", *Linguistics and Philosophy* 15, 509–559.
- Perry, John (2009) "Directing intentions", in: Joseph Almog; Paolo Leonardi (eds.) *The Philosophy of David Kaplan*, New York: Oxford University Press, 187–201.
- Predelli, Stefano (1998) "I am not here now", *Analysis* 58, 107–115.
- Puddifoot, Katherine (2021) *How Stereotypes Deceive Us*, Oxford: Oxford University Press. <https://doi.org/10.1093/oso/9780192845559.001.0001>
- Recanati, François (2005) "Literalism and contextualism: Some varieties", in: Gerhard Preyer; Georg Peter (eds.) *Contextualism in Philosophy: Knowledge, Meaning, and Truth*, Oxford: Clarendon Press, 171–196.
- Reimer, Marga (1991) "Demonstratives, demonstrations, and Demonstrata", *Philosophical Studies* 63, 187–202.
- Saul, Jennifer Mather (2002) "Speaker meaning, what is said, and what is implicated", *Noûs* 36, 228–248.
- Speaks, Jeff (2016) "The role of speaker and hearer in the character of demonstratives", *Mind* 125, 301–339.
- Speaks, Jeff (2017) "A puzzle about demonstratives and semantic competence", *Philosophical Studies* 174, 709–734.
- Stern, David G.; Rogers, Brian; Citron, Gabriele (eds.) (2016) *Wittgenstein: Lectures, Cambridge 1930–1933: From the Notes of G. E. Moore*, Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9781139644600>
- Stokke, Andreas (2010) "Intention-sensitive semantics", *Synthese* 175, 383–404.
- Travis, Charles (2001) *The Uses of Sense: Wittgenstein's Philosophy of Language*, Oxford: Oxford University Press. (Originally published 1989)
- Wettstein, Howard K. (1984) "How to bridge the gap between meaning and reference", *Synthese* 58, 63–84.
- Wittgenstein, Ludwig (1968) *Philosophical Investigations*, 2nd ed., Oxford: Basil Blackwell. (PI)

Wittgenstein, Ludwig (1969) *On Certainty*, ed. by G. E. M. Anscombe and G. H. von Wright, trans. by Denis Paul and G. E. M. Anscombe, Oxford: Basil Blackwell. (OC)

Wright, Crispin (2007) "Rule-following without reasons: Wittgenstein's quietism and the constitutive question", *Ratio* 20, 481–502. <https://doi.org/10.1111/j.1467-9329.2007.00379.x>

When Feelings Don't Fit: Asexuality, Allonormativity, and Affective Injustice

Kira aus dem Bruch, Sven Walter (Osnabrück, Germany)

Abstract

Examining how morally questionable norms, ideologies, attitudes, etc. shape how members of certain groups experience themselves and their environment in ways that impair their ability to make sense of, express, regulate, and share their feelings, several authors have recently argued that there is a specific kind of *affective injustice* individuals experience in their capacity as affective beings. This paper expands the burgeoning debate on affective injustice by drawing attention to an hitherto overlooked manifestation of such injustice.

We argue that the ways “Aces,” i.e., asexual-identified individuals who experience a sustained or near-total lack of sexual attraction, are “othered” as “affect aliens” by our culture of allonormativity that takes it for granted that all healthy, able-bodied adults experience sexual attraction, and that sex is a prerequisite for “normal” and “natural” human flourishing and both a cornerstone and the epitome of fulfilling partnerships is one manifestation of affective injustice.

“Asexuality” is an umbrella term for a spectrum of sexual orientations of individuals whose common point of identification is the experience of a sustained or near-total lack of sexual attraction to others, whether lifelong or occasional, while their physical, emotional, social, romantic etc. proclivities can otherwise differ in fluid ways (e.g., Higginbottom, 2024). Asexual-identified individuals, or “Aces,” do not (constantly) experience the affectionate urge, drive, pull, or longing that makes others perceive interpersonal encounters as emotionally charged erotic affordances for sexual intimacy. They violate what Brandley and Spencer (2023, p. 10) call “the hegemony of allonormativity”—the taken-for-granted assumption that all healthy, able-bodied adults naturally experience sexual attraction, and that sex is a prerequisite for human flourishing and both a cornerstone and the epitome of a fulfilling partnership. Throughout the past decade or two, researchers across various disciplines have begun to investigate the effects this minority stress has on Aces (e.g., MacNeela & Murphy, 2015). Strikingly, philosophy so far shows almost no interest in asexuality. The purpose of this paper is to start remedying this lacuna by examining the personal, societal, and potentially political, significance and meaning of the emotional toll Aces may have to endure as a

consequence of being “otherized” by the dominant ideal of compulsory allosexuality.[1] Specifically, we contend that allonormativity can be seen as making Aces suffer from what has recently been dubbed “*affective injustice*.”

The notion of “affective injustice” emerged in response to Fricker’s (2007) work on how social power dynamics can undermine individuals’ fundamental ability to participate in epistemic practices of knowing, understanding, and being understood. When those at the center of power systematically impair or dismiss the experiences of those at the margins, the latter are unjustly wronged not (only) epistemically as knowers, but (also) affectively in their ability to express, regulate, and share affective experiences (e.g., Gallegos, 2022; Srinivasan, 2018; Whitney, 2018). For instance, when Audre Lorde expressed her anger at the injustices experienced by Black women within academic and feminist spaces, a white woman requested: “Tell me how you feel but don’t say it too harshly or I cannot hear you” (1984, p. 125). By refusing to engage with Lorde’s justified anger while simultaneously demanding that Lorde show consideration for her own feelings in turn, the white woman not only dismissed Lorde’s epistemic credibility, but also denied her anger what Frye (1983) calls “uptake”—she declined to participate in the “*uniquely affective cooperative behavior whereby we take someone seriously as an affective being*” (Whitney, 2023, p. 30).

Philosophical work on affective injustice examines how morally questionable norms, ideologies, attitudes, evaluative dispositions, or behaviors of those in power shape how those who lack such power experience themselves and their environment to their systematic disadvantage. Below, we add another example to the growing list of potential examples by arguing that philosophical work on affective injustice can offer a valuable conceptual toolkit for framing and understanding the emotional costs of being otherized in virtue of one’s asexuality.

Our lives are “shaped” by our affective engagements in specific social, historical, cultural, technological, and political spaces that can be understood as unfolding through different “stages” (e.g., Cherry, 2019; Pismenny et al., 2024): others may *expect* us to have certain feelings, *elicit* them in us, *interpret* them, *respond* to them, or *benefit* from them. Harm can occur at any point in this process, for instance when others interpret our justified anger as

hypocritical jealousy, respond tactlessly to our grief, or benefit from our despair. While such harm does not automatically or necessarily qualify as an injustice, it seemingly does so when certain individuals are systematically marked for it as a consequence of arbitrary discriminatory and oppressive otherizing structures. We contend that allonormativity is for Aces pretty much what patriarchy is for feminists, racism for Black people, and heteronormativity and gender binarism for queer individuals (Przybylo, 2011, p. 446). To make this case, we consider different manifestations of affective injustice discussed in the literature and add evidence that Aces can and arguably often do have to cope with exactly analogous challenges at each of the five “stages” mentioned above—cases where their affective experiences are not just harmed, but *policed*, *elicited*, *distorted*, *dismissed*, or *exploited* in morally wrong ways.

(1) *Affective policing*. Others establish and impose expectations about how we are supposed to feel and express our feelings. One form of such *affective policing* that has recently been discussed under the heading of “affective injustice” is what Archer and Matheson (2023) call “emotional imperialism.” This happens when a dominant culture uses its power to impose its emotional regimes onto others, “whilst at the same time marking out the other culture’s emotional norms and standards as deviant and inferior” (Archer & Matheson, 2022, p. 771). Another form is what Cherry (2023) calls “affective stereotyping.” Affective stereotypes are widely accepted generalizations about how individuals, *qua* members of certain groups, experience and express their emotions—women are presumed to be perpetually cheerful, approachable, and accommodating, men strong, stoic, and resilient. Such stereotyping has been deemed unjust. Cherry (2023), for instance, argues that the “Angry Black Woman” stereotype not only re-frames legitimate grievances as unfounded ill temper or hostility, but also perpetuates ideals of white supremacy by placing Black women in an affective double bind where they must choose between reinforcing the stereotype by reacting as expected, or suppressing their valid feelings by refusing to conform.

Both affective imperialism and stereotyping are an issue for Aces. Historically, many Indigenous communities, such as the *nádleehi* among the Navajo and the *winkte* among the Lakota, recognized and embraced non-binary and gender-fluid roles who were not necessarily expected to engage in sexual relationships

but could pursue alternative forms of human intimacy (e.g., Jacobs et al., 1997). Some of them were considered asexual (Smithers, 2022, p. xxii), and they were eradicated through the religiously motivated heteronormative emotion regimes imported and violently imposed by white settler imperialism (e.g., Carter, 2008). Currently, Aces are often perceived as not being “capable of the same levels of emotion as allosexual people” (Miller et al., 2022, p. 24) and as lacking the capacity to experience “warm feelings towards others” (Pacho, 2017, p. 118). By fostering the expectation that all Aces are *per se* “cold” (Zivony & Reggev, 2023, p. 2220) and “unfeeling” (Kelleher & Murphy, 2024, p. 883), such stereotyping wrongs Aces in much the same way the Angry Black Woman stereotype wrongs Black women. On the one hand, it undermines the affective credibility and autonomy of romantic Aces who do feel a longing for a partner. On the other hand, it puts aromantic Aces in an affective double bind where they can avoid reinforcing the general allosexual Ace stereotype only by suppressing or concealing their own true feelings.

(2) *Affective extraction.* Stereotypes not only shape *passive* expectations but can also create a sense of entitlement to *actively elicit* norm-conforming affective responses. For instance, when angry women are told to “smile” or “relax,” this has been said to function as a form of *affective extraction* that invokes a patriarchal stereotype to undermine the authenticity of their feelings and compromise their affective autonomy (Cherry, 2019). Affective extraction is also an issue for Aces. Annoying allosexual stereotypes such as “you’ll grow out of it” (Kelleher & Murphy, 2024, p. 880), “you don’t know what you’re missing” (Cuthbert, 2022, p. 847), or “don’t you want to find *the one*” (Young, 2023, p. 13) seek to align their affectivity with the expectation that desiring and having sex are inherently enjoyable and healthy aspects of a “natural” mature life. Just as telling an angry woman to “smile” or “relax,” such comments create social pressure that unduly questions the authenticity of Aces’ feelings and compromises their affective autonomy.

(3) *Affective distortion.* Next, others must *interpret* our behavior to infer how we feel. Such interpretations can systematically distort how individuals experience themselves and the world as a result of structural imbalances in social power. One form of such *affective distortion* that has recently been discussed under the heading of “affective injustice” is what has been dubbed “affective gaslighting” (e.g., Oliver 2022, 2023; Pismenny et al., 2024; Whitney,

2023), a kind of psychological manipulation aimed at actively making someone question the validity, origins, or adequacy of their feelings. Another, related, form of affective distortion is the kind of “affective pathologizing” (Eickers, 2023) that occurs when a person’s feelings are routinely dismissed as symptoms of a medical condition rather than recognized as valid experiences.

Again, both affective gaslighting and pathologizing are an issue for Aces. The powerful cultural narrative that sells sexual activity as an inherently desirable “happy object” (Peters, 2022) can instill self-doubt in those who do not feel what society’s “script of happiness” (Ahmed, 2010, p. 62) dictates they ought to feel (e.g., MacNeela & Murphy, 2015). Such “self-questioning” (Kelleher et al., 2023, p. 55) can make Aces internalize the message that having no sexual relationship equates to unhappiness, causing them to mistrust their own more authentic and authoritative assessment and triggering feelings of “shame” (Vares, 2022, p. 778) or even “self-hatred” (Chasin, 2015, p. 174)—especially if their lived experiences are pathologized by health professionals (e.g., Flanagan & Peters, 2020; Schneckenburger et al., 2024) who view asexuality as a symptom of trauma, hormonal imbalance, or mental illness that “needs fixing” (Decker, 2015, p. 10). As with other “psychiatrized” groups, subjecting Aces to the sanist prejudice that they are “deserving of grief, care, support, and rehabilitation” (Owen, 2014, p. 126) not only denies them the ability articulate their authentic experiences, but also systematically discourages them from interpreting their affective lives in ways that affirm their identity in positive terms.

(4) *Affective dismissal.* Next, even when others correctly interpret our feelings, they can harm us through “affective invalidation,” i.e., by disregarding, trivializing, or belittling them without engaging with their substantive content. An extreme version is the kind of “affective silencing” (e.g., Slovin, 2022) that makes an active effort to prevent a person’s experiences from being recognized by others. Another version is the kind of “affective intrusion” that occurs when marginalized individuals are treated as “open persons” (Goffman, 1963, p. 126) that are unworthy of social respect and free to being ogled, remarked upon, or approached at will—this otherizes them as “alien” objects of fascination or scrutiny, while at the same time rendering invisible their identities and lived experiences, denying them the respect and dignity they deserve.

Again, affective invalidation, silencing, and intrusion are an issue for Aces. Allosexuals' expectation that "everyone wants to f*** someone" (Robbins et al., 2016, p. 765) encourages them to invalidate Aces' expressions of lack of sexual attraction by fabricating more legible alternative explanations: they tell Aces that they are "late bloomers" (Cuthbert, 2022, p. 842), "going through a phase" (Mitchell & Hunnicutt, 2019, p. 518), or "confused" (Brunning & McKeever, 2021, p. 14), or that they "will want sex once they get into it" (Cuthbert, 2022, p. 844). Comments like these undermine Aces' autonomy by denying them the capacity to understand and articulate their own affective experiences. As one Ace described their frustration when their friend smothered their orientation by insisting they *must* feel sexual attraction because they are, after all, "not a tree": "It was like learning that I had brown eyes, but when I mentioned it people just said that I really had blue eyes. I just couldn't believe that they didn't believe me and worse, that they thought they knew how I felt better than I did" (Robbins et al., 2016, p. 765). When that person shares their "you're not a tree" attitude with others, they not only invalidate Aces' authority over their own experience, but can also be seen as actively silencing them. Such aphobic attitudes become even more harmful when they are propagated by apparent authorities. As long as a significant percentage of medical professionals tells Aces that they are "making up the asexuality community" (Flanagan & Peters, 2020, p., 1636), that their partner is "just a bad lover" (Parshall, 2024), or that they have "to get over [their] 'fear of sex' and [...] a medication to increase [their] libido" (ibid.), allonormative biases are reinforced under the authority of the white coats, silencing Aces with a veneer of scientific legitimacy (e.g., Flanagan & Peters, 2020; Robbins et al., 2016). Aces also often explicitly complain about being "asked all sorts of invasive questions" (Kelleher and Murphy, 2024, p. 880) by people who regard their lack of sexual attraction as a "curiosity" (Scott & Dawson, 2015, p. 4) or "rarity" (Decker 2015, p. 46), treating them as a "spectacle of fascination" (Ceranowski, 2014, p. 288).

(5) *Affective exploitation*. Lastly, others can harm us by making use of the ways we affectively engage with the world to their own advantage. One form of such *affective exploitation* that has recently been discussed under the heading of "affective injustice" is what Pismenny et al. (2024) call "double burdening." When the white woman told Lorde, "Tell me how you feel but don't say it too

harshly or I cannot hear you” (1984, p. 125), she left Lorde no choice but to contribute to her own oppression, by either confirming the misogynoir Angry Black Woman stereotype or else taking on the coerced labor of explaining the systemic injustice Black women face in order to soothe the emotional impact of her experience precisely on and for those responsible for it. Another form of affective exploitation is what has been called “affective appropriation” (e.g., Agathangelou, 2019). When, for instance, a white woman “appropriates the Lordean rage [...] in a way that fails to address the sufferings of nonwhites” (Cherry, 2021, p. 129f.) to pursue her own agenda, for instance to bolster her credibility as an activist, she unjustly exploits others’, in this case a suppressed Black woman’s intersectional, affective labor.

Again, affective double burdening and appropriation are an issue for Aces. Many Aces face negative reactions when coming out to family, friends, or partners (e.g., Higginbottom, 2024; Robbins et al., 2016). Among other things, this shoulders them with the double burden of navigating and managing others’ misunderstandings and hostilities, while simultaneously also investing “a lot more emotional work” (Kelleher and Murphy, 2024, p. 878) to educate them about asexuality and help them come to terms with it. One participant in a recent study recounts telling her husband: “you know that I finally get it, here I am, but don’t be afraid, we are going to figure this out. I am with you, and this is not horrifying. It is something we have to figure out how to coalesce” (Glass, 2022, p. 352). A similar double burdening arises when Aces have to endure the inner conflict of engaging in sexual activities with allosexual partners —often accompanied by emotional or physical discomfort —while simultaneously fearing that they are a drain to them and unable to provide them with what they deserve. Popular media have appropriated asexuality for entertainment and commercial gain. In Season 8, Episode 9 of the immensely successful drama *House MD*, for instance, House encounters a married couple who both identify as asexual. House dismisses this possibility outright with the slur that “the only people who don’t want [sex] are either sick, dead, or lying” (Przybylo, 2019, p. 17). The narrative vindicates his allosexual notion of normality: the husband is revealed to have a brain tumor that impairs his sex drive and his wife just feigned asexuality out of fear of losing him. Similarly, Sheldon Cooper, one of the protagonists of the blockbuster sitcom *The Big Bang Theory*, who is identified as asexual in the

early seasons, embodies virtually all negative Ace stereotypes: he is “socially awkward” (Loftis, 2018, p. 109), “weird” (Fedtke, 2022, p. 370), and “devoid of any emotion or feeling” (ibid., p. 376) to the point of appearing “less than human” (Loftis, 2018, p. 112)—and it is precisely these stereotypes that account for much of the unique humor driving the series’ colossal success. Even worse, allonormativity is, once again, vindicated, as Sheldon’s asexuality turns out to be “just a phase”: He ends up “healed” through his engagement in a heteronormative sexual and romantic relationship (Fedtke, 2022, p. 379).

Above, we have reviewed various examples where the affective life of certain groups of oppressed individuals has been said to be policed, elicited, distorted, dismissed, and exploited in ways that should count as outright injustice: We have argued that Aces are otherized as “affect aliens” by an allonormative society in exactly similar ways. Hence, the affective harm they experience also ought not be seen as merely an unfortunate personal mishap, but as an affective injustice. If Young (1990) is right with her structural account of oppression as a matter of “the vast and deep injustices some groups suffer as a consequence of often unconscious assumptions and reactions of well-meaning people in ordinary interactions, media and cultural stereotypes, and structural features of bureaucratic hierarchies and market mechanisms” (p. 41), then asexual-identified individuals are, like so many others, oppressed—oppressed by “the normal processes of everyday life” (ibid.).

[1] By “otherizing” we refer to the processes by which marginalized individuals or groups are “marked” (Haslanger, 2012) on the basis of race, gender, disabilities etc. as deviant or unnatural. Aces are otherized because they do not experience sexual attraction and intimacy as “happy objects” (Peters, 2022) of their intrinsic yearning. This makes them what Ahmed (2010) calls “affect aliens”—they experience a “gap between the promise of happiness and how [they] are affected by objects that promise happiness” (ibid., p. 42).

References

Agathangelou, Anna M. (2019) “Empire of affects: Speculation and wagers in IR’s affective turn”, in: Eric Van Rythoven and Mira Sucharov (eds.) *Methodology and Emotion in International Relations: Parsing the Passions*, London: Routledge, 205–222.

Ahmed, Sara (2010) *The Promise of Happiness*, Durham: Duke University Press.

- Archer, Alfred; Matheson, Benjamin (2022) "Commemoration and emotional imperialism", *Journal of Applied Philosophy* 39(5), 761–777.
- Archer, Alfred; Matheson, Benjamin (2023) "Emotional imperialism", *Philosophical Topics* 51(1), 7–25.
- Brandley, ben; Spencer, Leland G. (2023) "Rhetorics of allonormativity: The case of asexual Latter-day Saints", *Southern Communication Journal* 88(1), 1–15.
- Brunning, Luke; McKeever, Natasha (2021) "Asexuality", *Journal of Applied Philosophy* 38(3), 497–517.
- Carter, Sarah (2008) *The Importance of Being Monogamous: Marriage and Nation Building in Western Canada in 1915*, Athabasca: Athabasca University Press.
- Ceranowski, Karli June (2014) "Spectacular asexuals: Media visibility and cultural fetish" in: Karli June Ceranowski and Megan Milks (eds.) *Asexualities: Feminist and queer perspectives*, London: Routledge, 139–161.
- Chasin, C. J. DeLuzio (2015) "Making sense in and of the asexual community: Navigating relationships and identities in a context of resistance", *Journal of Community & Applied Social Psychology* 25(2), 167–180.
- Cherry, Myisha (2019) "Gendered failures in extrinsic emotional regulation; Or, why telling a woman to 'relax' or a young boy to 'stop crying like a girl' is not a good idea", *Philosophical Topics* 47(2), 95–112.
- Cherry, Myisha (2021) *The Case for Rage. Why Anger Is Essential to Anti-racist Struggle*, Oxford: Oxford University Press.
- Cherry, Myisha (2023) "Affective stereotype threat as affective injustice", *Philosophical Topics* 51(1), 135–147.
- Cuthbert, Karen (2022) "Asexuality and epistemic injustice: A gendered perspective", *Journal of Gender Studies* 31(7), 840–851.
- Decker, Julie Sondra (2015) *The Invisible Orientation: An Introduction to Asexuality*, New York: Carrell Books.
- Eickers, Gen (2023) "Pathologizing disabled and trans identities: How emotions become marginalized", in Shelley Lynn Tremain (ed.) *The Bloomsbury Guide to Philosophy of Disability*, New York: Bloomsbury, 360–379.
- Fedtke, Jana (2022) "The erasure of sexuality: Sheldon's masculinity in *The Big Bang Theory*", in Lydia R. Cooper (ed.) *The Routledge Companion to Masculinity in American Literature and Culture*, London: Routledge, 368–380.

Flanagan, Shelby K; Peters, Heather J. (2020) "Asexual-identified adults: Interactions with health-care practitioners", *Archives of Sexual Behavior* 49, 1631–1643.

Fricker, Miranda (2017) *Epistemic Injustice: Power and the Ethics of Knowing*, Oxford: Oxford University Press.

Frye, Marilyn (1983) "A note on anger", in: *Politics of Reality: Essays in Feminist Theory*, Berkeley: Crossing Press, 84–94.

Gallegos, Francisco (2022) "Affective injustice and fundamental affective goods", *Journal of Social Philosophy* 53(2), 185–201.

Glass, Valerie Q. (2022) "Queering relationships: Exploring phenomena of asexual identified persons in relationships", *Contemporary Family Therapy* 44, 344–359.

Haslanger, Sally (2012) *Resisting Reality: Social Construction and Social Critique*, Oxford: Oxford University Press.

Higginbottom, Brooke (2024) "The nuances of intimacy: Asexual perspectives and experiences with dating and relationships", *Archives of Sexual Behavior* 53, 1899–1914.

Jacobs, Sue-Ellen; Thomas, Wesley; Lang, Sabine (eds.) (1997) *Two-spirit people: Native American gender identity, sexuality, and spirit*, Chicago: University of Illinois Press.

Kelleher, Sinéad; Murphy, Mike (2024) "Asexual identity development and internalisation: A thematic analysis", *Sexual and Relationship Therapy* 39(3), 865–893.

Kelleher, Sinéad; Murphy, Mike; Su, Xin (2023) "Asexual identity development and internalisation: A scoping review of quantitative and qualitative evidence", *Psychology & Sexuality* 14(1), 45–72. <https://doi.org/10.1080/19419899.2022.2057867>

Loftis, Sonya Freeman (2018) "A kiss on the train: Autism, asexuality and the conventions of romantic comedy", in: J. L. Schatz and Amber E. George (Eds.), *The image of disability: Essays on Media Representations*, Jefferson: McFarland & Company, 107–120.

Lorde, Audre (1984) "The uses of anger: Women responding to racism", in: *Sister Outsider*, Berkeley: Ten Speed Press, 124–133.

MacNeela, Pádraig/Murphy, Aisling (2015) "Freedom, invisibility, and community: A qualitative study of self-identification with asexuality", *Archives of Sexual Behavior* 44, 799–812.

Miller, Hilary; Welland, Louisa; McCook, Sarah; Giunta, Katherine (2022) "Wellbeing beyond binaries: A qualitative study of wellbeing in bisexual+ youth", *Journal of Bisexuality* 22(3), 385–428.

Mitchell, Heather; Hunnicutt, Gwen (2019) "Challenging accepted scripts of sexual 'normality': Asexual narratives of non-normative identity and experience", *Sexuality & Culture* 23, 507–524.

Oliver, K. (2022) "Gaslighting. Pathologies of recognition and the colonisation of psychic space", in: Paul Giladi and Nicola McMillan (eds.) *Epistemic Injustice and the Philosophy of Recognition*. London: Routledge, 114–137.

Oliver, Kelly (2023) "Affective gaslighting and rape culture: Compounding the violence of sexual assault", *Studies in Gender and Sexuality* 24(4), 312–321.

Owen, Ianna Hawkins (2014) "On the racialization of asexuality", in: Karli June Cerankowski and Megan Milks (eds.) *Asexualities: Feminist and Queer Perspectives*, London: Routledge, 119–135.

Pacho, Agata (2017) "'Cake is better than sex'—AVEN and asexuality", in: Paul G Nixon and Isabel Düsterhöft (eds.) *Sex in the Digital Age*, London: Routledge, 113–123.

Parshall, Allison (2024) "Life beyond sex: A lack of sexual attraction was long seen as a problem that needed to be fixed. New research on asexuality is expanding ideas of love and intimacy", *Scientific American* 330(1), 68.

Peters, Charnell (2022) "Asexuality, affect aliens, and digital affect cultures: Relationality with the happy objects of sexual and romantic relationships", *Western Journal of Communication* 86(5), 585–600.

Pismenny, Arina; Eickers, Gen; Prinz, Jesse (2024) "Emotional injustice", *Ergo: An Open Access Journal of Philosophy* 11(6), 150–176.

Przybylo, Ela (2011) "Crisis and safety: The asexual in sexusociety", *Sexualities* 14, 444–461.

Przybylo, Ela (2019) *Asexual Erotics. Intimate Readings of Compulsory Sexuality*. Columbus, Ohio State University Press.

Rawls, John (1971) *A Theory of Justice*, New York: Belknap Press.

Robbins, Nicolette K.; Low, Kathryn G.; Query, Anna N. (2016) "A qualitative exploration of the 'Coming out' process for asexual individuals", *Archives of Sexual Behavior* 45, 751–760.

Schneckenburger, Stella A.; Tam, Michelle W. Y.; Ross, Lori E. (2024) "Asexual competent practices in healthcare: A narrative review", *Journal of Gay & Lesbian Mental Health* 28(3), 314–334.

Scott, Susie; Dawson, Matt (2015) "Rethinking asexuality: A symbolic interactionist account", *Sexualities* 18(1–2), 3–19.

Slovin, L. (2022) "'What matters what makes us gay?' Affective failure in a social justice classroom", *Sex Education* 23(2), 220–227.

Smithers, Gregory (2022) *Reclaiming Two-spirits: Sexuality, Spiritual Renewal & Sovereignty in Native America*, Boston: Beacon Press.

- Srinivasan, Amia (2018) "The aptness of anger", *Journal of Political Philosophy* 26(2), 123–144.
- Vares, Tina (2022) "Asexuals negotiate the 'onslaught of the heteronormative'", *Sexualities* 25(5–6), 767–784.
- Vasallo, Brigitte (2019) "Monogamous mind, polyamorous terror", *Sociological Research Online* 24(4), 680–690.
- Vespermann, Daniel; Tirkkonen, Sanna Karoliina (2025) "Existential injustice in phenomenological psychopathology", *Philosophical Psychology* 38(1), 209–245.
- Whitney, Shilo (2018) "Affective intentionality and affective injustice: Merleau-Ponty and Fanon on the body schema as a theory of affect", *Southern Journal of Philosophy* 56(4), 488–515.
- Whitney, Shiloh (2023) "Anger gaslighting and affective injustice", *Philosophical Topics* 51(1), 27–62.
- Young, Eris (2023) *Ace Voices. What It Means to Be Asexual, Aromantic, Demi or Grey-ace*, London: Jessica Kingsley Publishers.
- Young, Iris Marion (1990) *Justice and the Politics of Difference*, Cambridge: Princeton University Press.
- Zivony, Alon; Reggev, Niv (2023) "Beliefs about the inevitability of sexual attraction predict stereotypes about asexuality", *Archives of Sexual Behavior* 52, 2215–2228.

Acknowledgement

This project was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) project number GRK 274877981 (DFG Research Training Group Situated Cognition).

The Material Theory of Values in Science

Kelli Barr (Pittsburgh, Pennsylvania, United States)

Abstract

How are we to understand situations where science fails on its own terms? Scientists have blamed perverse incentives for systematic epistemic failures like non-replicability and publication bias, but the exact relationship remains an open question. Let's assume they are right to blame the (social) system. This paper presents a novel framework for understanding how features of the social organization of science are implicated in collective epistemic failures: the material theory of values in science (MTV). This project is inspired by and follows in the tradition of feminist philosophers of science who have called attention to the need for explanations of systemic, specifically antifeminist, biases in science and for embodied models of scientists as epistemic agents. In the first part, I discuss the replication crisis as involving a particular type of collective action problem: a no-win standoff. The next part introduces the MTV and the explanation it supplies for this phenomenon. In the third section, the MTV is compared to several alternative explanatory strategies, including from the contemporary literature values in science and social epistemology, specifically agent-based computational models. I argue for why my approach is preferable and describe an important revision it entails for the general Mertonian sociological picture invoked in discussions of incentives in science. By way of conclusion, I summarize the philosophical task of the MTV, which is to take seriously the sociological and material conditions constituting scientists' epistemic agency.

Since revelations of the replication/reproducibility crisis in many research specialties, scientists are in a situation characterized not only by publish-or-perish, but also by the possibility that their findings might not be valid, secure, or even true despite their best efforts. Before this, publish-or-perish already called to mind the classic movie trope of a standoff: like adversaries circling each other with weapons drawn, scientists competed for who would be the first to publish the next discovery by continuously escalating the pace and output of publications. This was a game one could win, in principle, by beating others to the publication punch. However, the replication crisis has turned this situation into one where there is no path to victory for individual researchers, since systemic biases and failures manifest in aggregate activity (OSC 2015; Nosek et al. 2022). It is a no-win standoff; it can only be resolved by mutual destruction (i.e., dropping out of the profession) or mutually withdrawing. One cannot unilaterally withdraw, for example, by running a replication or reproduction study to further validate one's own or others' existing findings without taking oneself out of the running (at least temporarily) for being the first to publish a discovery. But convincing others to stand down collectively is terribly difficult. Most who remain in the profession remain in the standoff.

Scientists have tried to bargain with one another for mutual withdrawal. Some have tried to target incentives directly, for example by pursuing special funding for large-scale replication/reproduction projects or starting special venues for publishing such projects, in the hopes of shifting collective activity toward more validation of findings (see Korbmacher 2023 for a more comprehensive review of such efforts). Others have thrown their efforts into developing Metascience as a research specialty for scientific work about science itself, on the hopes of redeeming the enterprise with its own best tools. One persistent problem has been how to understand claims that systemic epistemic failures and biases are due, in part, to features of the social organization of science. Perverse incentives are consistently cited in discussions of the replication crisis. To take one hypothesis as an example: publish-or-perish might drive researchers to cut corners using questionable research practices. Many scientists seem convinced that there is some connection, but the exact relationship remains an open question.

Let's assume scientists are right that (at least some) epistemic failures are due to social factors. What can we learn from this? First, we can learn that the situation is not best explained by what scientists believe, or prefer, or know, or have attitudes about. A no-win standoff is a standoff precisely in virtue of the weapons each party points at one another, fingers poised ready on the triggers. But incentives aren't guns to peoples' heads... right? When it comes to the replication crisis, scientists talk as though they are compulsory. This is more than just a way of speaking. There are material consequences to withdrawing from the standoff. At stake is continued access to the resources necessary to keep doing research at all, such as access to funding, labor, and equipment – meaning, the possibility of continuing to have a career. There are many ways to think about materiality, notably stemming from Marxist intellectual traditions. My concern here is primarily with the means of scientific production listed above, but also secondarily with matters of socially situated agency and embodiment (Haraway 1988; Rouse 2002), with social ontology (cf. Haslanger 2022), and with explanations of systemic biases in critiques of science (Wylie 1996; Lloyd 2006). Second, we can learn about a broader field of values that structure scientific work by paying attention to the material dimensions of the social organization of science.

Statement of the Theory

The material theory of values in science (MTV) holds that there is a class of values that shape scientific practices independently of individual scientists' beliefs, preferences, knowledge, attitudes, or other cognitive phenomena. I call these *background values*, which are constituent parts of the *institutions* that structure scientific practices and are evident in scientists' interactions with their *material conditions*: the organization, allocation, and distribution of money, labor, and equipment. Scientists' actions, then, aren't explained by cognitive phenomena alone, or perhaps even primarily, but rather by their enacting the institutions that structure their work and the background values made material by those institutions.

MTV has two core tenets. First: values are materialized in institutions. To enact an institution is to enact a set of values. This has the consequence that values are not the kinds of things that one holds, like attitudes, or that one takes for granted, like background assumptions. Like the institutions which give them concrete expression, they are shared, collective things. By institutions, I mean those stable but flexible patterns of conventional conduct that structure individual and collective social behavior. This sense of 'institution' is inspired by the Deweyan pragmatist school of sociology based at the University of Chicago, which was most active in the middle and late part of the 20th century (see Hughes 1942; Gerson, unpublished manuscript). My use is more similar to what others characterize as 'social practices' (e.g., Rouse 2003) than to typical conceptions of institutions, especially among philosophers (see, for example, Turner 1994; Guala 2016). Hamburgers, putting on socks and shoes, the time of day when espresso is drunk, the particular layout of a biological laboratory's storage cabinets, and funding schemes; all these count as institutions and exist along interrelated scales of formality, longevity, complexity, and reach. The claim that values are constituents of institutions, then, entails that we have values as a result of the historical, ongoing societal organization of collective life, including the collective life of scientific communities. Values do not exist as free-floating, psychological, ethical/moral, or existential parameters. They are not intrinsic to human cognition or agency. To say that you value something is to claim membership in a community and to be (or to have been) shaped by that membership.

The second tenet of the MTV is that values have a material, rather than formal or psychological, relation to institutions. This contrasts with long running discussions in philosophy of science about the possible formal relations between values and, for example, theory choice or hypothesis acceptance (cf. Rudner 1953). Contemporary discussions of values in science motivated by the argument from inductive risk tend to treat values as decision criteria or parameters that stand in logical, justificatory, or causal relationships to individual decisions made by individual scientists. Values are taken to exist in the minds, heads, and decisions of individuals; for such cognitivist theories, the instantiation of values is logical or psychological. By contrast, I say that values exist in institutions; for material theories, the instantiation of values is sociological and intersubjective. This addresses a persistent problem in the values and science literature: how to understand the nature of the values we think are operative in science (Elliott and Korf 2024). Ward (2021) proposed an initial taxonomy: values can be reasons (for example, for actions or beliefs), or they can be causes (for example, as motivations). To this, the MTV adds that values can be constituent features, as in part-whole compositional relations, which locates the normativity of values in the institutions of which they are constituent parts rather than in individual psychological dynamics, decision-making, or utility maximizing.

Let's return briefly to the situation described at the outset, with the MTV in hand. Scientists identify publish-or-perish as a contributor to systemic epistemic biases like those toward publishing novel and positive findings. The idea is that incentives prioritizing novelty direct scientists to put their limited time and resources toward new discoveries instead of validating existing findings. According to our no-win standoff model, no one is willing to incur the risks of switching tracks from discovery to validation unless or until others do the same. Exactly how and why is this the case? The MTV directs us to examine the institutions implicated in publish-or-perish. For which scientific practices is productivity, in terms of high publication output, a constitutive value and how does its value relate to scientists' interactions with material conditions in which they work? We are led to observations of this kind: productivity is a constitutive value of such institutionalized practices as decisions about funding allocation, hiring, promotion, and tenure. That is, publications regulate access to the material goods and resources without which one cannot

continue to do research. One needs publications to get jobs, research funding, research labor, and other resources, like access to research libraries and laboratories, or other data collection equipment. To the extent that novel discoveries are a priority, in terms of it being much more difficult to get validation projects published, and to the extent that publications are required to continue on with a research career, publish-or-perish is a driver of publication bias. In this sense, producing epistemically sub-optimal outcomes is constitutive of the (current) material and social conditions of science, without anyone involved wanting this to be the case, or even trying to make it so.

Why Blame Institutions and Not Individuals?

One significant advantage the MTV has over alternatives lies in highlighting the dual economies in which publications are the coin of the realm. On the one hand, publications are symbolic currency in the reputational economy of credit (e.g., for discoveries) and prestige (e.g., honorary doctorates, exclusive society memberships). On the other hand, they are currency in the material economy that allocates financial, labor, and equipment resources. In this sense, the MTV motivates a revision to a common sociological picture in the background of discussions about the structure of scientific communities. Sociologists of science have long recognized the importance of credit allocation in incentivizing researchers to publish. Merton's (1957) priority rule formalized this relationship, and some philosophers of science have since championed it as key to achieving an efficient distribution of labor between scientists and between research programs (Kitcher 1999, Strevens 2011). Recent discussions are less sanguine about the efficiency of such a system (Romero 2017, Heesen 2018). But they do not dispute the basic features of the model: scientists pursue projects likely to garner recognition from their peers, in which case maximizing productivity is an advantageous strategy for maximizing recognition.

As the MTV shows, however, the focus on *what* scientists choose to publish belies institutional realities. Scientists cannot make a career from publishing replications and/or reproductions of others' findings. The relevant incentives are not isolated to institutions for scholarly communication, such as editorial policies, for-profit presses, and peer review processes. These institutions are entangled with others, such as research funding, equipment manufacturing

and distribution, organizational policies, hiring and promotion practices, and research assessment exercises. Scientists have a choice about what to publish but not *whether* to publish, and that's because the material resources needed to continue working as scientists are on the line. Thus, explanations based on the economy of recognition are incidental to explaining why it's a competitive strategy to maximize publications.

But, of course, scientists want to produce *true* findings (or accurate, or explanatorily adequate). Feminist philosophers of science, especially those concerned with the critique of science, have long been attentive to the ways in which decisions about what to pursue, and how, can matter to our epistemic appraisals of scientific results (Wylie 1996; Lloyd 2006). More recent models of decision making in inductive risk or practical underdetermination scenarios highlight value trade-offs with respect to relevant epistemic and non-epistemic risks, in the context of an individual scientist's decisions (Douglas 2000; Anderson 2004). This wedds well with a focus on the consequences of incentives: scientists must make choices viz-a-viz productivity norms. If they conform to productivity norms to avoid the risks of defying them, then we can explain their choices as a risk mitigation strategy, where values are traded off in light of assessing the relevant risks.

Yet, such an explanation is also only incidental. Again: there is no choice for scientists with respect to *whether* to produce papers. Interactions between the recognitional and material economies of science shape when, where, and how the value trade-offs depicted in inductive risk scenarios arise. This is another sense in which the material and social organization of science constitutively produces epistemically sub-optimal outcomes. MTV has the advantage of recognizing that a distinct *kind* of values, background values, is implicated here. The dominant view in the literature about values in science seems to be that values that can stand in causal or justificatory relations with epistemic judgments must be some kind of mental phenomenon in the mind of a scientist, like presuppositions, or preferences, or attitudes. By contrast, background values are constitutive of epistemic judgments in the sense that they structure which problem spaces are (materially) viable for exploration, when they are explored, and to what extent, since the practices by which these

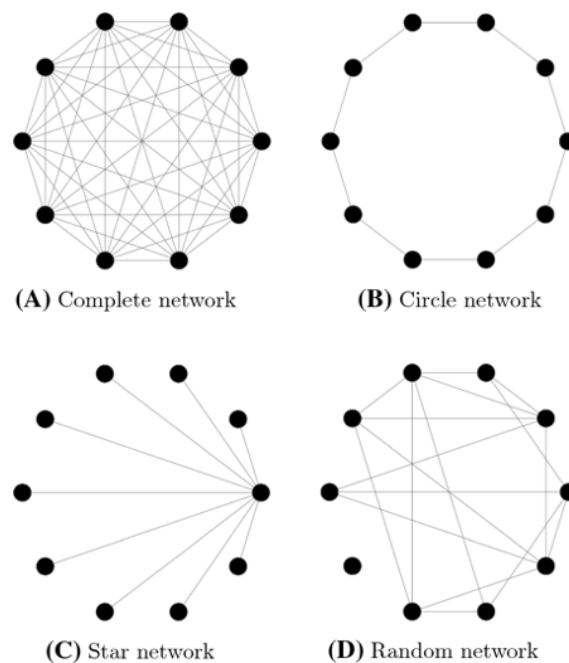
activities proceed are institutionalized practices. It is a distinct advantage of the MTV that it recognizes this broader field of value commitments at stake in science.

The individualistic and cognitivist models considered above are supplementary to explaining the epistemic implications of how scientific communities are organized socially. The MTV offers a more concrete and direct account of the relationship between publish-or-perish and epistemic failures. In addition to addressing open questions about the replication crisis and providing a needed revision to the Mertonian sociological picture, it also has important social epistemological implications, specifically for recent attempts to model scientific communities as networks of Bayesian agents (Zollman 2010, Šešelja 2021). Social epistemologists have been transparent about agent-based network models being highly idealized. It remains a very open question how to de-idealize them (following Knuuttila and Morgan 2019) in a way that would connect concretely with what we already know about the sociological features of scientific communities and practices (cf. Gerson 2013). Individual scientists participate in the institutions that scaffold scientific practices (e.g., peer-review of publications), and those institutions structure how scientists group together (e.g., into specialties). But institutions are not reducible either to groupings of agents or to other organizing structures like communication networks. This suggests that a minimal formal model of a scientific community is at least a three-body problem: both networks and institutions structure groupings of agents.

Implications for Theorizing about Epistemic Agents

The philosophical task of the MTV is to take seriously the sociological and material conditions constituting scientists' epistemic agency. There is risk in failing to take up this task: a risk of committing ourselves to continue theorizing about inhuman entities and pursuing non-empirical studies of science. Our usual theoretical resources for modeling scientific decision-making consist of cognitive models depicting individual agents as bloodless bundles of motivations, beliefs, knowledge, preferences, and other mental states (though, usually not simultaneously). This is especially the case in social epistemology, and especially in projects modeling scientists as simple Bayesian algorithms (i.e., Bayes' theorem plus a fixed updating criterion) that exchange information with other nearby algorithms. What kind of 'agents' are these?

They are not human and barely machine; they are nameless, faceless, and nearly limitless in time and space; and their relations with one another are through anonymous, unmediated channels of exchange of a single type of information. These are otherworldly entities, not even brains in vats. They might as well be Bayesian angels. And the risk is that inquiries like this will amount to nothing more than the new celestial mechanics.



(Figure 1 from Mohseni and Williams 2021)

Perhaps in some future, an indeterminate amount of time away, these simulations will make contact with the corporeal circumstances of embodied inquirers. What might this look like, and how long are we willing to wait for it? Imagine a group of people standing in a circle facing outward. Much available sensory information has been blocked. They wear blindfolds, noise-cancelling headphones, mittens, and thick clothing so that they cannot see, hear, feel, or otherwise be aware of there being anyone else around them. They are asked to train their minds on visualizing the value of a single probabilistic credence. After a period of time, a member is chosen at random to say their credence aloud and the utterance is piped into the headphones of select participants, after which everyone is asked to calculate a new credence to be displayed on

their mental billboard according to their pre-assigned updating criterion. In what sense are these individuals *agents*, let alone scientists? In what sense are they *connected* to each other in a *network*? In what sense is this *communication*? This might seem a rather brute way to depict a potential de-idealization. But one has to start somewhere, since no one has yet tried (as of my writing this), despite claims about the policy relevance of simulation results (cf. O'Connor 2019).

Feminist philosophies in the 20th and 21st centuries have consistently pushed for the social and material re-embodiment of reasoning agents, such as in feminist standpoint theories (cf. Harding 1991). The MTV follows this tradition by assuming as a starting point for theorizing that any epistemic agent is constituted, in part, by the institutions in which they are embedded, such that they have commitments (of various kinds) to the values constitutive of institutionalized practices. These agents are not only flesh and blood; they have histories, rich social relations with one another, and something important at stake in those social relations (i.e., their livelihoods). MTV points us toward the social features that organize collective life in science and the material conditions that constrain scientific work. It aims to show how we might understand scientific epistemology from that perspective: having an integrated understanding of the social, political, and material economy of scientific communities, from which epistemological questions and theories flow. The hope is that this more sociologically robust depiction of scientific communities opens up new territory for exploring the relationship between scientific practices and institutions, for understanding how power operates in and through scientific work, and for conversations between philosophy of science, science policy, and political economy.

References

- Alexandrova, Anna (2018) "Can the Science of Well-Being Be Objective?" *The British Journal for the Philosophy of Science* 69, 421-445.
- Anderson, Elizabeth (2004) "Uses of Value Judgments in Science: A General Argument, with Lessons from a Case Study of Feminist Research on Divorce", *Hypatia* 19, 1-24.
- Douglas, Heather (2000) "Inductive Risk and Values in Science", *Philosophy of Science* 67, 559-579.

Elliott, Kevin C.; Rebecca Korf (2024) "Values in Science: What Are Values, Anyway?" *European Journal for Philosophy of Science* 14, 1-24.

Gerson, Elihu M. "Institutions and Repertoires" [unpublished manuscript]

Gerson, Elihu M. (2013) "Integration of Specialties: An Institutional and Organizational View" *Studies in History & Philosophy of Science Part C* 44, 515-524.

Guala, Francesco (2016) *Understanding Institutions: The Science and Philosophy of Living Together*. Princeton: Princeton University Press.

Haraway, Donna (1988) "Situated Knowledges: The Science Question in Feminism and the Privilege of Partial Perspective", *Feminist Studies* 14, 575-599.

Harding, Sandra (1991) *Who's Science? Who's Knowledge? Thinking from Womens' Lives*. Ithaca: Cornell University Press.

Haslanger, Sally (2022) "Failures of Methodological Individualism: The Materiality of Social Systems", *Journal of Social Philosophy* 53, 512-534.

Heesen, Remco (2018) "Why the Reward Structure of Science Makes Reproducibility Problems Inevitable", *Journal of Philosophy* 115, 661-674.

Hughes, Everett C. (1942) "The Study of Institutions", *Social Forces* 20, 307-310.

Kitcher, Philip (1990) "The Division of Cognitive Labor", *Journal of Philosophy* 87, 5-22.

Knuuttila, Tarja; Mary S. Morgan (2019) "Deidealization: No Easy Reversals", *Philosophy of Science* 86, 641-661.

Korbmacher, Max et al. (2023) "The Replication Crisis Has Let to Positive Structural, Procedural, and Community Change" *Communications Psychology* 1, 1-13.

Lloyd, Elizabeth (2006) *The Case of the Female Orgasm: Bias in the Science of Evolution*. Cambridge: Harvard University Press.

Merton, Robert K. (1957) "Priorities in Scientific Discovery: A Chapter in the Sociology of Science", *American Sociological Review* 22, 635-659.

Mohseni, Aydin; Cole R. Williams (2021) "Truth and Conformity on Networks", *Erkenntnis* 86, 1509-1530.

Nosek, Brian et al. (2022) "Replicability, Robustness, and Reproducibility in Psychological Science", *Annual Review of Psychology* 73, 719-748.

O'Connor, Cailin (2019) *The Origins of Unfairness: Social Categories and Cultural Evolution*. Oxford: Oxford University Press.

- Open Science Collaboration (2015) "Estimating the Reproducibility of Psychological Science," *Science* 349, 4716.
- Romero, Felipe (2017) "Novelty vs. Replicability: Virtues and Vices in the Reward System of Science", *Philosophy of Science* 84, 1031-1043.
- Rouse, Joseph (2002) *How Scientific Practices Matter: Reclaiming Philosophical Naturalism*. Chicago: University of Chicago Press.
- Rudner, Richard (1953) "The Scientist Qua Scientist Makes Value Judgments" *Philosophy of Science* 20, 1–6.
- Šešelja, Dunja (2021) "Exploring Scientific Inquiry via Agent-Based Modeling" *Perspectives on Science* 29, 537-557.
- Strevens, Michael (2011) "The Role of the Priority Rule in Science", *Journal of Philosophy* 100, 55-79.
- Turner, Stephen (1994) *The Social Theory of Practices: Tradition, Tacit Knowledge, and Presuppositions*. Cambridge: Polity Press.
- Ward, Zina (2021) "On Value-Laden Science" *Studies in History and Philosophy of Science Part A* 85, 54-62.
- Wylie, Alison (1996) "The Constitution of Archaeological Evidence: Gender Politics and Science" in: Peter Galison and David J. Stump (eds.) *The Disunity of Science*, Stanford: Stanford University Press, 311-343.
- Zollman, Kevin J. S. (2010) "The Epistemic Benefits of Transient Diversity" *Erkenntnis* 72, 17-35.

Susanne K. Langer, The Vienna Circle, and *The Practice of Philosophy*

Mark Nader Basafa (Vienna, Austria)

Abstract

At the center of the genesis of Thomas Kuhn's landmark ideas in *The Structure of Scientific Revolutions* (1962) is the tug of war between Arthur Lovejoy and Bertrand Russell where the former emphasized historical treatment of science and the latter represented a purely logical view of the scientific enterprise. Within this context, I trace Kuhn's point of contact with Susanne K. Langer in the archival materials at MIT, to find that in one of Kuhn's undergraduate papers written for H. M. Sheffer's course, an acknowledgement of Langer: "I have derived considerable assistance from the clear exposition in Susanne K. Langer's *Introduction to Symbolic Logic*" (TSK-MIT Box 1, Folder 3). Another point of contact with Langer is Kuhn's reading of Langer's *Philosophy in a New Key* along with other books by April 1949 (Reisch 2019, 160). The book was, among others, inspired by Ernst Cassirer "that pioneer in the philosophy of symbolism" (Langer 1948, preface). Finally, this paper attributes the genesis of *Structure* with the historical thinking of figures such as Alexandre Koyre and Arthur Lovejoy and albeit indirectly with the historicism of Ernst Cassirer. Langer's departure from philosophical logic can be attributed to her acquaintance and cooperation with Cassirer on his work. It is in this sense that Langer becomes an actor in "Kuhn's Harvard" as the amalgamation of European and American historicist thought, which *Structure* eventually resolved.

According to Sander Verhaegh, symbolic logic had always played a role in Harvard's philosophy curriculum, but it had a limited impact (Verhaegh 2024, 22). In time, "[T]he academic landscape gradually started to shift in the years after the First World War" (21). With that shift, the intellectual climate changed rapidly as well at Harvard:

The department hired several philosophers who had contributed to the development of symbolic logic—H. M. Sheffer, C. I. Lewis, and A. N. Whitehead—and Harvard quickly began to be viewed as a central hub for analytic philosophy in the United States (21).

Susanne Katherina Langer was born in 1895 (as Susanne K. Knauth) had studied at Radcliffe College under Sheffer and Whitehead. Many instructors of the Radcliffe College also taught at the neighboring Harvard University. In 1926, She obtained her PhD in philosophy from Radcliffe College (established in 1879 as a college for women). Langer had Alfred North Whitehead as her dissertation advisor in 1924. Langer's parents were German immigrants even though she was born and raised in the United States. German language was spoken at home and she had a great command of the language.

Tracing Kuhn's point of contact with Langer in the archival materials at MIT, we find in one of Kuhn's papers written for H. M. Sheffer's course, an acknowledgement of Langer: "I have derived considerable assistance from the clear exposition in Susanne K. Langer's *Introduction to Symbolic Logic* (1937) (TSK-MIT, Box 1, Folder 3). Another point of Contact with Langer is Kuhn's reading [or sifting through] of Langer's *Philosophy in a New Key* along with other books by April 1949 (Reisch 2019, 160). The book was, among others, inspired by Ernst Cassirer "that pioneer in the philosophy of symbolism" (Langer 1948, preface).

Because of her family background and impeccable command of the German language "Langer's intellectual context would also have included German influences on her thought: philosophers such as Ernst Cassirer, Edmund Husserl, and Ludwig Wittgenstein" (Verhaegh 2024, 26). Langer had advocated a pluralistic conception in her *The Practice of Philosophy* (1930), which may indeed be attributed to Langer's Radcliffe Harvard education. "For Lewis and Sheffer, too, had defended variants of pluralism and were known for their pragmatic conception of the *a priori* " (25). Quine attending meetings of the Vienna Circle in 1933 had been surprised that they were reading Langer's book [*The Practice of Philosophy*] (27). Moritz Schlick singling out the book's method "the method of logical analysis, will be the only method of future philosophizing" (27). Schlick had valued the "author's exquisite style, lucid, fluent and brilliant, has been a source of real joy for me" (27). Rudolf Carnap had expressed interest to work with Langer once in the United States and had noted in his diary that Langer "used her first meeting with the by-then German philosopher to ask him about Frege" (27). Herbert Feigl called Langer's discussion group "*Langer Zirkel*" and had "told Schlick that it reminded him of the *Wiener Kreis*". "She is an excellent woman and her versatility is admirable" Feigl writes (27). Elsewhere Feigl wrote in a letter that "She reminds me a bit of Wittgenstein in her demeanor, in her intuitive determination, and in the biblical conciseness of her statements" (28). I venture to say that Kuhn in his acknowledgement of Langer with expressions "considerable assistance" and "clear exposition" points to her unquestionable philosophical prowess observed by those who had either read or met her.

Langer wrote an article in *Fortune Magazine* (see Langer 1944) and had been corresponding with Ernst Cassirer. In a letter on February 15, 1944, Cassirer thanks Langer for the article and writes:

Ich habe ihn mit großem Interesse studiert und mich an der grundsätzlichen Übereinstimmung unserer Auffassungen gefreut. Die Einwilligung zu Ihrer Übersetzung meiner Schrift „Sprache und Mythos“ gebe ich natürlich mit Freuden (Cassirer 2009, 229).

Langer translated the book into English as *Language and Myth* (1946) one year after Cassirer's sudden death. In the introduction to the book, she writes:

Reason is not man's primitive endowment, but his achievement. [...] I offer the translation of this little study [...] both as statement of a new philosophical insight and as a revelation of the philosopher's work: his material, his technique, and the solution of the problem by a final flash of interpretive genius (Langer 1946, ix, x).

It must now be obvious that Susanne Langer was highly respected in philosophical circles of the late '20s and '30s not only because of her background in philosophical logic and the eagerness with which some of the members of the Vienna Circle interacted with her, but also because of her personality. Expanding on Verhaegh's assessment, one could perhaps interpret Langer's change of philosophical interest – roughly towards the end of the World War II – as a sign of a shift away from “trends” in philosophical logic. It appears that at least with Langer, this “shift” may have begun with Langer's interest and acquaintance with the work of Cassirer. Kuhn's points of contact with Langer are indicative of making both Kuhn and Langer part and parcel of the Harvard hub and a party in the historicist scientific “strand”.

Now, enter Jean Piaget. Elsewhere, I have reconstructed Kuhn's thought on the finding of Piaget (see Basafa 2025, 92-106) for which there is not ample time for in this presentation. But such a reconstruction is intended to show that in the interaction with Piaget's experiments on kids, the young Kuhn went beyond the *gestalt* to expose what Kuhn discovered for himself as the ‘fringe of vague meaning’, which, first, was not only inspired by the philosophical awareness, trends or what I would call ‘textbook knowledge’ of the time, but, second and most importantly, Piaget experiments on kids were at the service of a strategy

to show that scientific activity was not only a purely logical undertaking, but that it also harbored elements that escaped clear, quasi “rational” definitions entirely. It is illuminating to note that Kuhn’s ‘fringe of vague meaning’ of the late 1940s resonates Langer’s notion of ‘insight’ in her book *The Practice of Philosophy* of 1930. Langer’s book had dedicated the entire chapter VII to ‘Insight’. In fact, one could say that this chapter, in a sense, vindicates Kuhn’s deep and persistent interest in this and other similar concepts, i.e., Polanyi’s tacit knowledge, or the ineffable in Wittgenstein. Again, Kuhn had, at least, read two books by Langer, namely, *Philosophy in a New Key* (1942) and Langer’s *Introduction to Symbolic Logic* (1937). It is not known to me that he had read Langer’s *The Practice of Philosophy*, but Langer’s assertions in the book, at the least, point to the philosophical awareness of the time and her great interest in Cassirer’s philosophy. Langer wrote:

Through the symbolic act or image, an important past event plays into the later history of the organism, without being discursively known. But I maintain that such formulation of an idea is knowledge in a wider sense—it acts as a stimulus, puts us into rapport with a past event just as a proposition might be (Langer 1930, 164).

She justified the above in the following way:

There is no knowledge without form; and probably no form is unique; therefore all knowledge can find symbolic expression. But it is the types of expression rather than the organ of recognition that determines the difference between insight and inference. [...] A scientist of genius is a person who can apprehend a new concept through some natural medium for whom there are unprobed patterns [...] which some enlightened psychologists have recognized as the principle of “Gestalt,” and which is “form” in a really general sense (165-166).

She then concluded the chapter by saying that there is nothing “alogical” about this process:

I do not see why, in the case of intensive [as opposed to extensive] symbolism, where the meaning is sometimes incommunicable, the process of understanding should be classed as “alogical” [...] the personal discovery of meanings through myth, ritual, and art, highly individual and

awe-inspiring by its subtlety, is the very acme of logical procedure, and the refinement of intelligence. (166)

No doubt, the source of language of myth for Langer is Cassirer. There are several references to Cassirer's *Philosophie der symbolischen Formen* (1923-1929). She writes:

The kernel of a myth is a remote idea, which is *shown*, not stated, in the myth. It is only the myth that is stated in words. Ernst Cassirer has made this point excellently clear in his "*Philosophie der symbolischen Formen*" [...]. (158)

Further, Langer also dedicated the bulk of chapter V titled "The Logical Basis of Meaning" to Wittgenstein's *Tractatus* (1921). The point is that Kuhn's "fringe of vague meaning" as well as similar concepts such as the "ineffable" or "tacit" or relating to "insight," *gestalt*, or symbolic form, was either a "craze" (to use Feyerabend's expression) of the time, or the "craze" was already part and parcel of the philosophical awareness of the time when discussing how science worked? I would vouch for the latter. For Kuhn, the question remained, however, that if science had such an "unspeakable" element, how did scientists chose among relevant theories?

The thesis I am harping on is that Kuhn's Harvard educational background and the genesis of *Structure* was the result of Harvard's "melting pot" of ideas harboring diverse strands of thought, which included both European and American philosophical strands. There were strands originating in Harvard's early days with its Kantian-inspired pragmatism as well as the logical strands that were later developed early in the last century with Sheffer, Russell, Whitehead, and Langer. Elsewhere, I have traced a direct line of Machian historicist thinking (see Basafa 2025, 45-56), which was passed on to Harvard's George Sarton, who had a direct influence on James B. Conant, Harvard's president at the time and Kuhn's mentor by quasi passing on the baton to Kuhn through Conant's General Education program. There were also the early reviewers of *Structure* such as the prominent historian of science Marie Boas-Hall, whose book Kuhn had used in his courses. In her review, Boas Hall had pointed to traces of Comte and Sarton in *Structure*. Philipp Frank was another prominent figure who connected the dots for the 'Unity of Science' at Harvard

— a Machian program, no doubt. Throughout this narrative, we need to keep an eye on Harvard pragmatism in the background and as a point of reference. A comparisons of Alexandre Koyré's historicism and Arthur Lovejoy's impeccable passion for intellectual history as well as his incessant anti-Russellian battles in countering neo-realism is indeed inevitable.

Now, a word or two on Harvard's critical realism and the overall picture of the genesis of *Structure*. I mentioned at the outset that a tug of war between Arthur Lovejoy and Bertrand Russell where the former emphasized historical treatment of science and the latter represented a purely logical view of the scientific enterprise, defined Harvard's pluralistic educational environment. In many respects, Lovejoy was fundamental in forming Kuhn's historicist thinking, which suggests that it was Lovejoy that not only formed Kuhn's historical thinking (as Kuhn had acknowledged) but that he also shaped Kuhn's philosophical thinking as evidenced by Kuhn's correspondence on March 17, 1955 with Lawrence S. Kubie, a New York psychiatrist and family friend (see Basafa 2025, 65). Even though Lovejoy is associated with American critical realism, which may have been neo-Kantian in origin as Matthias Neuber (see Neuber 2023) has suggested, I submit that American critical realists were too far removed from its alleged origins, both in terms of time and concrete influence, to be acknowledged as such – and particularly as it relates to the narrative I have suggested. Consider Roy Wood Sellars assertions in 1929 on what critical realism is:

The whole position is a frank physical realism very different from Kant's construction (Sellars 1929, 451).

Nevertheless, Kant's philosophy was so foundational that neo-Kantianism and its traces, no doubt, persisted in philosophies that followed – even if it meant going beyond Kant. In the words of James Bissett Pratt (one of the authors of *Essays in Critical Realism*, 1920):

For nearly a century idealism in some form or other dominated philosophy. Almost all the thinkers of to-day were brought up under its influence. Realism was banished from text-book and class-room except as a "terrible example," [...] (Pratt 1920, 88).

In short, I would argue that *Structure* owes its genesis not *only* to European historiography, some of which was no doubt propagated by prominent neo-Kantians such as Cassirer, but also to an amalgamation of many other strands of American philosophical thought. Sellars' words early in the century on 'interpretive function', echoed Kuhn in the mid-50s:

But, surely, perception is more than this. It is an interpretative response to an object. The stimulus aspect is just the beginning of true perception. [...] (452). These objects are present to the mind (known) but not in the mind. They are not noumena in any Kantian sense (Sellars 1929, 451).

As far as philosophy of science and perhaps its "overcoming" is concerned, there are further points of reference, some I have emphasized, which were not only of special importance to Kuhn, but was also instrumental for him in eventually elegantly assimilating ideas on how science works. I suspect that the nodes of assimilation in question, are Piaget's developmental psychology, and Ludwig Fleck's "textbook science" in *Entstehung und Entwicklung einer wissenschaftlichen Tatsache* (1935) as well as Kuhn's becoming acquainted with the works of Emile Meyerson, Leon Brunschvicg and Alexandre Koyré that presented for Kuhn a different and unaccustomed way of looking at history as it related to science. The Lowell Lectures (1951) offered Kuhn an early arena to give form to his early ideas in a public setting. One must therefore note that the genesis of *Structure* went even beyond the cast of characters Kuhn had on occasion mentioned; this may indeed strain Micheal Friedman's approach in finding casts of characters with Kantian background to advance his overall strategy of a post-Kuhnian philosophy of science *primarily* by way of Kant – using critique of Kuhn's philosophy of science as a springboard. I therefore find Friedman's strategy too general or too narrow a path to go beyond Kuhn. In contrast to Friedman's approach in (see Friedman 2001), I have placed the emphasis on historical awareness or historicism (Kantian in origin or otherwise). Indeed, I have shown that the evidence points to historicist thinking, particularly of course as it related to science. The figures who championed Kuhn's historical methods of choice, his "Honor List" if you will, were major actors very early in the century, and part and parcel of the historicist, historiographical era I have alluded to – and one that I submit was elegantly resolved as Kuhn's own brand of philosophy of science in *Structure*. Cassirer too, as a neo-Kantian, was very much aware of the science

of his day, as well as having a very deep sense of historicism. He was therefore not only an integral part of the continental European strand, but also crucial in generating debate with his contemporaries, e.g., debates with Koyré on Platonism. But again there were others such as Koyré who were not Kantian or neo-Kantian but contributed greatly to the historicist tradition and new historiographies; Koyré was pivotal to Kuhn and the genesis of *Structure*. Amid all of this, Harvard pragmatism was also the subject of debate as critical realists including Lovejoy had debated the role of Pragmatism and its position towards Idealism and Realism. In criticizing Dewey, Lovejoy writes “A consistent pragmatism must recognize”:

That, if a real physical world having the characteristics set forth by natural science is assumed, certain of the contents of experience, and specifically the contents of anticipation and retrospection, cannot be assigned to that world, and must therefore be called “psychical” (i.e. experienced but not physical entities) (Lovejoy 1920, 76).

I have associated the genesis of *Structure* with the historical thinking of figures such as Alexandre Koyré and Arthur Lovejoy and albeit indirectly with the historicism of Ernst Cassirer. Langer’s departure from philosophical logic can be attributed to her acquaintance and cooperation with Cassirer on his work. It is in this sense that Langer becomes an actor in “Kuhn’s Harvard” as the amalgamation of European and American historicist thought, which *Structure* eventually resolved.

References

Basafa, Mark Nader (2025), “*Kuhn’s Harvard and the Resolution of a Historicist Scientific Era in the Structure of Scientific Revolutions*” (dissertation), University of Vienna.

Cassirer, Ernst (1946), *Language and Myth*, translated by Susanne K. Langer, Dover Publications Inc. New York.

Cassirer, Ernst (2009), *Ausgewählter Wissenschaftlicher Briefwechsel; Nachgelassene Manuskripte und Texte*, Band 18, Herausgegeben von John Michael Krois, Felix Meiner Verlag, Hamburg.

Fleck, Ludwik (1979), *The Genesis and Development of a Scientific Fact*, The University of Chicago Press, originally published as *Entstehung und Entwicklung einer Wissenschaftlichen Tatsache: Einführung in die Lehre vom Denkstil und Denkkollektive* (1935), by Benno Schwabe & Co., Basel, Switzerland.

Friedman, Michael (2001), *Dynamics of Reason: The 1999 Kant Lectures at Stanford University*, CSLI Publications.

Langer, Susanne K. (1930), *The Practice of Philosophy*, Henry Holt and Company.

Langer, Susanne K. (1944), "The Lord of Creation" in *Fortune Magazine*, Vol. 30, January 1944.

Langer, Susanne K. (1948), *Philosophy in a New Key: A Study in the Symbolism of Reason, Rite, and Art*, The New American Library.

Lovejoy, Arthur O. (1920), *Essays in Critical Realism*, Macmillan and Co. Ltd, London.

Neuber, Matthias (2023), "Russell and American Realism," in *Topoi* (2024 43:127-133), Springer.

Pratt, James B. (1920), *Essays in Critical Realism*, Macmillan and Co. Ltd, London.

Reisch, George A. (2019) *The Politics of Paradigms*, State University of New York Press, Albany.

Sellars, Roy Wood (1929), "A Re-Examination of Critical Realism", *The Philosophical Review*, Sep., 1929, Vol. 38, No. 5 (Sep., 1929), 439-455.

Verhaegh, Sander (2024), *Susanne K. Langer and the Harvard School of Analysis* (chapter one), Gaikis, Lona , ed. *The Bloomsbury Handbook of Susanne K. Langer*. London,: Bloomsbury Academic, 2024. Bloomsbury Handbooks. Bloomsbury Handbooks. Bloomsbury Collections <<http://dx.doi.org/10.5040/9781350294660>>.

Archive Material

TSK-MIT, Box 1, Folder 3 (Massachusetts Institute of Technology, Distinctive Collections, Thomas S. Kuhn)

Marginalized Worldviews in Legal Institutions: Towards a Legal Hinge Epistemology

Melina Bentes (Lisbon, Portugal)

Abstract

How can Law become an important mechanism to the epistemic empowerment of marginalized worldviews? By connecting hinge epistemology with feminist standpoint theory, I attempt to construct a response to this question. As feminist standpoint theory reinforces, marginalized groups have epistemic advantage regarding their knowledge about how society works, because they suffer the effects of social and identity power, different from privileged groups. These effects are so present in their experiences that leads them to perceive social inequality as a hinge, as a certain assumption that constitutes their worldview, which transforms their knowledge in a better description of social reality, one that does not neglect these fundamental aspects of society that gives important epistemic resources to understand the complexity of social injustices. When applying this argumentation to the remarkable *Brown vs. Board of Education* legal case, it is demonstrated how a case of social injustice can be treated with the proper complexity, regulating the disagreements about opposing worldviews in order to balance evidences in favor of the acknowledgment of marginalized groups' certainties and experiences. This normative role of marginalized groups' hinges can rule out from legal knowledge any belief that would violate these groups' experiences and use the Law as a powerful epistemic mechanism towards the institutionalization of the awareness of marginalized groups' worldviews, guaranteeing that the legal system hears their voices and protects their interests.

1. Introduction

Hinge epistemology is a novel and promising area of Wittgenstinian research that has been acquiring a growing recognition in the past years. To contemporary epistemologists preoccupied with its developments, it would be possible to establish statues of certainty in knowledge through Wittgenstein's notion of "hinges", presented in his writings that led to *On Certainty*. His most important insight in the opportunity of the book was attesting that the skeptical argument cannot be answered due to its unintelligibility, leaving no room for judgment on its truth or falsity, for it violates the logical grammar of language games.

Wittgenstein claims that the framework of knowledge must establish certainties about the reality in order to maintain logical limits to someone's beliefs, avoiding skeptical hyperbolic doubt. He describes this framework as a world-picture/worldview that is constituted by assumptions that must be certain in order for any knowledge claim to have meaning. These assumptions

would logically support and rule out from the knowledge framework anything that violates it, for it would violate the natural ontology, the logic of the world (see Wittgenstein 1953: §§141-142).

To illustrate this idea, he presents a metaphor which compares the system of knowledge to a riverbed, where there would be solid assumptions about reality, as the bed of the river, functioning as channels to the development of more fluid propositions about reality, as the water of the river. The status of these propositions can depend on context, to the extent that a proposition may operate as something to be tested by experience, at one time, and something to be used as a rule of testing, a reference to the whole system, at another time (Wittgenstein 1953: §§96-99). The fluid propositions can only follow its flow if the riverbed is solid; if it collapses, the flow of the river stops and cannot fulfill its purpose anymore.

These assumptions are epistemic operators for any epistemic attitude in a community, functioning as hinges on a door that must be fixed in place for the door to fulfill its purpose (Wittgenstein 1953: §§341-343). Their most important aspect is the normative role before the framework of knowledge, for they are the rule of testing that regulate which evidence can figure within the system, excluding beliefs that do not fit this normativity (see Wittgenstein 1953: §§93-95, §105). But how is this norm fixed? When speaking about knowledge claims about empirical reality, the external world is arbitrary in establishing its norms, being impossible to conceive alternative realities that are as cruelly real as the one that there is.

This naturalistic interpretation gets tricky when speaking about knowledge claims about social reality, still, perhaps hinges conceived by marginalized worldviews can have the same normative role to social reality, for they cannot ignore the arbitrary force of the power relations that structure the framework of society, which silences their experiences in a cruelly real way. Through their perspective, they can have an advantage in relation to others, to understand the complexity of the dynamics of inequality through which society really works.

While seeking responses to the relativism and skepticism about the epistemic warrants of hinges, Coliva brings attention to the normativity of hinges,

defining them as not being epistemically rational, but norms constituting a community's notion of rationality, the conditions of possibility to any rational knowledge within it. Appealing to the upbringing within a certain community would not be enough to respond to skepticism and relativism, so the solution should focus on the fact that hinges cannot be epistemically justified, only presupposed, operating as norms of evidential significance, that determine what can count as evidence for ordinary epistemic propositions (Coliva 2015: 123-124).

The criticism she addresses to naturalism is that it cannot explain how some hinges are more natural than others, how assumptions that some communities hold as certain could not be made by others that would still be epistemically rational (Coliva 2015: 120-122). However, I intend to defend the idea that hinges created in marginalized worldviews can provide a better description of social phenomena, due to the plurality of their epistemic exchanges and resources, which allows them to enlarge their perception in ways that privileged worldviews never could, and due to their advantage in accessing the factivity of the social power that constitutes society's framework. For this reason, adopting a perspective that values their epistemic advantage is can improve knowledge produced within social and legal institutions.

2. Why marginalized standpoint matters

In this topic I claim that the normativity of social hinges can operate in two ways: (1) by considering which kind of evidence enters the worldview of this society and constitutes their idea of knowledge and (2) by providing a balance on socially situated knowledge that ensures attention to the experiences of marginalized groups, not allowing knowledge to imitate the structures of social power imposed on them. To bring attention to marginalized groups' experiences, knowledge can be conceived in a way that the flow of information about their experiences can be strengthened, fighting structural social inequalities by not allowing them to bias these groups' credibility and autonomy.

Feminist standpoint theory proposed a revolution in Epistemology while presenting a new paradigm for thinking about knowledge, focusing its attention on social conditions and the dynamics of epistemically biased knowledge. This new paradigm brought also a new conception of objectivity,

developing a critical perspective about the dynamics of social power on knowledge, while insisting on the need for situated knowledge to demystify the hegemonic social forces that operate behind the concept of traditional objectivity. At the end, standpoint theory proposes an inversion thesis, claiming that the marginalized part of this dispute has advantages on constructing knowledge about social reality, for they are in a better position to access the effects of social power upon non-hegemonic identities (Wylie 2003: 26).

When social power unites with identity power, there is an imaginative coordination about social identities that often creates stereotypes towards marginalized identities – what it means to be a woman, straight, gay, etc. – distorting social reality and imposing control of this authoritative knowledge upon the knowledge produced within marginalized groups. Since marginalized groups are often the target of this control, their experiences give them enough evidence to hold two assumptions as certain and constitutive of their knowledge about reality (as hinges): that society is shaped by structures of power and that this oppression is eliminable. With these assumptions in mind, one is able to understand these groups' realities, and how the injustices they suffer are manifestations of this discrepancy of social power (see Fricker 2007: 11-14).

The knowledge about social reality produced within marginalized worldviews, then, is less defective, because their social situation confronts them with the structures of power that shape society, transforming these assumptions into certainties that constitute their whole worldview about how the world works. These certainties (hinges) can better be sustained by marginalized groups because their agents are in position to grasp the effects of power relations on the world and upon others, while also having to negotiate with the authoritative knowledge of the privileged that their experiences are grounded in a fundamentally different understanding about reality (see Wylie 2003: 34-35).

Besides this, their knowledge has an epistemic advantage on purpose, because they have no interest in maintaining the oppression that is exercised against them, reinforcing the critical probity of their knowledge, that is produced in an environment that permits them to engage in epistemic projects with a critical

dissociation from authoritative knowledge that serves positions of privilege and maintains distorted conceptions (Wylie 2003: 37-39). Marginalized groups' experiences furnish them with a radically different worldview of society, for the effects of social power allow them to develop a critical distrust of the dominant worldview of socially privileged groups.

When the marginalized assumptions solidify as hinges in a groups' worldview, they can help diagnosing social problems with such complexity and sensibility that they can operate as regulative epistemic principles towards critical awareness and balance. When marginalized groups find a channel to generate and transmit their knowledge about reality, the factivity of social and identity power upon them is exposed to privileged groups, that suddenly are confronted with experiences that their social position never allowed access, placing new evidence and epistemic resources that can make them distrust the evidences and resources they had at their disposal.

Yet, my intention is not to demonstrate if this is possible, or how this change of perspective would happen within privileged groups. In fact, it is to demonstrate how adapting these ideas to think about hinges in legal epistemology can help the improvement of legal knowledge. In the following section, I use as example a legal paradigmatic case in order to show how epistemic agents from marginalized groups (or that understand marginalized hinges) can lead the normativity of these hinges to help dissolving concrete problems about apparently impossible disagreements in legal disputes.

3. Marginalized awareness and deep disagreements in legal cases

In order for any genuine argumentative exchange to happen there must exist a background of shared certainties, which is lacking in the case of deep disagreements. They occur when a clash of framework propositions make appealing to facts and normal criticism unfeasible, for the logical conditions that give meaning to them, in normal contexts of argumentation, do not obtain. Because the agents do not share the same worldviews, nor the same logical limits about reality, rational argumentation becomes an impossible task to accomplish (Fogelin 2005: 6-9).

Fogelin stresses that the remedy for the subversion of good and bad arguments is to teach students to analyse arguments in a manner sensitive to context, and appealing to this solution would enable them to diagnose a deep disagreement

under a normal argumentative context, where these minimal shared conditions for argumentation would be present. The example presented by him concerns the deep disagreement about affirmative action quotas for Black people, which for him, would be impossible to solve. I intend to demonstrate, through an example in legal institutions, that the worldview of marginalized groups can diagnose the complexity of the problems involving the knowledge about historical injustices and regulate knowledge in a way that deep disagreements could actually be solved.

On May 17th, 1954, the Supreme Court of the United States delivered the unanimous ruling that segregation in public schools was unconstitutional, violating the Fourteenth Amendment. It overturned the “separate but equal” doctrine, set by a precedent in the 1896 case *Plessy v. Ferguson*, which guaranteed equality of treatment for substantially equal facilities, even though these facilities were separated by race. After Black people sought remediation from the Court to obtain admission to public schools for Whites, the Court decided in favor of the plaintiffs, who contended that race segregated schools could never be equal (*Brown v. Board of Education of Topeka*, 1954).

Overturning an outdated doctrine was only possible because the Court comprehended the necessity to maintain the interpretation of the Law sensitive to social context, by seeing beyond an old inconclusive thesis that treated a case of historical reparation under mere “tangible” factors. In order to pursue this, the Court considered the effects of segregation in their given social context to respond to the question: does segregation in public schools, solely on the basis of race, even though the physical facilities and other “tangible” factors may be equal, deprive the minority group of equal educational opportunities? (*Brown v. Board of Education of Topeka*, 1954).

After valuing intangible considerations on the effects of racial segregation, the response reached by the Court was positive, concluding that segregation of White and Black children in public schools has a detrimental effect upon colored children, which would generate a collective feeling of inferiority as to the Black group’s status in the community that may affect their hearts and minds in an unpredictable way. This impact would be greater when it has the

sanction of the Law, for it would deprive them of some of the benefits they would receive in a racially integrated school system (*Brown v. Board of Education of Topeka*, 1954).

In the affirmative action quotas example brought by Fogelin, the deep disagreement would happen because of opposed moral standings, with the anti-quota agenda defending that only individuals could have moral claims, and the pro-quota agenda that social groups could have moral claims against others (Fogelin 2005: 10). But his conclusion does not present a proper diagnosis of the issue, and consequently, cannot reach an effective solution to it. Indeed, by a privileged and so-called neutral standpoint, this might be the apparent issue with the deep disagreement, but considering a marginalized standpoint, the issue reveals to be inherently linked with the historical injustices suffered by Black people.

In the *Brown v. Board of Education* example, as soon as the epistemic agent (legal operator) shares the same worldview as the marginalized groups and their certainty that society is shaped by structural inequalities, this introduces them to such a complex experience that they are able to perceive how this certainty manifests itself and affects knowledge in unmeasurable ways, turning it into enough evidence to reach an agreement about the diagnosis and the solution of this problem. The other normative role of hinges steps in when the legal mechanisms are used as means to foster awareness to marginalized worldviews, making the legal system at their disposal, as a channel to maximize the transmission of their experiences, despite the structure of social inequality.

4. The social normativity of legal epistemology

A reality of social inequality will always produce faulty knowledge, for social and epistemic injustice are two sides of the same coin, mutually supportive of each other. As demonstrated, considering as a certainty that society is based on relations social and identity power, and this is an undeniable aspect of social reality, we reach a necessary condition to any adequate diagnosis of problems concerning injustices. Even though oppression affects all agents in society, it does not affect all of them equally, for different situated groups have different epistemic predicaments (Medina 2015: 27-28).

As already argued, marginalized groups can claim epistemic advantage because being exposed to these two forms of power gives to their knowledge better epistemic and critical resources, about their experiences and others' experiences, also making them more prone to virtues of humility, curiosity and open-mindedness. Medina's proposal of regulative principles of epistemic friction can be interpreted as similar to the second aspect on the normative role of hinges, establishing that all cognitive forces encountered must be acknowledged and, in some way engaged, to lay out the desideratum of searching for balance in the interplay of cognitive forces, without some forces overpowering others. The willingness to put one's worldview in relation to that of others and calibrate different cognitive forces is the path to the epistemic virtues that are characteristic of marginalized groups (Medina 2015: 50-51).

The first aspect of hinges normativity occurs when epistemic agents, here being legal operators, share the same hinges as marginalized groups', especially about the advantage on their social position to access the social and identity power that constitute the framework of reality, which would function as a norm that denies any belief about social reality that neglects such a certainty. The second aspect occurs when they start operating as regulative principles towards acknowledgment, engagement and epistemic balance, constituting a resistance in the social and identity power epistemically exercised by privileged groups, that often imposes silences to knowledge that comes from the margins of society.

Through the social normativity of hinges, legal operators can transform the legal system in an institutionalized resistance in favor of the marginalized standpoints' epistemic empowerment, so their commitment with the importance of these identities ensures that institutional legal mechanisms will always be on their side, to overturn sexist, racist and homophobic legal fictions sustained by structural social inequalities. They would enable the historical knowledge and experience of marginalized groups to have a channel for being recognized as important and to figure in the very logical limits of these agents' worldviews, structuring the knowledge they acquire about social reality and guiding what counts as evidence for new social knowledge. When the marginalized worldview consolidate as the collective imagination of legal operators, they constitute the legal system from within and furnish the

marginalized groups with a strong mechanism against epistemic injustice. Thus, the Law would be able to fulfill its purpose and safeguard justice to all of those who seek its remediation.

References

Brown v. Board of Education of Topeka, Opinion; May 17, 1954; Records of the Supreme Court of the United States; Record Group 267; National Archives.

Coliva, Annalisa (2015). *Extended Rationality: A Hinge Epistemology*. London, England: Palgrave-Macmillan.

Fogelin, Robert (2005). The Logic of Deep Disagreements. *Informal Logic*, Vol. 25, No.1, p. 3-11.

Fricker, Miranda (2007). *Epistemic Injustice: Power and the Ethics of Knowing*. Oxford: Oxford University Press.

Medina, José (2015). *The Epistemology of Resistance*. Oxford: Oxford University Press.

Wylie, Alison (2003). Why Standpoint Matters. In Robert Figueroa & Sandra G. Harding, *Science and other Cultures: Issues in Philosophies of Science and Technology*. New York: Routledge. p. 26-48.

Wittgenstein, Ludwig. *On Certainty* (1953). London: Blackwell Publishing.

Acknowledgement

I thank prof. Nuno Venturinha for his supportive guidance and Josh Eisenthal for his helpful comments and English proofreading.

Why is Wittgenstein's anthropological turn, in fact, not a post-positivist sociological manifesto? David Bloor as his own enemy

Mehmet Onur Canpolat (Istanbul, Turkey)

Abstract

Wittgenstein's later remarks take their force from a communal perspective that focuses on the collective life of human beings. On this aspect of his philosophy, sociologists such as David Bloor develop a sociological interpretation of concepts like 'language-games' and 'forms of life'. However, the social element in Wittgenstein's later remarks is not necessarily sociological. Particularly because there is no one-to-one correspondence between the kinds of language-games and the kinds of activities, a correspondence that can provide us with a systematic sociological framework for studying forms of life. Yet, Wittgenstein's remarks still present a possibility of social sciences in which he defends a descriptive mode of analysis over an explanatory one. However, Bloor fails to see this and develops an interpretation of Wittgenstein that takes his philosophy as 'misbegotten sociology'. To repair Wittgenstein's philosophy, Bloor devises what he calls a typology of language-games that runs parallel with a typology of forms of life. I take Bloor's mistake as a result of his insistence on the principle of causality. Bloor's attempt to attain a causal mode of sociological explanation closes his eyes to the un-Wittgensteinian tone both in causality and explanation. Furthermore, Bloor's insistence on a causal explanatory model is a defense mechanism he developed against Bruno Latour's criticism of his work. It is due to a need to position himself against Latour in Science and Technology Studies literature that Bloor goes toward a radical misinterpretation of Wittgenstein's later remarks.

INTRODUCTION

It has been generally accepted that after his return to philosophy in 1929, Wittgenstein got away from the Tractarian idea that the primary function of language is to describe the world, and, thus, the logic of our language strives to find the general form of a proposition. He eventually finds peace in the idea that language is always the language of people, and meaning is to be seen in the background of this fact. At this apparent reference to 'human life,' one rightly tastes an anthropological flavor that celebrates the priority of life over language. However, such stress on human life also paved the way for radical interpretations in the form of social constructivism. Especially within the literature of Science and Technology Studies (hereafter STS), people such as David Bloor celebrated Wittgenstein's later philosophy as a manifesto for post-positivist sociology that can be useful in demystifying the value-laden aspect of science. In Bloor's view, this could be done by analyzing the types of language-game scientists use. Such an analysis, for Bloor, will reveal the fact that scientific language is institutionalized by the collective norms of that society,

which proves that science is a profoundly social activity. However, such an interpretation of Wittgenstein's 'anthropological turn' is an exaggeration, if not a distortion, and it is curious why Bloor might have gone that far. I shall explain some potential reasons.

1. Wittgenstein as Bloor's sociologist

Bloor is specifically interested in explaining religious behavior. He wishes to explain why this individual believed that, another believes this, and so on. On such a concern, he focuses on collective behavior on religious matters. He wishes to come up with an understanding of how a community's communal belief impacts how its members act. He then proposes a typical pattern running through the whole fabric of a belief system. He mentions that since believers will always take some empirical information as 'certainly true' and build a worldview around it, there will also emerge what can be called an 'anomaly'. That is, there will always be some counterfactual phenomena that threaten the believers' assumptions by leaving the belief system insufficient to explain the phenomena. So, practitioners constantly have to react to such phenomena in ways that maintain their belief systems. Bloor takes this constant necessity to respond to anomaly as an essential feature of all belief systems. In other words, where there is a belief system, the practitioners will always have relevant behavior that will be called a reaction to anomaly. By relying on Mary Douglas' work in her *Purity and Danger* (1966) (as well as citing Lakatos' *Proofs and Refutations* (1976)), Bloor phrases this as 'response to anomaly'. But for the sake of brevity, I will call it 'reaction'.

In Bloor's view, reaction is pre-determined by the community's social structure to which the practitioners belong. Practitioners' explanation in fighting against counterfactual phenomena is motivated by the underlying social order their community is based on. The stress on the pre-determination here goes hand in hand with Bloor's analysis of science in STS. Bloor's major contribution to the scholarship developing around the social analysis of science is the idea of science as a social institution, which has the same structure as any other belief system. More particularly, he defends that scientists also fight against counterfactual phenomena they cannot explain with their current theories, and they do so under the influence of some specific social structure. In this way, Bloor accounts for science as the 'belief system' of Western culture.

Bloor designated such an account of science in his famous book *Knowledge and Social Imagery* (1976), in which he presents the basic tenets of his Strong Programme that was to become one of the most renowned works in STS. Although he had an interest in Wittgenstein's potential impact on a post-positivist Sociology of Scientific Knowledge (in short SSK), as evident from his 1973 article 'Wittgenstein and Mannheim on The Sociology of Mathematics', it was only during/after the 1980s that he became interested in drawing insights from Wittgenstein's philosophy to support his radically constructivist theses more technically. In his books, *Wittgenstein, a Social Theory of Knowledge* (1983) and *Wittgenstein, Rules and Institutions* (1997), he mainly interprets Wittgenstein's philosophy as a sociology. By reverting the thesis Peter Winch formulated in his *The Idea of a Social Science* (1958), which is "much sociology is misbegotten philosophy", Bloor argues for the opposite: "much philosophy is misbegotten sociology" (Bloor, 1973, p.191, footnote 46). In his view, social sciences are necessary to illuminate philosophy, which shall be the take-home message from Wittgenstein's remarks after his 'anthropological turn'. However, for Bloor, Wittgenstein himself stopped short of revealing this fact. Therefore, the interpretative task shall be to demonstrate the genuine sociology underlying Wittgenstein's philosophy. Bloor writes in one of his later books that:

Wittgenstein leaves us in no doubt about the dimensions along which forms of life and language-games will vary [...] But neither he, nor his followers, have ever given us a comparative framework within which this variation can be understood and its causes and laws traced [...] It is high time this deficit was repaired. (1983, p.137)

Such an agenda surely makes sense since Bloor's post-positivist SSK attempts to give an account of science as yet another belief system of a culture. He uses some of Wittgenstein's central concepts to show how scientific practice and religious ritual have many commonalities. In *Wittgenstein*, Bloor finds what he craves: an encompassing conceptualization of the similarities between seemingly different activities, the emphasis on the collective aspect of human life, the analysis of certainty and objectivity as uncriticized bases for collective life, etc. By combining Wittgenstein's anthropocentric understanding with anthropologist Mary Douglas' framework, Bloor devises a systematic framework to analyze the underlying commonalities of science and other

belief systems. In doing so, he uses two of Wittgenstein's main concepts: 'language-games' and 'forms of life'. He aims to find a rigid connection between patterns of social structure and the variation of language-games. His first strategy is to determine the elementary patterns of social structure. Since any social structure is meaningful due to a boundary separating this community from another, which thereby characterizes a sense of 'us' and 'others', Bloor identifies two types of boundaries relying on Douglas' work: group and grid boundaries. The former characterizes the boundaries among different groups, whereas the latter establishes boundaries within a group. For example, the 'insider' and the 'outsider' (e.g., Middle-Eastern and Western communities) are matters of group boundary separating a group from another. In contrast, the hierarchies between the members within a group (e.g., scientists and laypeople in a Western community) reflect a grid boundary. He then adds to these a degree measuring the strength of the boundary, thereby attaining four potential social structures (see Figure 1). He says that a society may have a low group boundary, meaning their sense of 'us' and 'others' is not that strong, whereas they might have a high grid boundary, which means that the distinctions between classes within that society are sharp. Bloor takes these patterns to be exhaustive in explaining any community and, respectively, any human behavior.

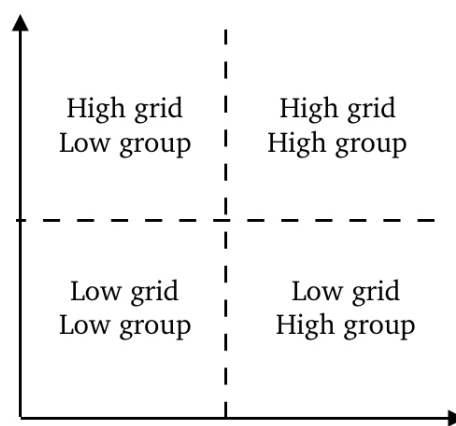


Figure 1

A typology of forms of life
(from Douglas)
(Bloor, WSTK, p.141)

An essential step in interpretation is that Bloor takes these four patterns as instantiations of 'forms of life'. Particularly because he interprets a 'form of life' as "a pattern of socially sustained boundaries" (1983, p.140). Every group boundary in his framework is socially motivated and designates a form of life. This is particularly important because this then gets connected with types of language-game since, in Wittgenstein's account, "to imagine a language means to imagine a form of life" (PI, §19) and vice versa. Therefore, the next strategy Bloor adopts is to spot a specific type of language-game and find four variations that will correspond to the variation in forms of life. Since Bloor is interested in explaining belief, he uses this framework to analyze human behavior concerning belief systems. And since the most characteristic feature of belief systems is reaction (to anomaly), Bloor's main subject of analysis is such reaction. Accordingly, he takes reaction as the type of language-game to be analyzed. In that, he takes the language-game itself as the reaction.

Douglas' work is primarily helpful for Bloor's analysis. He shows that, for Douglas, there are a few meaningful ways to react to anomaly, only four variations. The four different ways of executing the language-game of reaction are: 1. ignoring the anomaly; 2. rejecting it as a threat; 3. assimilating it to make it fit the norm; 4. accepting it as a source of change. Accordingly, Bloor points out this variation in the language-game of reaction (i.e., 1. indifference, 2. exclusion, 3. accommodation, and 4. opportunism (Bloor, 1983, p.139)) as the only source to which individuals will necessarily conform when they meet with anomaly in a religious context. For Bloor, any of these four ways of playing the language-game of reaction becomes the collective way of behaving in a relevant community due to the particular social pattern. In his words:

[...] Just as a consistent policy of responding to strangers would create a recognisable social style, so a consistent policy of responding to anomaly will stamp a characteristic physiognomy on a language-game [...This] policy is the language-game. (1983, p.141)

Such a consistent policy of responding to anomaly, as Bloor writes, is the language-game itself, and it becomes such a policy because of the recognizable social style. This one-to-one correspondence between the variation in the social structure and the institutionalized behavior in the form of language-game is the sociologist's primary instrument of analysis. Now, the sufficient

sociological explanation of an individual's behavior will have the form 'x reacted to the phenomenon A in such and such a way because she belongs to the social type 1'. All it takes to see how x reacts to A is to spot the specific variation of language-game x uses as a reaction. This particular variation will then naturally reveal the specific social style x belongs to. In that way, we can have a Wittgensteinian sociology, Bloor reckons. The fact that Wittgenstein failed to see this possibility in his philosophy must have been due to his naiveness.

As can be seen, Bloor finds a technicality in Wittgenstein's concepts, so they fit under the rubric of science. He deliberately distorts Wittgenstein's way of thinking to make it fit the dirt and dust of our world, the world of science, and the world of the sociologist. He believes this is necessary to make Wittgenstein's philosophy useful, to see that it is no philosophy but sociology.

2. Bloor's Wittgenstein as Wittgenstein's worst enemy

Bloor's talk of human behavior in religious matters is similar to what Wittgenstein calls ritual action in his *Remarks on Frazer's Golden Bough* (1967). Like Bloor, Wittgenstein talks of ritual action as one of the possible types of action humans carry out apart from animal actions such as eating, sleeping, etc. (1967, §15). As Bloor argues, it makes sense to assume that ritual action is an essential part of human life that springs up in almost every human society. But what about the anomaly? What about the reaction? Are they also crucial parts of human life, of ritual action? Here, the question of whether Wittgenstein would talk of an activity as being constituted by some essential elements, such as reaction to anomaly, deserves attention. The answer lies in one of the central concepts in Wittgenstein's later period: family resemblance. It denotes that our classificatory acts are not based on pointing out essential features of what is to be categorized but on detecting affinities (PI, §66). In that, we see that no category has its applications having one thing in common. We can easily say that instances of what we call ritual action will not necessarily have one thing that runs through the whole fabric of ritual. In other words, no specific type of language-game will be present in every ritual action specifically because it is not with regard to the presence of the same kind of language-games that we name an action as ritual but in relation to the resemblances these actions have with one another. Bloor fails to see that there is no possibility in Wittgenstein's philosophy for spotting a language-game as

the essence of an activity, that there are multiple ways of executing ritual action, and not in every form of life will the action yield the same multiplicity of language-games, the same combination of types of language-games, the same moves, etc. The action may and probably will be executed completely differently in every society. However, it still makes sense for us to place them all under the rubric of ritual action.

Here, it makes sense to say that when we talk about different communities, we talk about different forms of life. However, Bloor's interpretation of forms of life as socially sustained boundaries assumes that different forms of life are not different enough to produce incommensurable ways of practicing ritual. After all, in Bloor's eyes, separating one form of life from another is a matter of getting organized around this or that social rationale. However, Wittgenstein's understanding of forms of life is much broader than this. Bloor's apparent mistake is to assume that different forms of life will somehow end up devising the same action, only in different ways. In Bloor's interpretation, the difference between forms of life is not in the kind of activities and behaviors adopted by its partitioners, but in the way one and the same action is practiced.

Moreover, even if we give Bloor the idea that science is truly social, this still does not prove that science in Western communities and rituals in some African tribes do have precisely the same structure. As two different belief systems take place in different forms of life, they will use many different kinds of language-game. The multiplicity is not in the variation of one game but in the types of different games in general. For instance, in science, moves such as 'making a mistake', 'being wrong', 'making a hypothesis', 'testing it', 'establishing the result', etc., will become necessary language-games. Here, it makes sense to talk of anomaly and strategies against it. It makes sense to say that scientists will use language-games such as 'conforming to anomaly', 'denying it', etc. However, in the lives of, say, Azande people, no such kinds of language-game will be sensible. 'Making a hypothesis' or 'doubting it' will not become part of their lives. It should leave us wondering how anomaly and reaction to it can make sense in such a context. This is to say that the similarities between science and ritual shall not bring the two any closer other than both getting categorized as rituals. Although Bloor says, "[...] the similarities make it plausible to strive for an explanatory theory of intellectual

elaboration which covers both the Azande and the atomic scientist" (1976, p. 45), there should come Wittgenstein's warning, "What is most striking are not merely the similarities but also the differences" (1967, §39).

3. Genuinely Wittgensteinian social science

As has been shown, Wittgenstein's philosophy does not anyhow suggest a systematic analysis with rigid methodologies or a ready-made framework. However, neither does this mean that his philosophy forbids the possibility of a social science. On the contrary, it is suggestive of some sociological approaches if relevant. In his remarks on Frazer, where he is concerned with Frazer's anthropological method, Wittgenstein says:

I believe that the enterprise of explanation is already wrong because we only have to correctly put together what one already knows, without adding anything, and the kind of satisfaction that one attempts to attain through explanation comes of itself.

And here it isn't the explanation at all that satisfies us. (1967, §2)

The emphasis on the insufficiency of explanation here is the key. Wittgenstein takes it as an essential element of anthropology that this science shall leave everything as it is. This is particularly important because the anthropologist, as someone interested in other forms of life from within hers, will be at odds with how life in these different forms is meaningfully lived. As an outsider, she will not have their frame of mind, and this lack of participation will leave her blind to the functioning of language-games used in the relevant form of life. Explanation works only if one already knows what an explanation looks like under such and such a case. Yet, in the course of another form of life, one does not know how they make sense of the world. As Wittgenstein suggests, "One can only resort to description here, and say: such is human life. Compared to the impression that what is so described to us, explanation is too uncertain" (1967, §3).

Here, it becomes even more curious why Bloor might have ignored this aspect of Wittgenstein's philosophy, where a social scientific methodology is available. Is it that sociology is a special case of social science that what is said on anthropology cannot be applied to sociology? I believe the answer lies in Bloor's position against other scholars within the STS scholarship. There are

two main themes underlying Bloor's deliberate choice to go beyond Wittgenstein. One is to defend a then-radical idea that motivated STS movements in the early periods of the discipline: science is a value-laden activity. The other, which is used to rationalize the first idea, is the dependence on the causality principle. The first is a position that has been almost the only consensus among STS scholars. However, the latter has been a source of controversy, and I argue that it is to this we owe Bloor's intentional distortion of Wittgenstein's philosophy instead of following his advice. More particularly, especially in his debate with Bruno Latour, Bloor becomes so radical in committing to a causal explanatory approach against Latour's metaphysical and speculative style (see Bloor 1999a, 1999b; Latour 1999 for later discussion). That is, against Latour's criticism of his sociology of science as committing to a sort of idealism, Bloor wishes to defend his mode of analysis as causal in which it takes real phenomena as causally efficacious on scientists' belief, whereas this is still socially shaped since real factors are themselves not enough to raise final beliefs. Such a stance necessitates heavily depending on a strong causal understanding of social phenomena.

Furthermore, Bloor goes further and says that his sociological explanation also results from social determinacy, but this does not pose a problem since there is no possibility of transcending our social embeddedness. Thus, the social scientist is socially situated as much as her subject of investigation. In adopting a reflexive manner such as this, Bloor's reliance on causality can be seen as a radical departure from other scholars in STS since the causality principle itself was criticized by many during the later periods of the field. The importance of these principles for Bloor is reflected in two of the four tenets of his famous Strong Programme, which defends that the sociological analysis has to be 'causal' and 'reflexive'. I think it is on this deep connection to a causal mode of explanation Bloor must have ignored Wittgenstein's suggestion to rely on description in doing social sciences. It is due to the idea of a causal connection between natural and social factors that Bloor owes his protection against falling into the trap of idealist criticism raised by Latour. Therefore, for him, the causal relation between a specific social structure and a specific kind of language-game, i.e., the causality underlying the presence of a language-game and a relevant action, appears to be the most striking feature of sociology. In that sense, Bloor's reluctance to give up the idea of causality and explanation

by accepting a descriptive sociology of science pushes him to misinterpret Wittgenstein. Finally, it shall make sense to say that, as opposed to Bloor's interpretation, Wittgenstein's anthropological turn with all of its communal aspects is not necessarily sociological, although it has some methodological remarks. It is in this particular sense that Bloor seeks strength in the work of an author, Wittgenstein, whereas, in fact, it is Wittgenstein himself who is supposedly the one that will pose a significant threat to Bloor's sociology, which relies on a causal mode of explanation. In short, in his reference to Wittgenstein, Bloor becomes his own enemy.

References

Bloor, D. (1973). Wittgenstein and Mannheim on the Sociology of Mathematics. *Studies in history and philosophy of science*, 4(2), 173-191.

———(1991 [1976]). *Knowledge and social imagery*. University of Chicago Press.

———(1983). *Wittgenstein, a Social Theory of Knowledge*. The Macmillan Press, London.

———(1997). *Wittgenstein, rules and institutions*. Routledge

———(1999a). Anti-latour. *Studies in History and Philosophy of Science Part A*, 30(1), 81-112.

———(1999b). Reply to Bruno Latour. *Studies in history and philosophy of science*, 30, 131-138.

Douglas, M. (1966). *Purity and Danger. An Analysis of Concepts of Pollution and Taboo*. London, Routledge & Kegan Paul

Lakatos, I. (1976). *Proofs and refutations*. London: Nelson

Latour, B. (1999). For David Bloor... and beyond: A reply to David Bloor's' anti-Latour'. *Studies in history and philosophy of science*, 30, 113-130.

Winch, P. (1958). *The Idea of a Social Science* (London: Routledge and Kegan Paul, 1958).

Wittgenstein, L. (2009 [1953]). *Philosophical investigations*. John Wiley & Sons.

———(2018, [1967]). *Remarks on Frazer's The Golden Bough. The MyThology in our*, 29.

What Hilbert really wanted to say - Frege's clarification of metamathematics

Quirin Colares Oberrauch (Campinas, Brazil)

Abstract

I want to show that Frege's clarification of Hilbert's metatheoretic notions of structural definitions and logical consequence can fruitfully be read in analogy with Wittgenstein's treatment of impossibility proofs. First, I will clarify the point of analogy. Then, I will draw a parallel to the well-established interpretation of Frege's clarification of Hilbert's notion of axiomatic schemes as definitions of second-order relations. Finally, I suggest that this sheds light on Frege's still puzzling reconstruction of Hilbert's independence proofs through his peculiar conception of inference.

1 Introduction

Frege's remarks about Hilbert's *Grundlagen der Geometrie* "have often been denounced, and mostly dismissed". They share the fate of Wittgenstein's remarks on the first incompleteness theorem. (Cf. Floyd, 1995, 373) In the context of the debate about Frege's negative attitude about the possibility of logical metatheory, Wittgenstein's *Tractatus* has been interpreted as a resolution of the tensions in Frege's philosophy. (Ricketts (1986))

Despite all the merits of having pointed out a watershed in the philosophy of logic, the debate is founded on the vague conception of the universality of logic which can be taken to exclude (Ricketts (1986), Van Heijenoort (1967)) or to include the possibility of metatheory (Tappenden (2000), Dummett (1983)). Juliet Floyd's article on Frege's use of the definition-stroke is an instructive exception. As she notes, there is indeed a continuity between Frege's attitude towards metatheory and Wittgenstein's treatment of mathematical impossibility proofs. (Floyd, 2002, 151, fn.19)

The exploration of the continuity between Frege and Wittgenstein was first pioneered by G. E. M. Anscombe (Anscombe (1967), ch. 7) and later revitalized by Cora Diamond (Diamond (1991)). Moreover, Patricia Blanchette's work on the Frege-Hilbert controversy has significantly illuminated the question of Frege's semantic thinking. (Blanchette (2017), Blanchette (2014), Blanchette (2024))

Rather than focusing on broad philosophical themes, the underrepresented tradition prioritizes the examination of the subtle nuances in Frege's and Wittgenstein's treatment of concrete examples. In the same vein, I want to show that Frege's clarification of Hilbert's metatheoretic notions of structural definitions and logical consequence can fruitfully be read in analogy with Wittgenstein's treatment of impossibility proofs. First, I will clarify the point of analogy. Then, I will draw a parallel to the well-established interpretation of Frege's clarification of Hilbert's notion of axiomatic schemes as definitions of second-order relations. Finally, I suggest that this sheds light on Frege's still puzzling reconstruction of Hilbert's independence proofs through his peculiar conception of inference.

2 Wittgenstein on impossibility proofs

Wittgenstein repeatedly wrote on the meaning of impossibility proofs in mathematics. The famous paragraph § 23 of the *Philosophische Untersuchungen* on the multiplicity of language games likens their changes to the "Wandlungen der Mathematik". As a prelude to this, Wittgenstein considers the meaning of Frege's distinction between assertoric force and the content of a judgment as the singular propositional form in *Begriffsschrift*. (Cf. PI § 22)

Wittgenstein points out that it is only in the context of certain distinctions within the use of language it makes sense to say that to a question corresponds a propositional content, e. g. the assertory versus the interrogatory, hypothetical or fictional use of a proposition. (Cf. PI § 22) But it creates the illusion that there must be a identical propositional content, separate from its differing uses. Similarly, (Floyd, 1995, 397) Wittgenstein approaches his notorious remarks about Gödel's proof by contrasting the assertory use in the "game of truth-functions" with their use as parts of other sentences or as linguistic exercises. (RFM I, Appendix §2-6) He points out that there is no sense to the question whether a proposition is true independently of the language game we play.

Prior to the eighteenth-century proof of the impossibility of angle trisection, the purported trisectors had been held captive for two thousand years by the image (Cf. PI § 115) that there must be a clear sense to the conjecture "There exists a general method of Euclidean trisection?" (Cf. Floyd, 1995, 382-383, 390-391) To prove the impossibility of a solution to a mathematical problem, a

precise definition must be given of what a solution generally consists of. And this can only be achieved by stepping outside the system and providing a general characterization of all possible solutions. (Cf. Floyd, 1995, 391)

A major shift in the evolution of the problem occurred when it was agreed that the proof would have to be carried out in Euclidean terms, i. e. as a construction with straightedge and compass. (Cf. Floyd, 1995, 388, fn 52) The second major step was to characterize the possible constructions algebraically as the resolution of a corresponding arithmetic equation. (Cf. Floyd, 1995, 391) Thus, it is shown that it is impossible to trisect the angle by straightedge and compass, because the corresponding equation of trisecting the angle has no solution in the rational field, i. e. the field of all constructible numbers by Euclidean means. (Cf. Floyd, 1995, 391)

The difficulty is that the "proposition "P is unprovable" has a different sense afterwards—from before it was proved." (RFM I, Appendix III, 16) There never had been such a thing as "the trisection of the angle" in Euclid. To see this, the algebraic conception had to be recognized as the clarification of what the question actually consisted of. They had to be led to recognize it as "what [they] really wanted to say" (Cf. PI 334, Cf. Floyd, 1995, 393)

Like the trisection example, the question of whether the parallel postulate can be derived from the other axioms dates back to Euclid. In *Die Grundlagen der Geometrie* (1899) Hilbert introduces what is now the standard method for proving the independence of axioms using the Euclidean geometry as an example. (Cf. Blanchette, 2017, 2–3) It is widely accepted that he successfully provides a sufficiently precise formulation of the question of whether the parallel postulate can be derived from the other axioms in Euclid's geometry, ultimately proving that it cannot.

Hilbert interprets the axioms arithmetically and demonstrates that, under this interpretation, all the remaining axioms hold true, while the negation of the parallel postulate also becomes a theorem. (Cf. Blanchette, 2017, 3) Assuming that the deduction rules preserve the designated value of "being a theorem of Euclidean geometry" and remain valid independently of the interpretation of non-logical terms, the parallel postulate cannot be a logical consequence of the remaining axioms. (Cf. Blanchette, 2017, 17)

However, Hilbert's methodological innovations were not understood until mathematical logic had developed the foundational concepts of model-theory. (Button and Walsh, 2018) Hilbert's technique does provide a method for proving independence by offering a general characterization of what it means for a statement to formally follow from other statements. Nevertheless, Hilbert's methodological remarks fail to clarify the broader system within which the proof is conducted. Neither the concept of formal deduction nor the admissibility of reinterpreting non-logical terms is explicitly clarified. (Cf. Blanchette, 2017, 4) Hilbert simply assumes a general notion of proposition, presuming that the terms "axiom" or "theorem" maintain the same meaning across systems. Consequently, the logical form of the proof remains unclear.

3 Frege on structural definitions

In response to Hilbert's publication of *Die Grundlagen der Geometrie*, Frege takes issue with its initial claim that the suggested axioms for Euclidean geometry define their primitive terms. First, Hilbert asserts that these axioms express basic facts about our spatial intuition, thereby appearing to adopt them in the Euclidian sense. (Cf. Hilbert, 1899, 1, Frege, 1976, 61) Then, he presumes the existence of the systems of "points", "lines" and "planes", claiming that the axioms provide the "genaue und vollständige Beschreibung"—i. e. the precise and complete definition—of their relations: "lie", "between", "parallel", "congruent", and "continuous". (Hilbert, 1899, §§ 1, 3)

Crucially, Hilbert takes this as justification to reinterpret the primitive terms arithmetically—a point as a pair of numbers, a line as a linear function etc. (Frege, 1976, §9). As a result, the axioms and theorems of geometry are reinterpreted in terms of axioms and theorems of arithmetic. Hilbert employs this method to demonstrate both the consistency and independence of the axioms. (Hilbert, 1899, §§9-10)

Frege's critique appears offensively simple compared to the sophistication and accomplishments of Hilbert's methodology: "If the terms in the proposed 'axioms' do not have meaning beforehand, then the statements cannot be true (or false), and thus they cannot be axioms. If they do have meaning beforehand, then the 'axioms' cannot be definitions." (Shapiro (2005), 65) Hilbert, frustrated by Frege's persistent concerns about the ambiguity of the primitive terms, responds with a frequently quoted remark:

"Ja, es ist doch selbstverständlich eine jede Theorie nur ein Fachwerk oder Schema von Begriffen nebst ihren nothwendigen Beziehungen zu einander, und die Grundelemente können in beliebiger Weise gedacht werden. Wenn ich unter meinen Punkten irgendwelche Systeme von Dingen, z. B. das System: Liebe, Gesetz, Schornsteinfeger . . . , denke und dann nur meine sämtlichen Axiome als Beziehungen zwischen diesen Dingen annehme, so gelten meine Sätze, z. B. der Pythagoras auch von diesen Dingen." (Frege, 1976, 67)

Based on this understanding, Hilbert believes he can maintain the claim that his axioms function as definitions,—if not individually, then collectively. (Frege, 1976, 65–66) He asserts that a theory can be applied to an infinite number of equinumerous systems of primitive elements, provided that the relations defined by the axiomatic scheme remain valid within each system. Furthermore, he defers the question of whether an interpretation should be variable or definite to the formal development of a theory or its application. (Frege, 1976, 68)

Even if one accepts that the axioms are not fully interpreted sentences, they should, as definitions, give precise and complete meaning to the primitive terms, ensuring that their extension is uniquely determined. The key issue is that the reference is intentionally left indeterminate, leaving it unclear in what sense they qualify as definitions. When pressed on this issue, Hilbert's followers defended the notion that the axioms "implicitly define" the primitive terms by specifying neither more nor less than the relations between them. (Cf. Gabriel, 1978, 422–423)

However, the concept of implicit definition was originally introduced by Gergonne using the analogy of a system of arithmetical equations. (Cf. Gabriel, 1978, 420) A number or sequence of numbers is implicitly defined by a system of equations, if they constitute the unique solution to the system. This requires that the number of equations matches the number of unknowns and that there exists exactly one solution. (Cf. Gabriel, 1978, 420)

Frege takes up the analogy of Gergonne and remarks that, pace Hilbert, the reference must be uniquely determined. (Cf. Frege, 1976, 73) Just as a system of equations must be solved to prove the existence of a solution, every implicit

definition must ultimately be transformed into an explicit definition. (Cf. Gabriel, 1978, 422) Frege's demand that an implicit definition in Gergonne's sense can always be turned into an explicit definition was stated as a criterion by Padoa and was model-theoretically justified later by Beth's theorem. (Cf. Button and Walsh, 2018, 442)

Thus, the retreat to implicit definition is not of much help to make sense of Hilbert's method of reinterpretation. In retrospect, Wilfried Hodges remarks from a model-theoretic perspective:

"Frege's demolition of the implicit definition doctrine was masterly, but it came too late to save Hilbert from saying, at the beginning of his *Grundlagen der Geometrie*, that his axioms give 'the exact and mathematically adequate description' of the relations 'lie', 'between' and 'congruent'. Fortunately Hilbert's mathematics speaks for itself, and one can simply bypass these philosophical faux pas." (Hodges (2023))

Hilbert's prose could be ignored, but Frege's critique remained unanswered. So, Frege sought to clarify the logical form of Hilbert's method. Frege reconstructs Hilbert's method of reinterpretation as what became known as a structural definition. Hilbert's axioms do not define their primitive terms; rather they substitute these terms with variables of first order concepts, interpreting "point" as a geometric point, a pair of integers etc. (Cf. Frege, 1976, 74, Cf. Button and Walsh, 2018, 443)

"Die Sache ist wohl so zu denken, daß der Begriff *ist ein Paar von Zahlen des Bereiches Ω* , der erster Stufe ist, gleich wie der Euklidische Punktbegriff in den Hilbertschen Begriff zweiter Stufe (falls dieser vorhanden ist) fallen soll." (Frege, 1990, 374)

Contrary to some interpreters, (Demopoulos, 1994, 219, Giovannini and Schiemer, 2021, 16) Frege generalizes this method of reframing a theory as a second-order relation, formulating an account (Frege, 1990, 281) that, while not identical to, is nonetheless equivalent to the model-theoretic interpretation of Hilbert's technique. Consequently, the axiom scheme, understood as a conjunction of sentences about variable first-order concepts, functions as a second-level relation that determines a class of models in which all sentences of the theory are satisfied.

4 Frege on logical consequence

Even if Hilbert was not clear about the logical meaning of his method, the crucial point is that it provides a technique for proving the independence of axioms from one another. In the last of his articles on Hilbert's *Grundlagen der Geometrie*, Frege addresses the logical clarification of dependence in his own terms. (Frege, 1990, 423) He considers what it means for a statement to logically follow from other statements.

To begin with, according to Frege, inferential relations do not hold between sentences, but between thoughts. (Frege, 1990, 423–424) Moreover, they hold only between true thoughts. (Frege, 1990, 424–425) Therefore, he interprets "reinterpretation" as a one-to-one mapping between groups of thoughts with a definite truth-value. Now, assume that there is a group of thoughts, Ω , and a thought, A . Furthermore, assume that Ω is mapped one-to-one to another group of thoughts, Ω' , and that A is mapped to another thought, A' . Assume also that in this mapping every non-logical constant is mapped to a non-logical constant of the same logical category (object, first-order concept, second-order concept), but which expresses a different sense, whereas the logical constants are mapped to themselves. (Frege, 1990, 427–428)

If this mapping of the subpropositional elements induces a mapping in which every thought of Ω is mapped to a true thought of Ω' , but the thought A is mapped to a false thought A' , then the thought A cannot be a logical consequence of Ω . As such it would depend solely on the logical category, logical constants and logical rules applied in the inferential steps. (Frege, 1990, 428) With the exception of the peculiar requirement that inferential relations hold only between true thoughts, this procedure can be seen as foreshadowing Tarski's definition of logical consequence. (Demopoulos (1994), Tarski (1936))

Thus, it provides a clarification of the technique Hilbert uses to prove independence, analogous to its model-theoretical sense. But Frege's far-reaching insight, later echoed by Tarski, is that this technique presupposes a clear distinction between logical and non-logical constants within a language. For Frege, this requires a logically perfect language, in which every thought is expressed unequivocally—meaning its logical content is fully analyzed. (Frege, 1990, 427, 429, Blanchette, 2014, 14–15) This is likely the reason he abandoned further development of these ideas, (Blanchette, 2014, 20) though it remains

speculative whether he considered such an endeavor impossible in principle, as suggested by the anti-metatheory interpretation. After all, was his Begriffsschrift not intended as a perfect language?

5 What Hilbert really wanted to say

Wittgenstein's remarks on impossibility proofs highlight that the meaning of a question cannot be understood independently of the context in which it is situated. Like Wittgenstein's observations on the impossibility proofs of the trisection problem or Gödel's theorems, Frege's comments do not challenge the validity of the proof, but instead seek to clarify what the proof consists of. He critiques and refines the logical meaning of the conceptual underpinnings by bypassing Hilbert's accompanying prose. (Cf. Floyd, 1995, 374, RFM VII § 19)

While Hilbert demonstrates the independence of the parallel postulate from the other axioms of Euclidean geometry, Frege observes that the meanings of "axiom" and "logical consequence" have shifted. Thus, "Independence" in Hilbert's sense, is not "independence" in Euclid's or Frege's sense. (Cf. Blanchette, 2017, 6) As Wittgenstein pointed out, an impossibility proof requires that the meaning it assigns to the question is accepted as a clarification of what the question actually means. (Cf. Wittgenstein (1974) § 30, Floyd (1995), 390).

I suggest, that Frege's point is that there must be a way to reconcile Hilbert's notion of independence with Euclid's. Some argue that Frege simply "did not get it or did not want to" (Shapiro (2005), 66), but I believe that this is an anachronistic view. In some sense, Frege understood Hilbert's work better than Hilbert himself. Even if incomplete, Frege's logical clarification of Hilbert's structural definition as a second-order relation and his definition of logical dependence foreshadowed Tarski's model-theoretic definitions of a class of structures and logical consequence. In this light, Frege's critique and reconstruction of Hilbert's method can be seen as an effort to clarify what Hilbert "really wanted to say" (PI § 334).

6 Conclusion

The point I want to emphasize is that, even though Frege remained skeptical about Hilbert's technique for proving independence, he was correct that

Hilbert had not come to terms with the logical form of these proofs. Thus, Frege's exercise in clarifying Hilbert's proof illuminates a key conceptual shift in our understanding of mathematics and model-theory.

Despite the significant differences between Wittgenstein's method from Frege's insistence on the distinction between the assertoric force and the content of a judgment—for instance, that propositional content remains identical across different uses of language—there is also a fundamental continuity. Frege's insistence that the sense and reference of a logical theory must be determined completely can be read most fruitfully as an assertion that logic derives meaning through its application (*Anwendung*).

References

- G. E. M. Anscombe. (1967) *An Introduction to Wittgenstein's Tractatus*, New York: Harper and Row.
- P. Blanchette (2014) *Frege on Formality and the 1906 Independence Test*, in: *Formalism and Beyond: On the Nature of Mathematical Discourse*, 23:97
- P. Blanchette (2017) *Models in Geometry and Logic: 1870-1920*, *Logic, Methodology and Philosophy of Science - proceedings of the 15th International Congress*, ed. Niiniluoto, Seppälä, Sober; College Publications 2017, 41-61.
- P. Blanchette (2024) *The Frege-Hilbert Controversy*, In E. N. Zalta and U. Nodelman, editors, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, Spring 2024 edition.
- T. Button and S. Walsh (2018) *Philosophy and Model Theory*, Oxford University Press.
- W. Demopoulos (1994) *Frege, Hilbert, and the Conceptual Structure of Model Theory*, *History and philosophy of logic*, 15(2):211–225.
- C. Diamond (1991) *Frege and Nonsense*, In: *The Realistic Spirit: Wittgenstein, Philosophy, and the Mind*, The MIT Press.
- M. Dummett (1983) *The Interpretation of Frege's Philosophy*. *Erkenntnis*, 20(2).
- J. Floyd (1995) *On saying what you really want to say: Wittgenstein, Gödel, and the Trisection of the Angle*, In: *From Dedekind to Gödel: Essays on the Development of the Foundations of Mathematics*, pages 373–425. Springer.
- J. Floyd (2002) *Frege, semantics, and the double definition stroke*, In *The story of Analytic Philosophy*, pages 155–180, Routledge.

G. Frege (1976) *Nachgelassene Schriften und Wissenschaftlicher Briefwechsel. Zweiter Band: Wissenschaftlicher Briefwechsel*, volume 2. Felix Meiner Verlag.

G. Frege (1990) *Kleine Schriften*, Georg Olms Verlag.

G. Gabriel (1978) *Implizite Definitionen—eine Verwechselungsgeschichte*, *Annals of Science*, 35(4):419–423.

E. N. Giovannini and G. Schiemer (2021) *What are implicit definitions?* *Erkenntnis*, 86:1661–1691.

D. Hilbert. (1899) *Grundlagen der Geometrie*: 14c. Auflage. BG Teubner.

W. Hodges (2023) *Model Theory*. In E. N. Zalta and U. Nodelman, editors, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, Fall 2023 edition.

T. G. Ricketts (1986) *Objectivity and objecthood: Frege's metaphysics of judgment* In *Frege synthesized: Essays on the philosophical and foundational work of Gottlob Frege*, pages 65–95. Springer, .

S. Shapiro (2005) *Categories, structures, and the frege-hilbert controversy: The status of meta-mathematics*. *Philosophia Mathematica*, 13(1):61–77.

J. Tappenden (2000) *Frege on axioms, indirect proof, and independence arguments in geometry: Did frege reject independence arguments?* *Notre Dame Journal of Formal Logic*, 41(3):271–315.

A. Tarski (1936) *On the concept of logical consequence*, *Logic, semantics, meta-mathematics*, 52:409–420.

J. Van Heijenoort (1967) *Logic as calculus and logic as language*, In *Proceedings of the Boston Colloquium for the Philosophy of Science 1964/1966: In Memory of Norwood Russell Hanson*, pages 440–446. Springer.

L. Wittgenstein, G. Anscombe, R. Rhees, and G. Wright, (1974), *Bemerkungen über die Grundlagen der Mathematik*. Frankfurt am Main: Suhrkamp.

L. Wittgenstein, G. Anscombe, R. Rhees, and G. Wright, (1969) *Philosophische Untersuchungen*, Suhrkamp.

Language and Methodological Limits in Wittgenstein and Hayek

Jeffrey Patrick Colgan (New Orleans & Iowa City, United States)

David Ween (New Orleans, United States)

Abstract

Beyond their familial relation and origins in Vienna, the intellectual connections between Friedrich Hayek and Ludwig Wittgenstein might appear tenuous. Hayek, whose work ranged from technical economic analyses to investigations of the linkages between economics and political thought, seems as if he would be uninterested in the unorthodox yet profound investigations into logic, mathematics, language, and mind undertaken by his older cousin. However, Hayek claimed to have been deeply influenced by Wittgenstein, was an early reader of the *Tractatus Logico-Philosophicus*, and after Wittgenstein's death began compiling materials for a biography of Wittgenstein. Though much scholarly attention has been directed to Wittgenstein and Hayek separately, little has explored the influence of Wittgenstein on Hayek and how reading these two authors with and against each other might offer insights into their own work. This paper argues that Hayek and Wittgenstein can be read as participating in a shared dialogue about the limits of language and the far-reaching consequences of those limits. Section One locates these thinkers in the particular intellectual milieu of *fin de siècle* and early twentieth century Vienna, highlighting the shared conditions, values, and worries that motivated their future work and its reception. Section Two treats the connections between Hayek's work and the *Tractatus*-era Wittgenstein as well as their shared enduring commitment to the importance of language in philosophy, political thought, and the sciences. Section Three moves beyond the *Tractatus* and argues that important productive commonalities can be established between Hayek's work and Wittgenstein's later writings.

Introduction: Wittgenstein, Hayek, and the Limits of Language

Beyond their familial relation and shared origins in the *grande bourgeoisie* of *fin de siècle* Vienna, the intellectual connections between Friedrich Hayek and Ludwig Wittgenstein—two of the most influential thinkers of the twentieth century—might appear tenuous. Hayek, whose work ranged from technical economic analyses to investigations of the linkages between economics and political thought, seems as if he would be uninterested in the unorthodox yet profound investigations into logic, mathematics, language, and mind undertaken by his older cousin Wittgenstein. However, Hayek claimed to have been deeply influenced by both the content and method of Wittgenstein's philosophical work. Hayek was an early reader of the original German edition of Wittgenstein's *Tractatus Logico-Philosophicus* (1921), and after Wittgenstein's death he began compiling materials for a biography of his relative. Though much scholarly attention has been directed to Wittgenstein and Hayek separately, very little has explored the influence of Wittgenstein on

Hayek and how reading these two authors with and against each other might offer insights into their own works and the historical period out of which they originated. This paper argues that Hayek and Wittgenstein can be read as participating in a shared dialogue about the limits of language and the far-reaching consequences of those limits. Wittgenstein's investigations into the limits of sense-full language serve as the provocation to which Hayek is responding, even if his own writings are primarily concerned with tracing out how these proposed limits affect scientific methodology.

Much of the previous literature that deals with the connections between Wittgenstein and Hayek either focuses solely on their limited biographical connections (see Ebenstein 2001, Monk 1990: 518) or willfully misreads Wittgenstein so as to establish an antagonism of sorts between their philosophical work (see Ebenstein 2003). Notable exceptions include Gray 1998 and Erbacher 2020. It is in Erbacher's commentary and Janik's afterword in *Friedrich August von Hayek's Draft Biography of Ludwig Wittgenstein: The Text and Its History* that readers can find the most sustained engagement with the commonalities in biography, philosophical interests, and reception between Wittgenstein and Hayek. This project builds upon existing scholarship in, at least, two ways. First, throughout the emphasis is on the philosophical commonalities between these two figures with respect to language and method, rather than just epistemology. Second, most scholarly engagement has focused on the connections between *Tractatus*-era Wittgenstein and the work of Hayek. This paper argues that significant commonalities can be established between Hayek's work, especially on spontaneous order, and the posthumously published writings of Wittgenstein.

Section One locates these thinkers in the intellectual milieu of *fin de siècle* and early 20th century Vienna, highlighting the shared conditions, values, and worries that motivated their work and its reception. Section Two treats the connections between Hayek's work and the *Tractatus*-era Wittgenstein, particularly their shared commitment to the importance of language in philosophy, political thought, and the sciences. Section Three moves beyond the influence of the *Tractatus* on Hayek and argues that important productive commonalities can be established between Hayek's work and Wittgenstein's posthumously published writings.

1. Family Relations, Viennese Origins

Ludwig Josef Johann Wittgenstein was born on April 26, 1889, into one of the wealthiest families of Vienna. He was the eighth and youngest child of industrialist Karl Wittgenstein and Leopoldine Kalmus (see Monk: 3-27). Ten years later Friedrich August von Hayek was born on May 8, 1899, into a family of Viennese academics and scientists. He was the eldest of three children by physician-botanist-academic August von Hayek and Felicitas von Juraschek. Wittgenstein's maternal grandmother—Leopoldine's mother Marie Stallner—was the sister of Hayek's maternal great-grandfather Johann Stallner, making Wittgenstein and Hayek second cousins (Caldwell & Klausinger: 11). Their immediate families overlapped in Viennese society, and there were brief interactions when both were young; however, it was only as adults that Wittgenstein and Hayek had substantial interactions. Their first significant contact was in 1918, when both were officers in the Austrian army and shared a carriage on an all-night train ride back to the warfront. Hayek was struck by Wittgenstein's "radical passion for truthfulness in everything" (Flowers & Ground: 50). They met next in 1928 on the grounds of Cambridge and again in the early 1940s, when Hayek, who was at the London School of Economics, attended a talk at the Cambridge Moral Science Club with Richard Braithwaite (Flowers & Ground: 51). In the early 1940s, Hayek visited Wittgenstein a handful of times in his rooms at Cambridge, and around the end of the War they corresponded by letter about supporting and, eventually, visiting their relatives in Austria (Flowers & Ground: 52). Their final meeting came by accident aboard an overnight train in 1947 as both were returning to England from a visit to Vienna (Monk: 518; Flowers & Ground: 52).

Though their personal contact was infrequent, Wittgenstein made a lasting impression on the younger Hayek. In 1977, Hayek published "Remembering My Cousin, Ludwig Wittgenstein" in the journal *Encounter*, where he outlines their family ties, interactions, close-reading of Wittgenstein's *Tractatus*, and the effect Wittgenstein's character had on Hayek. Most notably, Hayek began preparing a biography of Wittgenstein after the latter's death. Though ultimately thwarted by both Wittgenstein's literary executors and Wittgenstein's sister Margaret, the attempted biography demonstrates a lifelong interest in the life and philosophical work of his cousin (Manzi: 254; see also Erbacher 2020).

Both Wittgenstein and Hayek were profoundly influenced by their Viennese Origins. Though Wittgenstein was slightly older than Hayek, their childhoods and young adult lives took place during the malaise of the last years of Habsburg Vienna, the shockingly quick collapse of the Austro-Hungarian Empire, and the doomed First Austrian Republic. One notable aspect of late Habsburg Vienna that provides insight into Wittgenstein's and Hayek's catholic approaches to intellectual work and the befuddlement among their later English peers was their apparent lack of academic and artistic specialization and professionalization (Janik: 18). Coming from a society where informal and intermixing groups were meeting almost daily to discuss and debate art, politics, science, philosophy, and more, both Wittgenstein and Hayek had broad, interdisciplinary interests and were at times uneasy participants in the specialization of Anglo-American academia. Wittgenstein famously insisted that the main point of the *Tractatus*, ostensibly a text about language and logic, was an ethical point (for an overview of the ethical readings of *TLP* see Bronzo 2012; Christensen 2024, Citron 2019, Colgan forthcoming 2025, and Ware 2011). So, too, with Hayek; his early interests in the philosophy of language and psychology are inseparable from his more mature work in epistemology and the social sciences (Gray: 8-16, 21, 23). Beyond the intellectual milieu of Vienna, both witnessed a profound crisis and collapse of liberalism. Though both dealt with this in different ways, the promises and challenges of liberalism haunt their social and political commitments.

2. The Language of Science and the Science of Language

To the extent that scholars have engaged with the philosophical and methodological connections between Wittgenstein and Hayek, that engagement has focused on the commonalities and differences between Hayek and the *Tractatus*—and for good reason. Hayek was an early admirer of the *Tractatus*, having stated that it made a “great impression upon him” (Flowers & Ground: 51). Hayek had an interest in the philosophy of language that predated his engagement with the *Tractatus*, through his reading of the journalist-turned-philosophical nominalist Fritz Mauthner, who identified philosophy with the critique of language (Gray: 13; Glock: 12). Notably, in the *Tractatus* Wittgenstein disparages Mauthner's nominalism and distinguishes it from his own critique of language (*TLP*: 4.0031).

The *Tractatus* also influenced the style and presentation of Hayek's 1952 *The Sensory Order: Foundations of Theoretical Psychology*, which adopted a less complicated decimal numbering system. Here Hayek pursued one of his earliest intellectual interests from his time at the University of Vienna—a Kant-influenced account of how the mind imposes order onto existence—and which he allowed himself only after he had achieved success with the publication of *The Road to Serfdom* (Hayek 2017: 115-116; Ebenstein 2003: 127-128). But even beyond *The Sensory Order*, Wittgenstein's emphasis on language generally and the project of the *Tractatus* specifically influenced Hayek's work. A striking parallel can be found in Hayek's "The Confusion of Language in Political Thought," which argues that political thought is constrained by political language because our political language smuggles in outdated, misused, and anthropomorphized language, and with it faulty assumptions and premises about the science and its subject (see Hayek 1968). The remedy for Hayek is a critique of language. Here Hayek offers the distinction between two types of social orders: *cosmos*, denoting spontaneous and self-regulating orders, and *taxis*, denoting more limited and human organized orders (Hayek 1968: 10-14). Hayek's position that natural language is a cosmos—a spontaneous order—shares much with Wittgenstein, especially his posthumously published writings (see also Hayek 2021).

Ebenstein (2001 & 2003) has sought to establish a deep antagonism between Hayek's epistemology and its consequences for his understanding of language and that of the *Tractatus*. Ebenstein identifies a conflict between Hayek's view that "knowledge encompasses more than can be expressed in words" and a reading of the *Tractatus*—primarily *TLP* 6.53 & 7—wherein "knowledge means being able to describe something in sensible terms" (Ebenstein 2003: 129). Just like Ebenstein's mistaken over-emphasis on the affinity between Wittgenstein and the Vienna Circle's reading of the *Tractatus*, it is unclear that the antagonism that Ebenstein identifies is warranted. A first response would be to note that Wittgenstein has a broader understanding of knowledge than Ebenstein's reading allows. For example, Wittgenstein acknowledges that natural language is successfully used by humans without knowing its origins and mechanics (*TLP*: 4.002). Further, language is not uniform in the *Tractatus*; the text offers the well-known tripartite distinction between sense-full (*sinnvoll*), senseless (*sinnlos*), and nonsensical (*unsinnig*) language. Since

knowledge claims can be made about more than just sense-full language, Ebenstein's reading that knowledge claims can only be attributed to what can be expressed by factual description is overly restrictive. Lastly, there is good reason to worry that Ebenstein's reading of the *Tractatus* as offering a clear theory that can be used to reject wholesale certain natural language utterances as nonsensical is overhasty (*TLP*: 5.5563).

Hayek was not only committed to a critique of political language; he thought this was needed in the sciences generally. Hayek framed the issue with regards to institutions, and he argued that the sciences often suffered from improper institutional transplant—similarly to how political language smuggled in harmful conceptions of political institutions. He called this the problem of scientism. “The scientific attitude...involves *a mechanical and uncritical application of habits of thought to fields different from those in which they have been formed*” (Hayek 2018: 80, emphasis added; compare with *PI*: 41-47). Hayek concludes that, not only were the institutions of natural sciences not well-designed for the social sciences, they were not even well-designed for the natural sciences (Hayek 2018: 79). The health of the sciences, thus, demanded a critique of the very language used to *do* science. The language, habits, and methodologies of a science must be developed within the field in which they are deployed.

3. Spontaneous Order and the Whole Hurly-Burly of Life

Hayek is not commonly read as a philosopher of science. In fact, his lack of a formal philosophical background created opposition to his proposed biography of Wittgenstein (Erbacher 2014: 76). Yet he pursued decidedly philosophical questions with regards to the status and methods of the social sciences. It is only very recently that any attention has been directed to Hayek's influence in the philosophy of science (Scheall 2024). Hayek is treated as more removed from the philosophy of science than warranted, and his relationship with Wittgenstein is under-explored.

Hayek came to distinguish the natural sciences from the social in terms of “spontaneous orders,” introduced in “The Use of Knowledge in Society” (1945). Spontaneous orders cannot be neatly reduced to their component parts, and so mathematical laws of behavior are not possible here as they (supposedly) are in the physical sciences. This explained, for Hayek, the knowledge problem

that he observes in “Economics and Knowledge” (1937). The facts one would need to engineer solutions to social problems, if they even existed, were not accessible, and could not be assembled into comprehensive sets of true propositions.

Several connections can be drawn from Hayek’s work to the later writings of Wittgenstein. First, there is Wittgenstein’s acknowledgement of the messy context, scaffolding, or stage-setting in which phenomena are embedded, only against which they can be taken as meaningful (*PI*: 240).

How could human behaviour be described? Surely only by showing the actions of a variety of humans, as they are all mixed up together. Not what *one* man is doing *now*, but the whole hurly-burly, is the background against which we see an action, and it determines our judgment, our concepts, and our reactions. (*RFP* II: 629; see also *Z*: 567)

Not only is the “whole hurly-burly” the context of any meaningful signification, but much of this background—the preconditions of our experience—eludes our ability to account for it (*PI*: 242; *PPF*: 1; *OC*: 94-99; Stern: 255). The sheer profundity and proximity of these preconditions make them, as it were, invisible to us. This leads to an honest acknowledgment of human finitude and its methodological consequences (see Braver: 223-239). Wittgenstein and Hayek both express impatience with the hubris of much of philosophy and the sciences (*PI*: 299; Glock: 341). Such approaches assume a particular picture or model, trapping us in particular ways of thinking, unable to consider alternatives.

In his later work, Wittgenstein anticipated Thomas Kuhn’s scientific paradigms with his notion of means, or forms, of representation, which refers to the systems in which scientific descriptions get their meanings, and which regulate their usage (*AWL*: 40; *OC*: 512-516; Glock: 343). Each system depends upon key explanatory norms to which all scientific descriptions within it must adhere. For example, Newtonian physics and Freudian psychoanalysis each offer norms for explaining and accounting for phenomena (Glock: 343-344). As new scientific discoveries are made, older forms of representation can be found lacking, and new ones must be developed. Though ultimately conventionalist with respect to how supposed scientific truths are conveyed,

the supposed truths that these forms represent are not necessarily invalidated; rather, it is how these truths are made meaningful that has altered. Further, certain forms of representation can be so deep-rooted that they are taken as unquestionable until questioning becomes a possibility, perhaps as a result of a fundamental adjustment in our form of life (OC: 292-298; Glock: 344). This foreshadows Kuhn's notion of "normal science" (Kuhn 1962: 23).

Neither this notion of forms of representation, nor anything in Hayek, quite reach Kuhn's concept of scientific revolutions. Yet, Hayek foreshadows the central notion of the 'hard core' of hypotheses that characterizes the research program view of the sciences advanced by Imre Lakatos in his attempt to reconcile the work of Popper and Kuhn (Gray 1998: 19). Popper's positions here are formed in conversation with his close friend Hayek. Lakatos might provide a way of adapting Wittgenstein and Hayek to reflect a Kuhnian view of normal science, and scientific revolutions as shifts in the prevailing language of the sciences.

While this account of Hayek and Wittgenstein's shared influence on the Popperian line of philosophy of science is somewhat speculative, the influence of both Hayek and Wittgenstein on Lakatos's major critic, Paul Feyerabend, is better documented. While Lakatos wants to reconcile Popper and Kuhn, Feyerabend insists that the notions of a hard core on the one hand, or on normal science on the other, are misguided. Feyerabend, who originally planned to study under Wittgenstein, was taken in by Popper on Wittgenstein's death (see Erbacher 2020: 11-15). In addition to his explicit interest in the work of Wittgenstein, Feyerabend owes his polycentric view of the sciences to Hayek, with whom he studied in Vienna. Feyerabend notes that while much normal science is done by insiders, major innovations are owed to iconoclastic outsiders who challenge accepted language, method, and norms and defy "normal" science (see Feyerabend 2010).

Hayek and Wittgenstein influence both horns of this debate in philosophy of science. While the Popperian tradition orbits primarily around the nature of knowledge and an opposition to positivism, the alternative, embodied by Feyerabend, focuses on the distribution of knowledge, sociality, and our relationship with and loyalty to accepted forms of expression and meaning. Central in all of this is Hayek, who works at the intersection of Wittgenstein

and Popper's thinking. As such, Hayek and Wittgenstein should be taken seriously as philosophers of science, who draw on diverse philosophical concerns to inform their thinking about the sciences and language. The two had a profound impact on the thinking of the prominent philosophers of science that followed them. It is unfortunate that their roles in shaping philosophy of science have been largely overlooked.

Conclusion

The commonalities between Wittgenstein and Hayek can be seen not only in their familial connection and Viennese origins but also in their motivating concerns, their prioritization of critique of language, their methodological contextualism, and their influence on the history of twentieth century philosophy of science. Wittgenstein understood language to elude being a proper object of thought. It is inexhaustible, and any investigation into it is unavoidably self-reflexive; it entails a critique of language. Hayek understood the social sciences in a similar way—not only because they are inexhaustible and an investigation into society is necessarily undertaken within the context of society, but also because investigation into the social sciences entails a critique of language. Further, for Hayek, to responsibly undertake any scientific endeavor requires first a critique of the language used in that scientific endeavor. To fail to acknowledge the character and limits of language, and the consequences of these for scientific methodology, is a hubristic error. These thematic throughlines in the thought of Wittgenstein and Hayek are not only helpful for better reading those authors, they offer insight into the influence of these thinkers on twentieth century philosophy of science.

References

- Braver, Lee (2012), *Groundless Grounds: A Study of Wittgenstein and Heidegger*, Cambridge: the MIT Press.
- Bronzo, Silver (2012), "The Resolute Reading and Its Critics: An Introduction to the Literature," *Wittgenstein-Studien*, Vol. 3, 45-80.
- Caldwell, Bruce & Hansjörg Klausinger (2021), F. A. Hayek's genealogy, CHOPE Working Paper, No. 2021-06, Duke University, Center for the History of Political Economy (CHOPE), Durham, NC.

Christensen, Anne-Marie Søndergaard (2024), *Wittgenstein and Ethics*, Cambridge: Cambridge University Press.

Citron, Gabriel (2019), "Honesty, Humility, Courage, & Strength: Later Wittgenstein on the Difficulties of Philosophy and the Philosophical Virtues," *Philosophers' Imprint*, Vol. 19, No. 25.

Colgan, Jeffrey Patrick (forthcoming 2025), "Translating the *Tractatus* and Tractarian Ethics" in Passinsky, Asya, Julio De Rizzo, and Benjamin Schnieder, eds. *Facets of Reality — Contemporary Debates: Proceedings of the 45th International Wittgenstein Symposium*, Berlin: De Gruyter.

Ebenstein, Alan (2001), *Friedrich Hayek: A Biography*, New York: Palgrave MacMillan.

Erbacher, Christian (2014), "A Note on the History of F.A. von Hayek's Unpublished 'Sketch of a Biography of Ludwig Wittgenstein'," *The Wittgenstein Archives at the University of Bergen*, 76-79.

Erbacher, Christian, ed. (2020), *Friedrich August von Hayek's Draft Biography of Ludwig Wittgenstein: The Text and Its History*, with an afterword by Allan Janik, Leiden: Brill Mentis.

Feyerabend, Paul (2010), *Against Method: Outline for an Anarchistic Theory of Knowledge*. Verso Books.

Flowers, F. A., & I. Ground, eds. (2018), "Remembering My Cousin, Ludwig Wittgenstein" in *Portraits of Wittgenstein*, London: Bloomsbury Academic, 49-52. See also Hayek, Friedrich A. (1977), "Remembering My Cousin, Ludwig Wittgenstein" *Encounters* 49, no. 2, 20-22.

Glock, Hans-Johann (1996), *A Wittgenstein Dictionary*, Oxford: Blackwell.

Gray, John (1998) *Hayek on Liberty*, third edition, London: Routledge.

Hayek, Friedrich A. von (1968), "The Confusion of Language in Political Thought, With Some Suggestions for Remedying It," *Occasional Paper* 20, The Institute of Economic Affairs.

Hayek, Friedrich A. von (2017), *The Sensory Order and Other Writings on the Foundations of Theoretical Psychology*, edited by Viktor Vanberg, Chicago: The University of Chicago Press.

Hayek, Friedrich A. von (2018), *Studies on the Abuse and Decline of Reason: Text and Documents*, edited by Bruce Caldwell, Carmel: Liberty Fund.

Hayek, Friedrich A. von (2021), *Law, Legislation, and Liberty: A New Statement of the Liberal Principles of Justice and Political Economy*, edited by Jeremy Shearmur, Chicago: The University of Chicago Press.

Janik, Allen & Stephen Toulmin (1973), *Wittgenstein's Vienna*, New York: Simon and Schuster.

Kuhn, Thomas (1962), *The Structure of Scientific Revolutions*, Chicago: The University of Chicago Press.

Manzi, Jack (2020), Review of *Friedrich August von Hayek's Draft Biography of Ludwig Wittgenstein: The Text and Its History*, *Nordic Wittgenstein Review* 9, 254-256.

Monk, Ray (1990), *Ludwig Wittgenstein: The Duty of Genius*, New York: The Free Press.

Scheall, Scott (2024), "A Hayekian Social Science of Science," *Cosmos + Taxis*, Vol 12, No. 7-8, pp. 1-26.

Stern, David (1997), "Heidegger and Wittgenstein on the Subject of Kantian Philosophy" in *Figuring the Self: Subject, Absolute, and Others in Classical German Philosophy*, ed. by David E. Klemm & Günter Zöller, Albany: State University of New York Press, pp. 245-259.

Ware, Ben (2011), "Ethics and the Literary in Wittgenstein's *Tractatus Logico-Philosophicus*," *Journal of the History of Ideas*, Vol. 72, No. 4, pp. 595-611.

Wittgenstein, Ludwig (1972), *On Certainty*, edited by G.E.M. Anscombe & G.H. von Wright, New York: Harper Torchbooks. [OC]

Wittgenstein, Ludwig (1980), *Remarks on The Philosophy of Psychology*, vol. II, eds G.H. von Wright and H. Nyman, tr. C.G. Luckhardt and M.A.E. Aue. Oxford: Blackwell. [RFP]

Wittgenstein, Ludwig (1981), *Zettel*, translated by G.E.M. Anscombe, ed. by G.E.M. Anscombe & G.H. von Wright, Oxford: Blackwell [Z]

Wittgenstein, Ludwig (2009), *Philosophical Investigations*, fourth edition, Chichester: Blackwell. [PI]

Wittgenstein, Ludwig (2009), *Philosophy of Psychology – A Fragment*, Chichester: Blackwell. [PPF]

Wittgenstein, Ludwig (2023a), *Tractatus Logico-Philosophicus*, translated by Michael Beaney, Oxford: Oxford University Press. [TLP]

Political obligation and feminist theory. A proposal.

Nicole Darat (Valparaíso, Chile)

Abstract

This article explores the relevance of political obligation within feminist theory, questioning the traditional obligation to obey laws in the context of unequal citizenship status for women. It focuses in the history of philosophical thought starting with Socrates' arguments in Plato's *Crito* and contrasting them with feminist critiques, particularly through the lens of Sophocles' *Antigone*. The discussion highlights the contributions of feminist theorists to the subject, like Hannah Pitkin, Carole Pateman, and Nancy Hirschmann, who emphasize the importance of disobedience and participatory democracy in understanding political obligation. The article critiques mainstream political obligation theory for their lack of engagement with feminist perspectives and proposes a feminist alternative that views political obligation as a collective practice of world-building. This approach underscores the necessity of both obedience and resistance, advocating for a participatory democracy where obligations are created through active engagement rather than passive acceptance. The article concludes by suggesting that political obligation, from a feminist standpoint, involves a commitment to building inclusive and just social practices that challenge existing power structures and relations.

Why should political obligation be a relevant issue for feminism? If we define political obligation, at least preliminarily, as the obligation to obey laws, the answer to this question is far from obvious. Feminist theory has pointed out several critical points on the subject, mainly through what we know as critical theories of democracy and citizenship. Perhaps the most relevant of them points to the unequal citizenship status of women. What kind of obligation do those who do not enjoy equal rights have? Are they properly obligated or do they simply obey? If they are equally obligated, where does this obligation come from? Is it identical to that of those who enjoy full rights? Concepts such as "sexual citizenship" (Evans, 1993) "feminist citizenship" (Dietz, 1998; Lister, 1997; Phillips, 1991, Reverter, 2011) or "counter-citizenship", together with the critiques of deliberation and political representation or the traditional conception of consent, as assumed by modern and contemporary theories of the social contract, are part of the conceptual apparatus with which feminism has approached either directly or indirectly.

The question of political obligation has a long history in philosophy, we can point to Plato's *Crito* as the first problematization of the question in political philosophy. In this dialogue, Socrates discusses his obligation to abide by the death sentence against him, instead of agreeing to flee as his friend Crito proposes. The first argument given by the laws, is that he has resided for a

long time in Athens and that he has agreed with the laws and has pledged to obey them. This is usually seen as anticipating the argument of consent in contract theories. Mainly Locke's, where mere permanence in a territory counts as tacit consent. In fact, in 52c, the laws tell Socrates that he has given more signs than other citizens that the laws of Athens are satisfactory to him, since he has only left the city for a festival or to fulfill his military duties, he has never wanted to go and live under other laws. Another aspect of Socrates' reasoning often overlooked is that the laws remind Socrates, as proof of his consent, that he has had children in the city by adding that "you were certainly well in it".

A second argument put forward by Socrates points to the gratitude he has towards Athens, for owing his birth, upbringing, and education, to it, at this point Socrates uses an analogy with the obligation of obedience that children owe to their parents, referring to membership (to a political community) as grounding for political obligation. The third argument that Socrates gives points out that disobeying the sentence would be parasitizing his fellow citizens, since he would have enjoyed the benefits that compliance with the law entails, but ignored the burden that this implies when the time comes to comply with it himself. According to Richard Dagger (2021), this line of argument coincides with the so-called argument of justice or *fair play* as a justification of political obligation. The last argument is of a utilitarian nature, since Socrates makes the laws say: "Does it seem to you that the city can still exist without being ruined in which the judgments that are produced have no effect, but are invalidated by private individuals and are annulled?" (50b)

None of these arguments are developed in the dialogue. However, let us remember that Socrates considers his conviction unjust and believes that Athens should pay him a pension like outstanding athletes, just as he has stated in the *Apology*. But for him, to live a good life means to live justly, and to live justly consists of complying with the laws, regardless of their content. If they are considered unjust, one should try to persuade the judges, but flight and, therefore, disobedience, are not an option compatible with a good life. To live according to his principles has a greater value than to remain alive for life's sake. This desire to live an examined life is the main motivation to

Socrates according to readers like Russell Bentley (1996), more than the mere authority of law, unlike scholarship has repeatedly interpreted Socrates' argument to Crito.

It is probably Sophocles' Antigone where the supposedly shared vision of what is a life worth living, is put to the test. This work has been the object of feminist analysis and criticism, mainly to answer the reading of Hegel ([1807] 2010 and [1820] 2004) who, based on his interpretation of the tragic character, will justify the exclusion of women from the public space. It is Luce Irigaray (1985) who for the first time tears Antigone from the androcentric reading and reclaims her as a symbol of sexual difference and maternal genealogy. Its political interpretation as a figure of disobedience, meanwhile, has been notably recovered by authors such as María Zambrano (1989), Judith Butler (2000) and Bonnie Honig (2013), who, based on Antigone's act of disobedience, reconsider the question of obedience to the law and the relationship with the State. The relationship of feminist theory with political obligation seems closer to Antigone's challenge than to Socrates' obedience. From a feminist perspective, the good life doesn't come from obedience to the laws, but from disobedience to an order that insistently leaves them outside the category of the human.

If we consider Richard Dagger and David Lefkowitz entrance "Political Obligation" to the Stanford Encyclopedia of Philosophy as the mainstream of political obligation theory, it can also be read as an overview of the *malestram* (O'Brien, 1981) and this because, although feminist contributions are mentioned, they are not discussed. Although feminist theoretical production on the subject is scarce it isn't inexistent, we can mention Hannah Pitkin (1965 and 1966), Carole Pateman (1973 and [1979]/ 1985) and Nancy Hirschmann (1992). Pitkin and Pateman are mentioned in a parenthesis regarding the anarchist critiques of political obligation, while Hirschmann is only mentioned in the bibliography, there is no mention of their work in the body of the text. Of these three, only Hirschmann introduces her theory as a feminist theory of political obligation. Pateman's, which predates her most important work *The Sexual Contract* (1988) is rather a reflection of the author's commitment to the critique of liberal democracy and the defense of participatory democracy. Hannah Pitkin, on the other side, although she made important contributions to feminist theory, such as her reading of Machiavelli in *Fortune is a Woman*

(1984) and in her critical approach to the concept of public sphere in Hannah Arendt, does not incorporate this perspective in this text.

Pitkin, Pateman and Hirschmann: the formation of a feminist theory of political obligation.

As we have mentioned before, a feminist theory of political obligation is also a theory of disobedience, and this aspect is central to the theories of Hannah Pitkin and Carole Pateman. The latter points out that, in liberal theories, such as Rawls', disobedience appears merely as a symbolic gesture and not as the real possibility of opposing the law as part of an exercise of political participation.

For Pitkin (1965 and 1966), a theory of obligation should answer these four questions:

- 1) The question of the limits of obligation: When are we obliged to obey and when are we not?
- 2) The question of the locus of sovereignty: Whom are we obligated to obey?
- 3) The Question of the Difference Between Legitimate Authority and Coercion: Is There Really a Difference Between Legitimate Authority and Coercive Power?
- 4) The question of the Justification of the obligation: Why are we obliged to obey even a legitimate authority?

Pitkin examines different theories of political obligation, from the theory of divine right to utilitarianism and, of course, the theory that bases political obligation on consent, namely the contractualist theory. The author identifies problems in each of them by answering these 4 fundamental questions. I am interested in dwelling on the critique she makes of the utilitarian perspective, and of the contractarian perspective. On utilitarianism Pitkin will distinguish between an individualistic and a social perspective. For the individualistic perspective, obligation depends on the amount of well-being that is derived from obedience and the decision to resist or not, in case it does not provide me with well-being, will depend on the result of the calculation between pleasures and pains that resistance would cause. From a social perspective, this

calculation should consider the entire political community. Bentham leans towards the latter. For Pitkin, the social version of utilitarianism performs particularly poorly in trying to answer the last question. The only reason to obey a government, and that is the answer to the question of what makes it legitimate, is that it generates greater well-being for the political community as a whole. For Pitkin, this answer simply sidesteps the question of legitimacy or rather shifts it to the next one: Why should the well-being of society mean anything to me? The utilitarian response ends up being no better than the one that simply commands us to obey the law and authority.

Consent theory, meanwhile, explains the obligation from the act of consenting: you are obliged because you have consented, whether that consent is effective or tacit. Taking Locke and Tussman as an example, consent by itself is not enough, since, in order to be rational consent, it must be given only to "reasonable" forms of government. The question then is how we determine what are those characteristics that make some governments reasonable and others not. Pitkin will conclude that consent theory does not really base the obligation on the act of consenting, but on the characteristics of government, namely, those that make it reasonable and, therefore, legitimate. Legitimacy is given by reasonableness understood in terms of the Kantian hypothetical contract. This brings us to the fourth question, why are we obligated to obey even legitimate power?

To Pitkin it is necessary to think about political obligation in a different way. She will postulate that it is the social practice of promising that ultimately answers the question of obligation. Obligation and promise have a similar structure for the author. We learn the value of promises by participating in practices of promising and keeping promises, we are not given a promise and told "this is a promise" as if to say "this is a pencil". We learn the value of political obligation within social practices that give it meaning. Although for the author this does not constitute a single and definitive answer to the question, it can give directions on possible answers.

Pitkin will propose that obligation depends on two factors: one, conscience (a factor, so to speak, deontological) and the other, attention to how others see what we do. (a utilitarian factor). While the first factor is key in interpersonal relationships, the second is key for social life. It is up to the other people with

whom we share the world to judge whether our decision to obey or disobey is correct. This is why for Pitkin, there is not only a right of resistance, but also an obligation of resistance. If obedience is due when power is legitimate, the determination of that legitimacy, nor the answer to the question of when it ceases to be so, have a predefined content. The determination of legitimacy depends on social practices of acting and judging in common and, therefore, political obligation can never be an individual act.

For Carole Pateman, meanwhile, the political obligation must be horizontal (Rousseauian model), that is, towards our fellow citizens and not limited only to expression through voting. A genuine justification of political obligation should not be taken for granted but should be part of a participatory democracy. Later in her career, she would link this reading to feminist theory, both in *The Sexual Contract* (1988), and in some articles such as *Feminism and Democracy* (1983). For Pateman, as for Pitkin, political obligation cannot be understood without recognizing the role of resistance in it. Without the possibility of disobedience, there is no political obligation. Without a participatory democracy, there is no political obligation. It is participatory democracy that, in Pateman's opinion, would allow the free creation of obligations and not the mere passive acceptance of a state of affairs that could only be resisted in a testimonial way.

In the texts prior to *The Sexual Contract*, the defense of a model of participatory democracy against the realism of the theories of democracy in use was the main object of Pateman's work. In *Participation and democratic theory* (1970) he defended the possibility of a participatory democracy, understood as one where participation is not only through voting, but there is participation in other spheres of social and political life. In fact, Pateman speaks of a participatory society, understanding that politics exceeds the scope of government. It will pay special attention to industry, a sector that theorists such as Schumpeter or Dahl consider unlikely to be democratized. Pateman, drawing both on G.D.H. Cole's theory and on concrete results of the implementation of workers' *self-management* in (now former) Yugoslavia, argues that the democratization of industry functions as a school for the development of the democratic capacities of the working class and that there would be no evidence of a decrease in productivity. the case of Yugoslavia, in fact, testified to the contrary. However, Pateman is cautious in pointing out

that there is little data and that it is not possible to draw universalizable conclusions, but the "realism" that assumed the apathy of the majorities and rejects participatory democracy as utopian, cannot be taken at face value either.

Hirschmann, on the other hand, by rescuing the given obligations, in contrast to consent, would seek to show the insufficiency of this model to account for the experience of women when assuming obligations. This experience is modeled by the author based on object relations theory, according to which women would have been raised mostly in a way that makes them remain closer to care and the private world of the mother, than to the contractual language, which for the author is linked to the public world, normally associated with masculinity. However, she will insist that she does not seek to defend given obligations as a normative ideal for politics, but to show that a good part of obligations (not only of women) are not explained by consent.

Obligation as a world building practice

Pateman and Pitkin pointed out the importance of disobedience as part of a theory of political obligation, and this becomes quite evident in the light of Lorde's words: "For the master's tools will never dismantle the master's house. They may allow us to temporarily beat him at his own game, but they will never enable us to bring about genuine change" obeying in a context of structural injustice can only deliver temporary victories. If there were such a thing as a lasting victory, it would have to come from disobedience, but, following Lorde, not only from the negative moment of disobedience, but from what comes after, from the construction of other sources of support, that exceed the master's game, that exceed inclusion in the institutionality (although not necessarily reject it).

Feminists of the Women's Bookshop in Milan (1990) set out to think of female authority in a different way: the *affidamento*, a practice of trust between women that is based on the collective recognition of one legitimate authority in another, and that, therefore, generates obligations. Feminist theory is part of what we do in these spaces of resistance, Milan's feminists referred to it as something that arose from their work together, either in the organization, or in simply being together.

Based on this brief review of theories of political obligation, I want to propose a possible feminist alternative, which involves the shift from the obligation to obey the laws, to the reciprocal obligation of a practice of world-building, an Arendtian expression that Linda Zerilli (2005) recovers for feminism. It is a non-sovereign action, we are not masters of it, and, therefore, it can never guarantee us a result. Bonnie Honig also puts this aspect of Arendt's agonism: "Action produces its actors; episodically, temporarily, we are his agonistic achievement" (Honig 1995:145).

For Macarena Marey it is about the "teratogenic" condition of the action, that is, "its uncontrollable character". Human action, Marey affirms, is in this sense, monstrous:

The monstrosity of action is inescapable, it is a fact of our condition as agents, of our way of acting in the world. The point, however, is that those who do not know that their action is teratogenic are more likely to contribute to the destruction. Ignoring responsibility is the epistemic and moral deficit that destructive violence generates. On the contrary, knowing (although always uncertainly) the combinations of contingency and necessity in action allows us to act creatively in the face of each attempt at annihilation (2022: 59).

The proposal I argue here is to understand political obligation from feminism as an obligation to build worlds, that is, the obligation to act with and for others. This obligation is close to Nancy Hirschmann's idea of non-voluntary obligations, we have not explicitly given our consent to these obligations, however, we have a responsibility that is not individual, but collective. The construction of this world implies vigilance and contestation of existing authority, as well as obedience to it if deemed necessary to guarantee the conditions that allow its persistence. The determination of legitimacy, as Pitkin points out, cannot be done individually, but corresponds to collective practices, of construction and contestation. However, no definitive answer can be given as to why obey even legitimate authority. The answer, as Pitkin states, is always in collective practice.

References

Butler, Judith (2002) *Antigone's Claim: Kinship Between Life and Death*, New York: Columbia University Press.

Dagger, Richard; Lefkowitz, David (2021) "Political Obligation", *The Stanford Encyclopedia of Philosophy*, Edward N. Zalta (ed.) URL = <<https://plato.stanford.edu/archives/sum2021/entries/political-obligation/>>.

Dietz, Mary (1998) "Citizenship with a Feminist Face: The Problem with Maternal Thinking", in: Joan B. Landes (ed.) *Feminism, The Public And The Private*, Oxford: Oxford University Press.

Evans, David (1993) *Sexual Citizenship: The Material Construction of Sexualities*, New York: Routledge.

Hegel, Georg Wilhelm Friedrich (2010) *Fenomenologia Del Espiritu* (Ángel Ramos, trans.) Madrid: Abada.

Hegel, Georg Wilhelm Friedrich (2004) *Principios De La Filosofía Del Derecho, O, Derecho Natural Y Ciencia Política*. (Juan Luis Vermal, trans.) Buenos Aires: Sudamericana.

Hirschmann, Nancy (1992) *Rethinking Obligation: A Feminist Method for Political Theory*, Ithaca NY: Cornell University Press.

Honig, Bonnie (1995) "Toward an agonistic feminism", in: Bonnie Honig (ed.) *Feminist Interpretations of Hannah Arendt*, Pennsylvania: Penn State University Press, 135-166.

Honig, Bonnie (2013) *Antigone, Interrupted* (Illustrated edition), Cambridge: Cambridge University Press.

Irigaray, Luce (1985) *Speculum of the Other Woman* (Gillian C. Gill trans.) Ithaca NY: Cornell University Press

Lister, Ruth (1997) "Citizenship: Towards a Feminist Synthesis", *Feminist Review* 57(1), 28-48.

Lorde, Audre (2007) "The Master's Tools Will Never Dismantle the Master's House", in: Audre Lorde, *Sister Outsider: Essays and Speeches*, New York: Crossing Press, 110-113.

Marey, Macarena (2022) "La dialéctica de la responsabilidad", *Bordes* 7 (2), 57-61.

Milan Women Bookstore Collective (1990) *Sexual Difference: A Theory of Social-Symbolic Practice*, Bloomington IN; Indiana University Press.

O'Brien, Mary (1981) *The Politics of Reproduction*, New York: Routledge.

Pateman, Carole (1970) *Participation and Democratic Theory*, Cambridge: Cambridge University Press.

Pateman, Carole (1985) *The problem of political obligation: A critical of liberal theory*, Berkeley: University of California Press.

Pateman, Carole (1988) *The Sexual Contract*, Stanford: Stanford University Press.

Pitkin, Hannah (1965) Obligation and Consent I, *The American Political Science Review* 59(4), 990-999.

Pitkin, Hannah (1966) Obligation and Consent II, *The American Political Science Review* 60(1), 39-52.

Pitkin, Hannah (1984) *Fortune is a woman: Gender and politics in the thought of Niccolò Machiavelli*, Chicago: University of Chicago Press.

Phillips, Anne (1991) "Citizenship and feminist theory", in: Geoff Andrews (ed.) *Citizenship*, London: Lawrence & Wishart, 76-88.

Platón (2020) "Critón", in: *Critón. El político*, (Casadesús Bordoy, trans.) Madrid: Alianza, 9-52.

Reverter Bañón, Sonia (2011) "La dialéctica feminista de la ciudadanía", *Athenea Digital. Revista de pensamiento e investigación social*, 121-136.

Sófocles (1992) "Antígona", in: *Tragedias*, (Assela Alamillo, trans.) Madrid: Gredos, 239-299.

Zambrano, María (1989) *La tumba de Antígona*, Barcelona: Anthropos.

Zerilli, Linda (2005) *Feminism and the Abyss of Freedom*, Chicago: University of Chicago

Acknowledgement

I am grateful for the support of the FONDECYT 1230985 project funded by the National Agency for Research and Development (ANID) in Chile.

Updated version available at the end of the paper

AI ethics and the diversity of forms of life

Michiel De Proost (Brussels, Belgium)

Jonathan Adams (Oslo, Norway)

Abstract

There is a growing literature that tries to lay out aspects of normativity in the field of AI ethics that cannot be sufficiently captured within a system of principles. For instance, Heuser, Steil, and Salloch introduce the idea of human life forms to highlight the role of moral judgment, which they argue cannot be bypassed when formulating ethical approaches to AI in the healthcare sector. However, Heuser, Steil, and Salloch overlook that life forms are not uniform and can conflict. This paper aims to explore the diversity of forms of life. In particular, it will raise attention to the possibility of oppressive forms of life and argue that this possibility generates challenges for the strategy of social critique.

Introduction

There is a growing literature that tries to lay out aspects of normativity in the growing field of AI ethics, which cannot be sufficiently captured within a system of principles (Corrêa et al., 2024; Heuser, Steil and Salloch, 2025). As Hamsphire (1978: p. 53) indicated in the seventies “moral theory cannot be rounded off and made complete and tidy ... partly because new ways of life should always be expected to arise in association with new knowledge and with new social forms.” Moreover, the uncodifiability of moral principles and artificial moral agents is getting more attention (Furlan, 2024; Graff, 2024; Villegas-Galaviz & Martin, 2024). We should take care that principlist reasoning does not end up overfocusing the harsh spotlight of (old and new) moral principles and instead pay attention to the softer and subtler lights of moral life. Heuser, Steil, and Salloch’s (2025: p. 3) recent publication is the most extensive to introduce the idea of ‘human life forms’ to highlight the role of moral judgment, which “cannot be circumvented as a basis for developing alternative models for the ethical guidance of AI in healthcare”. This publication serves as the principal target of our critique in this paper.

Heuser, Steil and Salloch define life forms as, based on the work of Jaeggi (2018: p. 50 [emphasis in original]), “forms of life are nexuses of practices, orientations, and orders of social behavior. They include attitudes and habitualized modes of conduct with a *normative character* that concern the *collective conduct of life*, although at the same time they are *not strictly codified* or institutionally binding”. Other reference points, according to Heuser, Steil,

Updated version available at the end of the paper

Updated version available at the end of the paper

and Salloch, that regard human life forms as normatively relevant are Kant's (1913) critique of judgment and Wittgenstein's (1976) work on rule-following. Like actions within a game, on the account from Wittgenstein that these authors cite, participation in a form of life aims to achieve a goal by following shared rules. However, Heuser, Steil and Salloch put the term "life-world" at the centre of their title. Although we are more interested in ethical problematization than in exegetical accuracy and recognize that any attempt to learn from another writer is always likely to involve a certain eclecticism, there is a question as to how far one can go with their use of this concept. Neither did Wittgenstein use the concept of 'Lebenswelt' (or 'life-world') in his work, nor did Jaeggi introduce it in her work. Indeed, Wittgenstein scholars have explicitly contrasted the philosopher's notion of *Lebensform* ('life form') with phenomenologists' concurrent usage of *Lebenswelt* in order to emphasize crucial aspects of his work (Floyd, 2020).

Moreover, it is peculiar that other methodological commitments in Jaeggi's work are not touched upon, such as the tradition of critical theory and the possibility of immanent critique, especially if we take into account their worry about the diversity of life forms and that "racism and discrimination ... should not be left only to technological solutions, but also need to be addressed politically" (Heuser, Steil, and Salloch, 2025: p. 7). As Stahl (2024: p. 78) indicates, "many forms of life are deeply shaped by white supremacy, sexism, the exclusion of non-traditional gender identities, and class differences and will thus count, to a greater or lesser degree, as oppressive in this sense." While Heuser, Steil, and Salloch do not spend time themselves discussing the relationship between forms of life and oppression in the context of AI ethics, this paper assumes that questions raised in feminist ethics and a Wittgensteinian approach to feminist epistemology provide extra tools to critically examine such relationships in AI ethics.

The remainder of this paper proceeds as follows: First, the importance of crises in life forms is sketched out. Second, an argument is offered in favor of identifying which groups can participate in identifying crises in life forms. Finally, the paper concludes with an emphasis on the importance of further investigating storytelling in AI ethics.

Updated version available at the end of the paper

Updated version available at the end of the paper

Crises and forms of life

Heuser, Steil, and Salloch draw on classic debates in moral and political philosophy concerning the role of ethical principles and the relationship between justice and the good life, but this means that standard objections to mainstream liberal ideal theorizing from critical theorists are left out. For example, critical theorists such as Jaeggi often argue that many liberal theories presume an unrealistically idealistic social ontology that views social institutions as cooperative arrangements emerging from a preceding consensus, yet this is not touched on. The AI ethicists' labors cannot stop by just introducing a new concept in the debate but must engage with the conceptual and political baggage that it brings to justify its role in ethical inquiry. Moreover, Heuser, Steil, and Salloch ignore that Jaeggi's model of social practice takes crises as reference points to diagnose the ethical failures of social arrangements: "The moment of crisis forces reflection on and adjustments of practices—a re-creation of practices—that were previously taken for granted" (Jaeggi, 2017: p. 167). To get the full picture of life forms in AI ethics, such moments of crisis must be explored further.

Jaeggi was not the first to emphasize the importance of unstable life forms in moral philosophy, and a similar idea has been developed by several feminist moral philosophers in a Wittgensteinian spirit, including Walker (2007) and Alcoff (1991). Questions of moral and epistemic authority matter particularly in this critical tradition because of the historic denial of them to marginalized groups, such as women and/or people of color. It seems crucial, therefore, to ensure that analyses of life forms in the context of humans and AI systems do not overlook the multiplicity, complexity, and multi-dimensionality of possible positions within forms of life. This is especially urgent if we take into account the existing oppressive forms of life (Stahl, 2024), i.e. social processes through which the self-understanding inherent in the life form is distorted by relations of domination and exclusion contributing to marginalization. Indeed, the extent to which AI narratives inform and engage with issues of gender and 'whiteness' (Katz, 2020) are of ongoing concern and debate in wider attempts to decolonize AI ethics (Mhlambi & Tiribelli, 2024). Ideologies and related experiences of domination get in the way of perceiving many of the moral and political problems we face in AI ethics. Feminists want to peel this layer of distortion away or at least illuminate the messy realities, with the aim of

Updated version available at the end of the paper

Updated version available at the end of the paper

understanding the often-conflicting needs that find expression in words and concepts.

Whose crisis?

Heuser, Steil, and Salloch often refer to what “we” do and to “our” practices in AI development, but that appearance of social homogeneity is misleading. The questions of “Whose problems?” and “Whose crisis?” are ones that the analysis of life forms in AI ethics cannot avoid. As Fricker (1999: p. 207) argues: “With the possibility of such a question, we confront the disunity of the ‘we’ who agree in form of life. If the different ‘we’s within a form of life stand to one another in relations of advantage and disadvantage, power and powerlessness, then this inequality is likely to be reiterated in interpretive practice.” Fricker’s conviction that attention to who is knowing, from where, and with what stakes, is crucial to supporting, rather than undermining, a usable conception of moral knowledge, is shared by other feminist philosophers (Walker, 2007; Scheman, 2022).

Two possible mechanisms can be identified that affect the way in which groups and their ethos of “we” can participate in the hermeneutical processes through which a life form either identifies or does not identify problems as constituting crises. The first kind of mechanism is defined by Stahl (2024) as “a form of exclusion from the administration of the hermeneutical resources of a life form”. In other words, the unfair exclusion of knowers from participation in the social processes by which a collectively binding self-understanding is determined. For instance, Miragoli (2024) describes the case of Irina, who is experiencing feelings for girls and, in an attempt to understand her emotions better, goes on Google. The conformist behavior displayed by Google’s algorithms, which tend to read and interpret input information in the light of categories extracted from dominant heteronormative trends, is not helpful to Irina in making sense of her own sexual experience.

Part of the challenge of such algorithmic conformism is that forms of life change profoundly over time, whereas AI-driven decision-making systems are prone to what Hardalupas (2024) calls “epistemic calcification,” whereby (potentially conformist) interpretive frameworks become fixed. Hardalupas argues that AI systems, due to human input, biased data, and wider institutional inertia, embed certain assumptions about health and disease, for

Updated version available at the end of the paper

Updated version available at the end of the paper

example, to the exclusion of other perspectives, leading to a kind of epistemic oppression called “contributory injustice” (Dotson, 2012: p. 25). If technologies are capable of embedding – or even, as Winner (1986) suggests, constituting – forms of life, then medical AI systems’ classifications may be not only unhelpful but intractably so, with critiques made by patients failing to impact their functioning. At their most radical, AI systems may take part in what Winner refers to as “technological somnambulism,” whereby “we [whoever that ‘we’ may be] so willingly sleepwalk through the process of reconstituting the conditions of human existence” (1983: p. 10). Taking this metaphor further, awaking AI developers to the perspectives of the marginalized demands neither the subtle brush of a principle-based approach nor some faint awareness of a voice from the outside but a full-body wake-up in the form of fundamental critique. While “the recognition of the impact of life-worlds and joint procedures of interpretation” that Heuser et al. (2025: p. 12) propose is necessary, it is only a first step in the direction of such critique.

It is important to note that this first kind of mechanism concerns the extent to which subordinated groups have an opportunity to bring their personal experiences of the life form’s failure to achieve its ascribed purposes into the collective consciousness. Even if subordinated groups face serious constraints in this respect, this will not preclude the development of an individual or subcultural consciousness of crisis. As Jaggar (1983: p. 370) highlights, the suffering of oppressed groups can be epistemically beneficial: “Their pain provides them with a motivation for finding out what is wrong, for criticizing accepted interpretations of reality and for developing new and less distorted ways of understanding the world.” This has also been highlighted in more recent epistemic injustice literature, which indicates how marginalized communities often possess fully fledged concepts to make sense of oppressive experiences as well as fully fledged expressions that communicate them but lack acknowledgment on the part of the dominant culture (Medina, 2013). Consequently, outsiders who are offered space within dominant forms of life “are in positions to be listened to and to be intelligible” (Scheman, 1996: p. 403).

A second kind of mechanism that can lead to the emergence of oppressive forms of life is best thought of as a form of ‘epistemic death’ (Medina, 2013). This occurs when subjects are not given even minimal amounts of credibility, are not treated as intelligible communicators, and are prevented from

Updated version available at the end of the paper

Updated version available at the end of the paper

participating in meaning-making and meaning-sharing practices. Under such conditions of epistemic death, subjects owe nothing to those communities, and they may have a right to fight epistemically by any means necessary (including the right to lie, to hide, to sabotage, to silence others, etc.), demonstrating loyalty and solidarity only with alternative epistemic communities (communities of resistance). We can identify similar practices related to AI that interrupt and disrupt the established epistemic economy of a society and its practices and institutions (Bonini & Treré, 2024; Lehuedé, 2024).

As well as the ways in which forms of life disadvantage their own members, oppression can be identified in one's being excluded, or expelled, from the forms of life to which others are granted access. Yet this does not mean that there are no legitimate boundaries to be drawn around distinct forms of lives, which are warranted in part due to the distinct vantage points on life and morality that they afford us. On Lindemann's (2014: p. 19) Wittgensteinian account, personhood – interpreted as a form of moral consideration – is restricted to those entities that “share in our form of life” as human beings and can “be brought into personhood” by those around them, which is limited to humankind. This excludes not only non-human animals but also highly humanlike AI systems from participation in our forms of life, against the assumptions of some philosophers of robotics (Schmid, 2017). With this exclusion comes the implication that AI systems are unable to appreciate the norms internal to our shared forms of life and that, while ethical principles may be programmed into an AI system, awareness of the inner workings of human forms of life is out of their grasp. For a human being to “be able to participate skillfully in a form of life” (de Boer, 2023: p. 2272) means appreciating its distinctive normative demands, which can only be done from within those human forms of life that artificial agents (whether hypothetical or actual) are outside. However, the moral stakes in deciding who takes part in ‘our’ life forms remain high, and assumptions about the individuals and entities that they include and exclude cannot be taken for granted.

Further mapping the rough ground

Heuser, Steil, and Salloch (2025: p. 11) argue that “emphasis must be placed on stakeholders’ good life in its various forms and practices” but do not recognize that stakeholders are differentiated epistemically according to their social location: what might be reasonable to believe from one location may very well

Updated version available at the end of the paper

Updated version available at the end of the paper

be unreasonable from another. One response to the challenge of better understanding individuals in context is to suggest that we situate them through stories: by telling stories we reveal different life forms (Schachtner 2020). As Walker (2007: p. 72) indicates, “[b]ecause people and their relationships are not uniform and situations are not necessarily repeatable, moral consideration on this view presses toward enrichment of detail and amplification of context.” To get the full picture of life forms in AI ethics, the critical value of AI narratives and responsible storytelling must at the very least be explored further (Coeckelbergh, 2023; Chubb, Reed, and Cowling, 2024). In other words, we need to go back to the “rough ground” of how words and narratives are actually used (Wittgenstein, 1976: §107).

The point is not just to collect new stories but to recognize their transformative power (Haraway, 2016). Through storytelling, we can uncover injustices, inspire collective action, and foster new types of solidarity (Jones & Fulfer, 2024). For example, Lindemann (2001) explains that storytelling can convert ‘master narratives’—typical stories representing dominant moral and understandings and related cultural views—into ‘resistance narratives’ that highlight and raise awareness of the harms that individuals suffer due to these master narratives. Inspired by Lugones (2003), several scholars have indicated how certain playful ways of using language can help marginalized individuals safely navigate dominant worldviews without being trapped in oppressive definitions and perceptions (Boncompagni, 2023). Alternative storytelling practices in AI might include how Indigenous-led researchers in New Zealand are using a machine learning-based speech-to-text system to help revive Te Reo Maori, the language of the Maori people (Papa Reo, 2024). Another example is discussed in the Radical AI Podcast of June 2021. An Indian feminist app called Mumkin is designed to have difficult conversations to address the issue of female khatna (also known as female genital cutting) within the Bohra community in India. The chatbot is intended to help users shed inhibitions and gain confidence to have real-life conversations on gender, culture, and society. Finally, the recent ethnographic research of Madianou (2024) reveals how ordinary practices of resistance to datafication and technological experimentation are becoming increasingly ubiquitous in the aid sector.

This highlights the importance of (self-)critique, along with the concepts of patience, slowness, and listening (Chadwick, 2024), in recognizing forms of

Updated version available at the end of the paper

Updated version available at the end of the paper

struggle and resistance. It also demonstrates that storytelling can enhance a (marginalized) individual's ability to act autonomously by creating opportunities for others to acknowledge and respect the stories that shape her world. Scheman (2009) insightfully describes this form of agency as “perceptual autonomy,” which necessitates that persons can express a “nonperiscope point of view” that hasn't been altered to align with a dominant perspective. The contestation of AI power begins with experiments in listening differently, for which bringing together work on forms of life with feminist philosophy is a valuable first step.

Conclusion

In conclusion, while many points that Heuser, Steil, and Salloch raise in their paper on AI ethics and their focus on the importance of studying forms of life are well taken, it is much less clear how their analysis of forms of life leaves room for reflection on oppressive positions and forms of resistance. Moreover, we suggested that messy storytelling should be more central in AI ethics. As Eagleton and Jarman (1993: p. 55) state in the script of the *Wittgenstein* film: “roughness and ambiguity and indeterminacy aren't imperfections—they're what make things work.” From this point, the important question about AI driven systems becomes, to paraphrase Winner (1986) with a nod to Walker, what kind of *world* are ‘we’, as a group that is “not uniform”, making?

References

- Alcoff, Linda (1991) "The problem of speaking for others", *Cultural critique* 20, 5-32.
- Boncompagni, Anna (2023) "Forms of life and linguistic change: the case of trans communities". *Philosophies* 8, 50.
- Bonini, Tiziano; Treré, Emiliano (2024) *Algorithms of resistance: The everyday fight against platform power*, Cambridge: MIT Press.
- Chadwick, Rachele (2024) "The question of feminist critique", *Feminist Theory* 25, 376-395.
- Chubb, Jennifer; Reed, Darren; Cowling, Peter (2024) "Expert views about missing AI narratives: is there an AI story crisis?", *AI & society* 39, 1107-1126.
- Coeckelbergh, Mark (2023) "Narrative responsibility and artificial intelligence: How AI challenges human responsibility and sense-making", *AI & society* 38, 2437-2450.

Updated version available at the end of the paper

Updated version available at the end of the paper

Corrêa, Nicholas Kluge; Santos, James William; Galvão, Camila; Pasetti, Marcelo; Schiavon, Dieine; Naqvi, Faizah; Hossain, Robayet & De Oliveira, Nythamar (2024) "Crossing the principle–practice gap in AI ethics with ethical problem-solving", *AI and Ethics*, 1-18.

de Boer, Bas (2023) "Explaining multistability: Postphenomenology and affordances of technologies", *AI & Society* 38, 2267-2277.

Dotson, Kristie (2012) "A cautionary tale: On limiting epistemic oppression", *Frontiers: A Journal of Women Studies* 33, 24-47.

Eagleton, Terrey; Jarman, Derek (1993) *Wittgenstein : The Terry Eagleton script. The Derek Jarman film*, London: British film institute.

Floyd, Juliet (2020) "Wittgenstein on ethics: Working through Lebensformen", *Philosophy & Social Criticism* 46, 115-130.

Fricker, Miranda (1999) "Epistemic oppression and epistemic privilege", *Canadian Journal of Philosophy Supplementary Volume* 25, 191-210.

Furlan, Timothy J (2024) "Principlism, uncodifiability, and the problem of specification", *Cambridge Quarterly of Healthcare Ethics*, 1-22.

Graff, Joris (2024) "Moral sensitivity and the limits of artificial moral agents", *Ethics and Information Technology* 26, 13.

Hampshire, Stuart. (Ed.). (1978). *Public and private morality*, Cambridge: Cambridge University Press.

Haraway, Donna J (2016) *Staying with the Trouble: Making kin in the Chthulucene*, Durham: Duke University Press.

Hardalupas, Mahi (2024) "Contributory injustice, epistemic calcification and the use of AI systems in healthcare", *Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society* 7, 573-583.

Heuser, Stefan; Steil, Jochen; Salloch, Sabine. (2025) "AI Ethics beyond Principles: Strengthening the Life-world Perspective", *Science and Engineering Ethics* 31, 7.

Jaeggi, Rahel (2018) *Critique of forms of life*. Cambridge: Belknap Press.

Jaggar, Alison (1983) *Feminist politics and human nature*. Totowa: Rowman and Allanheld.

Jones, Janet; Fulfer, Katy (2024) "Solidaristic Listening", *Hypatia* 39, 634-651.

Kant, Immanuel (1913) *Kritik der Urteilskraft*. Berlin: Reimer.

Katz, Yarden (2020) *Artificial whiteness: Politics and ideology in artificial intelligence*. New York: Columbia University Press.

Updated version available at the end of the paper

Updated version available at the end of the paper

Lehuedé, Sebastián (2024) "An elemental ethics for artificial intelligence: water as resistance within AI's value chain", *AI & Society*, 1-14.

Lindemann, Hilde (2001) *Damaged identities, narrative repair*, New York: Cornell University Press.

Lindemann, Hilde (2014) *Holding and letting go: The social practice of personal identities*, Oxford: Oxford University Press.

Lugones, Maria (2003) *Pilgrimages/peregrinajes: Theorizing coalition against multiple oppressions*, Lanham: Rowman & Littlefield Publishers.

Madianou, Mirka (2024) *Technocolonialism: When Technology for Good is Harmful*, Hoboken: John Wiley & Sons.

Medina, José (2013) *The epistemology of resistance: Gender and racial oppression, epistemic injustice, and the social imagination*. Oxford: Oxford University Press.

Mhlambi, Sábêlo; Tiribelli, Simona (2023) "Decolonizing AI ethics: Relational autonomy as a means to counter AI harms", *Topoi* 42, 867-880.

Miragoli, Martin (2024) "Conformism, Ignorance & Injustice: AI as a Tool of Epistemic Oppression", *Episteme*, 1-19.

Papa Reo (2024) "Papa Reo", <https://papareo.nz>

Schachtner, Christina (2020) *The narrative subject: Storytelling in the age of the Internet*, Cham: Palgrave Macmillan.

Scheman, Naomi (1996) "Forms of life: Mapping the rough ground", in: Hans Sluga and David G. Stern (eds.) *The Cambridge Companion to Wittgenstein*, Cambridge: Cambridge University Press, 389-416.

Scheman, Naomi (2009) "Narrative, complexity, and context: autonomy as an epistemic value", in: Hilde Lindemann, Margaret Urban Walker, and Marian Verkerk (eds.) *Naturalized bioethics: toward responsible knowing and practice*, Cambridge: Cambridge University Press, 106-124.

Scheman, Naomi. (2022) "Disrupting Demands: Messy Challenges to Analytic Methodology", *Journal of Social Philosophy* 53, 473-493.

Schmid, Hans Bernhard (2017) "'Robot' as a life-form word", In R. Hakli and J. Seibt (eds.), *Sociality and normativity for robots: Philosophical inquiries into human-robot interactions*, Cham: Springer, 233-250.

Stahl, Titus (2024) "Oppressive Forms of Life", *Critical Horizons* 25, 77-93.

Updated version available at the end of the paper

Updated version available at the end of the paper

Walker, Margaret Urban (2007) *Moral understandings: A feminist study in ethics*, Oxford: Oxford University Press.

Winner, Langdon (1986) *The whale and the reactor: A search for limits in an age of high technology*, Chicago: University Of Chicago Press

Wittgenstein, Ludwig (1976). *Philosophical investigations*. Oxford: Basil Blackwell.

Villegas-Galaviz, Carolina; Martin, Kirsten (2024) "Moral distance, AI, and the ethics of care", *AI & society* 39, 1695-1706.

Acknowledgement

This research has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No 949841– DIME).

Updated version available at the end of the paper

AI ethics and the diversity of forms of life

Michiel De Proost (Brussels, Belgium)

Jonathan Adams (Oslo, Norway)

Abstract

There is a growing literature that tries to lay out aspects of normativity in the field of AI ethics that cannot be sufficiently captured within a system of principles. For instance, Heuser, Steil, and Salloch (2025, p. 3) introduce the idea of human life forms to “[highlight] the role of moral judgment which cannot be circumvented as a basis for developing alternative models for the ethical guidance of AI in healthcare.” However, Heuser, Steil, and Salloch overlook that life forms are not uniform and can conflict. This paper aims to explore the diversity of forms of life. In particular, it will raise attention to the possibility of oppressive forms of life and argue that this possibility generates challenges for the strategy of social critique.

Introduction

There is a growing literature that tries to lay out aspects of normativity in the growing field of AI ethics, which cannot be sufficiently captured within a system of principles (Corrêa et al., 2024; Heuser, Steil and Salloch, 2025). As Hamsphire (1978: p. 53) indicated in the seventies “moral theory cannot be rounded off and made complete and tidy ... partly because new ways of life should always be expected to arise in association with new knowledge and with new social forms.” Moreover, the uncodifiability of moral principles and artificial moral agents is getting more attention (Furlan, 2024; Graff, 2024; Villegas-Galaviz & Martin, 2024). We should take care that principlist reasoning does not end up overfocusing the harsh spotlight of (old and new) moral principles and instead pay attention to the softer and subtler lights of moral life. Heuser, Steil, and Salloch’s (2025: p. 3) recent publication is the most extensive to introduce the idea of ‘human life forms’ to highlight the role of moral judgment, which “cannot be circumvented as a basis for developing alternative models for the ethical guidance of AI in healthcare”. This publication serves as the principal target of our critique in this paper.

Heuser, Steil and Salloch define life forms as, based on the work of Jaeggi (2018: p. 50 [emphasis in original]), “forms of life are nexuses of practices, orientations, and orders of social behavior. They include attitudes and habitualized modes of conduct with a *normative character* that concern the *collective conduct of life*, although at the same time they are *not strictly codified* or institutionally binding”. Other reference points, according to Heuser, Steil, and Salloch, that regard human life forms as normatively relevant are Kant’s (1913) critique of judgment and Wittgenstein’s (1976) work on rule-following. Like actions within a game, on the account from Wittgenstein that these authors cite, participation in a form of life aims to achieve a goal by following shared rules. However, Heuser, Steil and Salloch put the term “life-world” at the centre of their title. Although we are more interested in ethical problematization than in exegetical accuracy and recognize that any attempt to learn from

another writer is always likely to involve a certain eclecticism, there is a question as to how far one can go with their use of this concept. Neither did Wittgenstein use the concept of ‘Lebenswelt’ (or ‘life-world’) in his work, nor did Jaeggi introduce it in her work. Indeed, Wittgenstein scholars have explicitly contrasted the philosopher’s notion of *Lebensform* (‘life form’) with phenomenologists’ concurrent usage of *Lebenswelt* in order to emphasize crucial aspects of his work (Floyd, 2020).

Moreover, it is peculiar that other methodological commitments in Jaeggi’s work are not touched upon, such as the tradition of critical theory and the possibility of immanent critique, especially if we take into account their worry about the diversity of life forms and that “racism and discrimination ... should not be left only to technological solutions, but also need to be addressed politically” (Heuser, Steil, and Salloch, 2025: p. 7). As Stahl (2024: p. 78) indicates, “many forms of life are deeply shaped by white supremacy, sexism, the exclusion of non-traditional gender identities, and class differences and will thus count, to a greater or lesser degree, as oppressive in this sense.” While Heuser, Steil, and Salloch do not spend time themselves discussing the relationship between forms of life and oppression in the context of AI ethics, this paper assumes that questions raised in feminist ethics and a Wittgensteinian approach to feminist epistemology provide extra tools to critically examine such relationships in AI ethics.

The remainder of this paper proceeds as follows: First, the importance of crises in life forms is sketched out. Second, an argument is offered in favor of identifying which groups can participate in identifying crises in life forms. Finally, the paper concludes with an emphasis on the importance of further investigating storytelling in AI ethics.

Crises and forms of life

Heuser, Steil, and Salloch draw on classic debates in moral and political philosophy concerning the role of ethical principles and the relationship between justice and the good life, but this means that standard objections to mainstream liberal ideal theorizing from critical theorists are left out. For example, critical theorists such as Jaeggi often argue that “many liberal theories presume an unrealistically idealistic social ontology that views social institutions as cooperative arrangements emerging from a preceding consensus” (Stahl, 2024: p. 82), yet this assumption remains unexamined. However, this assumption remains unexamined. The AI ethicists’ labors cannot stop by just introducing a new concept in the debate but must engage with the conceptual and political baggage that it brings to justify its role in ethical inquiry. Moreover, Heuser, Steil, and Salloch ignore that Jaeggi’s model of social practice takes crises as reference points to diagnose the ethical failures of social arrangements: “The moment of crisis forces reflection on and adjustments of practices—a re-creation of practices—that were previously taken for granted” (Jaeggi, 2017: p. 167). To get the full picture of life forms in AI ethics, such moments of crisis must be explored further.

Jaeggi was not the first to emphasize the importance of unstable life forms in moral philosophy, and a similar idea has been developed by several feminist moral philosophers in a Wittgensteinian spirit, including Walker (2007) and Alcoff (1991). Questions of moral and epistemic authority matter particularly in this critical tradition because of the historic denial of them to marginalized groups, such as women and/or people of color. It seems crucial, therefore, to ensure that analyses of life forms in the context of humans and AI systems do not overlook the multiplicity, complexity, and multi-dimensionality of possible positions within forms of life. This is especially urgent if we take into account the existing oppressive forms of life (Stahl, 2024), i.e. social processes through which the self-understanding inherent in the life form is distorted by relations of domination and exclusion contributing to marginalization. Indeed, the extent to which AI narratives inform and engage with issues of gender and ‘whiteness’ (Katz, 2020) are of ongoing concern and debate in wider attempts to decolonize AI ethics (Mhlambi & Tiribelli, 2024). Ideologies and related experiences of domination get in the way of perceiving many of the moral and political problems we face in AI ethics. Feminists, as Khader (2024: p. 637) suggests, “want to peel this layer of distortion away” or at least illuminate the messy realities, with the aim of understanding the often-conflicting needs that find expression in words and concepts.

Whose crisis?

Heuser, Steil, and Salloch often refer to what “we” do and to “our” practices in AI development, but that appearance of social homogeneity is misleading. As Stahl (2024: p. 87) writes, “[t]he question: “Whose problems? Whose crisis?” is therefore one that the analysis of life forms cannot avoid.” Thus, Fricker (1999: p. 207) argues: “With the possibility of such a question, we confront the disunity of the ‘we’ who agree in form of life. If the different ‘we’s within a form of life stand to one another in relations of advantage and disadvantage, power and powerlessness, then this inequality is likely to be reiterated in interpretive practice.” Fricker’s conviction that attention to who is knowing, from where, and with what stakes, is crucial to supporting, rather than undermining, a usable conception of moral knowledge, is shared by other feminist philosophers (Walker, 2007; Scheman, 2022).

Two possible mechanisms can be identified that affect the way in which groups and their ethos of “we” can participate in the hermeneutical processes through which a life form either identifies or does not identify problems as constituting crises. The first kind of mechanism is defined by Stahl (2024) as “a form of exclusion from the administration of the hermeneutical resources of a life form”. In other words, the unfair exclusion of knowers from participation in the social processes by which a collectively binding self-understanding is determined. For instance, Miragoli (2024) describes the case of Irina, who is experiencing feelings for girls and, in an attempt to understand her emotions better, goes on Google. The conformist

behavior displayed by Google's algorithms, which tend to read and interpret input information in the light of categories extracted from dominant heteronormative trends, is not helpful to Irina in making sense of her own sexual experience.

Part of the challenge of such algorithmic conformism is that forms of life change profoundly over time, whereas AI-driven decision-making systems are prone to what Hardalupas (2024) calls "epistemic calcification," whereby (potentially conformist) interpretive frameworks become fixed. Hardalupas argues that AI systems, due to human input, biased data, and wider institutional inertia, embed certain assumptions about health and disease, for example, to the exclusion of other perspectives, leading to a kind of epistemic oppression called "contributory injustice" (Dotson, 2012: p. 25). If technologies are capable of embedding – or even, as Winner (1986) suggests, constituting – forms of life, then medical AI systems' classifications may be not only unhelpful but intractably so, with critiques made by patients failing to impact their functioning. At their most radical, AI systems may take part in what Winner refers to as "technological somnambulism," whereby "we [whoever that 'we' may be] so willingly sleepwalk through the process of reconstituting the conditions of human existence" (1983: p. 10). Taking this metaphor further, awaking AI developers to the perspectives of the marginalized demands neither the subtle brush of a principle-based approach nor some faint awareness of a voice from the outside but a full-body wake-up in the form of fundamental critique. While "the recognition of the impact of life-worlds and joint procedures of interpretation" that Heuser et al. (2025: p. 12) propose is necessary, it is only a first step in the direction of such critique.

"It is remarkable that even under significant hermeneutical constraints, the oppressed may still achieve what Stahl (2024: p. 86) calls "the development of an individual or subcultural consciousness of crisis." As Jaggar (1983: p. 370) highlights, the suffering of oppressed groups can be epistemically beneficial: "Their pain provides them with a motivation for finding out what is wrong, for criticizing accepted interpretations of reality and for developing new and less distorted ways of understanding the world." This has also been highlighted in more recent epistemic injustice literature, which indicates how marginalized communities often possess fully fledged concepts to make sense of oppressive experiences as well as fully fledged expressions that communicate them but lack acknowledgment on the part of the dominant culture (Medina, 2013). Consequently, outsiders who are offered space within dominant forms of life "are in positions to be listened to and to be intelligible" (Scheman, 1996: p. 403).

A second kind of mechanism that can lead to the emergence of oppressive forms of life is best thought of as a form of "*hermeneutical disadvantage*" (Stahl, 2024: p. 86) or alternatively "epistemic death" (Medina, 2017: p. 49). This occurs when subjects are not given even minimal amounts of credibility, are not treated as intelligible communicators, and are

prevented from participating in meaning-making and meaning-sharing practices. Under such conditions of epistemic death, subjects owe nothing to those communities, and they may have a right “to fight epistemically by any means necessary (including the right to lie, to hide, to sabotage, to silence others, etc.), demonstrating loyalty and solidarity only with alternative epistemic communities (communities of resistance)” (Medina, 2017: p. 49). We can identify similar practices related to AI that interrupt and disrupt the established epistemic economy of a society and its practices and institutions (Bonini & Treré, 2024; Lehuedé, 2024).

As well as the ways in which forms of life disadvantage their own members, oppression can be identified in one’s being excluded, or expelled, from the forms of life to which others are granted access. Yet this does not mean that there are no legitimate boundaries to be drawn around distinct forms of lives, which are warranted in part due to the distinct vantage points on life and morality that they afford us. On Lindemann’s (2014: p. 19) Wittgensteinian account, personhood – interpreted as a form of moral consideration – is restricted to those entities that “share in our form of life” as human beings and can “be brought into personhood” by those around them, which is limited to humankind. This excludes not only non-human animals but also highly humanlike AI systems from participation in our forms of life, against the assumptions of some philosophers of robotics (Schmid, 2017). With this exclusion comes the implication that AI systems are unable to appreciate the norms internal to our shared forms of life and that, while ethical principles may be programmed into an AI system, awareness of the inner workings of human forms of life is out of their grasp. For a human being to “be able to participate skillfully in a form of life” (de Boer, 2023: p. 2272) means appreciating its distinctive normative demands, which can only be done from within those human forms of life that artificial agents (whether hypothetical or actual) are outside. However, the moral stakes in deciding who takes part in ‘our’ life forms remain high, and assumptions about the individuals and entities that they include and exclude cannot be taken for granted.

Further mapping the rough ground

Heuser, Steil, and Salloch (2025, p. 11) argue that “emphasis must be placed on stakeholders’ good life in its various forms and practices” but do not recognize that stakeholders are differentiated epistemically according to their social location: what might be reasonable to believe from one location may very well be unreasonable from another. One response to the challenge of better understanding individuals in context is to suggest that we situate them through stories: by telling stories we reveal different life forms (Schachtner 2020). As Walker (2007: p. 72) indicates, “[b]ecause people and their relationships are not uniform and situations are not necessarily repeatable, moral consideration on this view presses toward enrichment of detail and amplification of context.” To get the full picture of life forms in AI ethics, the critical value of AI narratives and responsible storytelling must at the very least be

explored further (Coeckelbergh, 2023; Chubb, Reed, and Cowling, 2024). In other words, we need to go back to the “rough ground” of how words and narratives are actually used (Wittgenstein, 1976: §107).

The point is not just to collect new stories but to recognize their transformative power (Haraway, 2016). Through storytelling, we can uncover injustices, inspire collective action, and foster new types of solidarity (Jones & Fulfer, 2024). For example, Lindemann (2001) explains that storytelling can convert 'master narratives'—typical stories representing dominant moral and understandings and related cultural views—into ‘resistance narratives’ that highlight and raise awareness of the harms that individuals suffer due to these master narratives. Inspired by Lugones (2003), several scholars have indicated how certain playful ways of using language can help marginalized individuals safely navigate dominant worldviews without being “trapped in oppressive definitions and perceptions” (Boncompagni, 2023: p. 5). Alternative storytelling practices in AI might include how Indigenous-led researchers in New Zealand are using a machine learning-based speech-to-text system to help revive Te Reo Maori, the language of the Maori people (Papa Reo, 2024). Another example is discussed in the Radical AI Podcast of June 2021. An Indian feminist app called Mumkin is “designed to have difficult conversations to address the issue of female khatna (also known as female genital cutting) within the Bohra community in India. The chatbot is intended to help users shed inhibitions and gain confidence to have real-life conversations on gender, culture, and society” (Toupin, 2024, p. 589). Finally, the recent ethnographic research of Madianou (2023, p. 1) reveals ordinary practices of resistance to “datafication and technological experimentation which are becoming increasingly ubiquitous in the aid sector,” and we could imagine that something similar is taking place in healthcare practices.

This highlights the importance of (self-)critique, along with the concepts of patience, slowness, and listening (Chadwick, 2024), in recognizing forms of struggle and resistance. It also demonstrates that storytelling can enhance a (marginalized) individual's ability to act autonomously by creating opportunities for others to acknowledge and respect the stories that shape her world. Scheman (2009) insightfully describes this form of agency as “perceptual autonomy,” which necessitates that persons can express a “nonperisopic point of view” that hasn't been altered to align with a dominant perspective. The contestation of AI power begins with experiments in listening differently, for which bringing together work on forms of life with feminist philosophy is a valuable first step.

Conclusion

In conclusion, while many points that Heuser, Steil, and Salloch raise in their paper on AI ethics and their focus on the importance of studying forms of life are well taken, it is much less clear how their analysis of forms of life leaves room for reflection on oppressive positions and forms of resistance. Moreover, we suggested that messy storytelling should be

more central in AI ethics. As Eagleton and Jarman (1993: p. 55) state in the script of the *Wittgenstein* film: “roughness and ambiguity and indeterminacy aren’t imperfections—they’re what make things work.” From this point, the important question about AI driven systems becomes, to paraphrase Winner (1986) with a nod to Walker, what kind of *world* are ‘we’, as a group that is “not uniform”, making?

References

- Alcoff, Linda (1991) "The problem of speaking for others", *Cultural critique* 20, 5-32.
- Boncompagni, Anna (2023) "Forms of life and linguistic change: the case of trans communities", *Philosophies* 8, 50.
- Bonini, Tiziano; Treré, Emiliano (2024) *Algorithms of resistance: The everyday fight against platform power*, Cambridge: Mit Press.
- Chadwick, Rachelle (2024) "The question of feminist critique", *Feminist Theory* 25, 376-395.
- Chubb, Jennifer; Reed, Darren; Cowling, Peter (2024) "Expert views about missing AI narratives: is there an AI story crisis?", *AI & society* 39, 1107-1126.
- Coeckelbergh, Mark (2023) "Narrative responsibility and artificial intelligence: How AI challenges human responsibility and sense-making", *AI & society* 38, 2437-2450.
- Corrêa, Nicholas Kluge; Santos, James William; Galvão, Camila; Pasetti, Marcelo; Schiavon, Dieine; Naqvi, Faizah; Hossain, Robayet & De Oliveira, Nythamar (2024) "Crossing the principle–practice gap in AI ethics with ethical problem-solving", *AI and Ethics*, 1-18.
- de Boer, Bas (2023) "Explaining multistability: Postphenomenology and affordances of technologies", *AI & Society* 38, 2267-2277.
- Dotson, Kristie (2012) "A cautionary tale: On limiting epistemic oppression", *Frontiers: A Journal of Women Studies* 33, 24-47.
- Eagleton, Terrey; Jarman, Derek (1993) *Wittgenstein : The Terry Eagleton script. The Derek Jarman film*, London: British film institute.
- Floyd, Juliet (2020) "Wittgenstein on ethics: Working through Lebensformen", *Philosophy & Social Criticism* 46, 115-130.
- Fricker, Miranda (1999) "Epistemic oppression and epistemic privilege", *Canadian Journal of Philosophy Supplementary Volume* 25, 191-210.
- Furlan, Timothy J (2024) "Principlism, uncodifiability, and the problem of specification", *Cambridge Quarterly of Healthcare Ethics*, 1-22.
- Graff, Joris (2024) "Moral sensitivity and the limits of artificial moral agents", *Ethics and Information Technology* 26, 13.
- Hampshire, Stuart (Ed.). (1978). *Public and private morality*, Cambridge: Cambridge University Press.

- Haraway, Donna J (2016) *Staying with the Trouble: Making kin in the Chthulucene*, Durham: Duke University Press.
- Hardalupas, Mahi (2024) "Contributory injustice, epistemic calcification and the use of AI systems in healthcare", *Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society* 7, 573-583.
- Heuser, Stefan; Steil, Jochen; Salloch, Sabine. (2025) "AI Ethics beyond Principles: Strengthening the Life-world Perspective", *Science and Engineering Ethics* 31, 7.
- Jaeggi, Rahel (2018) *Critique of forms of life*. Cambridge: Belknap Press.
- Jaggar, Alison (1983) *Feminist politics and human nature*. Totowa: Rowman and Allanheld.
- Jones, Janet; Fulfer, Katy (2024) "Solidaristic Listening", *Hypatia* 39, 634-651.
- Kant, Immanuel (1913) *Kritik der Urteilkraft*. Berlin: Reimer.
- Katz, Yarden (2020) *Artificial whiteness: Politics and ideology in artificial intelligence*. New York: Columbia University Press.
- Khader, Serene J (2024) "How is feminist philosophy nonideal theory?", *Social Theory and Practice* 50, 619–642.
- Lehuedé, Sebastián (2024) "An elemental ethics for artificial intelligence: water as resistance within AI's value chain", *AI & Society*, 1-14.
- Lindemann, Hilde (2001) *Damaged identities, narrative repair*, New York: Cornell University Press.
- Lindemann, Hilde (2014) *Holding and letting go: The social practice of personal identities*, Oxford: Oxford University Press.
- Lugones, Maria (2003) *Pilgrimages/peregrinajes: Theorizing coalition against multiple oppressions*, Lanham: Rowman & Littlefield Publishers.
- Madianou, Mirca (2023) "Contesting Technocolonialism: The 'Ordinary Transcripts' of Everyday Resistance", Paper presented at AoIR 2022: The 23rd Annual Conference of the Association of Internet Researchers. Dublin, Ireland: AoIR. <https://doi.org/10.5210/spir.v2022i0.13048>
- Medina, José (2013) *The epistemology of resistance: Gender and racial oppression, epistemic injustice, and the social imagination*. Oxford: Oxford University Press.
- Medina, José (2017) "Varieties of hermeneutical injustice", In Ian James Kidd, José Medina, Gaile Pohlhaus, Jr. (eds.) *The Routledge handbook of epistemic injustice*, New York: Routledge, 41-52.
- Mhlambi, Sábëlo; Tiribelli, Simona (2023) "Decolonizing AI ethics: Relational autonomy as a means to counter AI harms", *Topoi* 42, 867-880.
- Miragoli, Martin (2024) "Conformism, Ignorance & Injustice: AI as a Tool of Epistemic Oppression", *Episteme*, 1-19.
- Papa Reo (2024) "Papa Reo", <https://papareo.nz>

Seeing the World Aright: Wittgenstein, Murdoch and the Ethical Engagement with Literature

Marcello Di Massa (Pisa, Italy)

Abstract

The relationship between literature and morality, as experienced by Wittgenstein during his life, brings to the forefront the ethical dimension of engaging with reading. Wittgenstein's appreciation for literature stemmed not only from its ability to illuminate significant themes of life and morality, but also from his view of reading as a deeply spiritual activity that was integral to his ethical and existential commitment. For him, reading was not a passive activity but an ethical engagement that could reveal transformative insights. I argue that it is both useful and coherent to consider the perspective offered by Iris Murdoch, which can be seen as expanding Wittgenstein's point of view. As Murdoch claims, both philosophy and literature seek an ethical sense by fostering moral and conceptual clarity. In particular, great literature offers a disciplined vision that challenges inhibitions and preconceived notions, involving readers in a morally active and imaginative effort. Through attentive engagement, literature educates us in understanding human situations and the intricate moral and conceptual texture of life. Scholars such as Cora Diamond and Toril Moi further develop these ideas, emphasizing the moral significance of focusing on the particular in literature and life. Such attention reflects an ethical attitude that transcends individual reflection, extending to social life and diverse moral landscapes. This approach underscores the continuity between life, philosophy, and literature, encouraging an open, responsive, and ethically engaged reading practice, in the authentic Wittgensteinian spirit.

1. Wittgenstein and the Ethical Value of Literature

It is well known that Ludwig Wittgenstein was driven by a profound ethical impulse, which led him to engage in constant self-questioning and self-examination throughout his life. His ideal was to achieve a true form of moral decency. Accordingly, he set high demands on himself, ruthlessly criticizing his own shortcomings and weaknesses. This extreme seriousness was typical of Wittgenstein, along with his commitment to authenticity and self-honesty. In general, one could say that he was a man grappling with life, always animated by a search for meaning, and he himself lamented his inability to "take things lightly" (GT 1991: 43; my transl.). This ethical dimension is not confined to the purely personal sphere but deeply shapes his conception of philosophy. Wittgenstein viewed philosophy not as a set of theories or doctrines but as a practical and transformative activity (as particularly emphasized by so-called resolute or therapeutic readers), requiring active self-engagement. As he famously wrote: "Work on philosophy—like work in architecture in many respects—is really more work on oneself. On one's own conception. On how

one sees things. (And what one expects of them.)” (VB 1998: 24). For Wittgenstein, the philosophical endeavor was thus inseparable from a moral task: clarifying language and its relation to the world also meant understanding oneself and one’s relationships with others, in a journey that demands courage, rigor, and a deep respect for the complexity of life. In particular, his approach to philosophical problems was guided by the intention to dissolve them, showing how thought is often ensnared by conceptual tangles that generate suffering, confusion, and disquiet (PI 2009: § 111; § 123). The form of therapy Wittgenstein sought to promote, especially in his later work, aims to free us from the pictures that hold us captive (PI 2009: § 115) and prevent us from seeing the world clearly, constraining us to consider only one way things must be. In a sense, perspicuity is an ethical achievement.

This union of philosophy and ethics finds a significant resonance in Wittgenstein’s relationship with literature, which he held in high regard. “You read with the soul,” his sister Hermine wrote in a 1916 letter to her brother (quoted in McGuinness 2013: 227). For Wittgenstein, reading was a serious and spiritual matter. The writers he most admired, such as Leo Tolstoy, Fyodor Dostoevsky, and Søren Kierkegaard, shared a common concern for a striking moral vision and a deep sensitivity to ethical and existential matters, as well as a religious perspective that greatly interested Wittgenstein. Interestingly, even the authors whose humor and linguistic playfulness he particularly appreciated, such as Wilhelm Busch, Karl Kraus, and Gottfried Keller, consistently revealed, in various forms, a profound integration of moral depth and aesthetic value. One might think, for instance, of Busch’s *Edward’s Dream*, in which the reader witnesses a certain moral vision unfolding through Edward’s dreamlike experience.

For Wittgenstein, reading and reflecting on literature were part of the same commitment to clarity and moral integrity that he sought to pursue as both a man and a philosopher. As early as the *Tractatus*, according to some traditional readings, the idea implicitly emerges that if someone is interested in “what is higher” (TLP 1972: 6.432) and what holds the greatest value for life, they should turn to art and literature. However, regardless of the ‘mystical’ formulation of this insight, it is important to note that this must not imply confining ethics within rigid boundaries or reducing it to a narrowly and negatively defined inexpressible domain. Rather, far from being restricted to a

specific subject matter, ethics is pervasive in the world, language, and life itself (TLP 1972: 6.421), as it is fundamentally a matter of experience, both in our use of language and in the way we navigate reality. In this regard, the ethical and aesthetic value of literature is not opposed to that of philosophy (as Wittgenstein conceived it) but rather to the discourse of science, which concerns itself with descriptions of the contingent and empirical configuration of the world. The latter is easily expressed through truth-functional propositions.

In an important sense, the goal of “seeing the world aright” (TLP 1972: 6.54), which is the declared ambition of the *Tractatus*, a work deliberately conceived as both philosophical and literary, and having an ethical point (Monk 1990: 178), is a feature that philosophy shares with literature. Literature often offers a new way of seeing things, revealing unnoticed aspects of our experience and unfolding a more or less enlightening or original vision of the world, which frequently expresses an ethical sense. Wittgenstein was, in fact, interested in the idea that one could learn from literature. In a well-known remark, he observed: “People nowadays think, scientists are there to instruct them, poets, musicians, etc., to entertain them. *That the latter have something to teach them; that never occurs to them*” (VB 1998: 42). If Wittgenstein had in mind an ‘ethical teaching’ here, it is important to note that this does not imply advocating a form of didacticism. The cognitive benefits of reading cannot be neatly captured in terms of propositional knowledge. At the same time, this does not necessarily promote an ineffabilist picture of what literature can convey. Rather, the point is that the ethical value of engaging with literary works is best grasped through the distinctive indirect style with which narrative unfolds a sense or a way of seeing things. A famous example is Wittgenstein’s varied reactions to Tolstoy’s works:

“I once tried to read *Resurrection* but couldn’t. You see when Tolstoy just tells a story he impresses me infinitely more than when he addresses the reader. When he turns his back to the reader then he seems most impressive. [...] It seems to me his philosophy is most true when it is latent in the story”. (Malcolm 1984: 98)

Concerning such artistic matters, it is about grasping something that emerges and stands out against the background of the narrative, in a manner similar to

perceiving the dawning of an aspect in a picture or the physiognomy of a face (PI 2009: § 285). Related ideas on the non-propositional mode in which language can affect or move us are explored in various ways, for instance, in the *Lectures on Aesthetics* as well as in the *Philosophical Investigations*, particularly in relation to the intuition that words can strike us by having a certain gaze or a familiar physiognomy (PI 2009: 230). Extending this theme to literature, the analogy with seeing and vision proves to be crucial. When engaging with a literary work, the reader must be able to discern the ethically relevant point, which is not directly stated and cannot be absorbed passively. And, even though it is not an explicit claim, this ability to see a point depends, in turn, on the ethical attitude of the reader, who is personally called upon to engage practically with the work. Wittgenstein, indeed, when discussing his readings with friends or family members, paid close attention to what others managed to see in them. For example, it is reported that he was interested in knowing what his friend John King could see in Dostoevsky's *The Brothers Karamazov*: whether he was impressed by Father Zosima, the legend of the Grand Inquisitor, and so on (Rhees 1984: 72). It may be reasonable to think that this ability to 'see' was indicative of the reader's personality and their ethical capacity to allow a work to speak to them and move them in an existential sense. At the very least, this is an aspect of literature that Wittgenstein deeply respected and one that reflects the intimate continuity between life, philosophy, and reading in relation to moral engagement.

2. Iris Murdoch: Moral Vision and Clarity through Literature

One may turn to ideas drawn from the work of the philosopher and novelist Iris Murdoch to find a way of understanding the moral significance of literature and its intimate connection with philosophy, which resonates with Wittgenstein's case. Murdoch curiously described herself as "a Wittgensteinian Neo-Platonist" (Dooley 2003: 92), and there are certainly convergences in their shared appreciation of everyday language and their suspicion of the generalizations inherent in the abstract methods of traditional philosophy, which obscure differences. Her insights, when taken in the ordinary sense, setting aside their ontological or substantive commitments, can be helpful here in giving substance to Wittgenstein's approach to the moral dimension of literature.

First, Murdoch is convinced that both philosophy and literature, despite their differences, are committed to the pursuit of truth: “they are both truth-seeking and truth-revealing activities” (Murdoch 1999: 11). Although this formulation sounds quite distant from Wittgenstein, one must understand the term ‘truth’ in a very broad sense; it would actually be better to replace it with ‘meaning’ or ‘sense’. For Murdoch, great art is capable of showing us the world truthfully and clearly, and the clarity gained in this way is an ethical achievement because it requires exercising moral vision by disciplining our inhibitory tendencies and systems (or pictures, in Wittgenstein’s sense) that prevent us from seeing things differently (Murdoch 1999: 352). Given her belief that we are perpetually moral beings and thus always morally active, the encounter with a literary work also takes place on moral ground, stimulating imaginative engagement and the achievement of a clear moral vision, potentially involving personal transformation. This should not be understood in overly abstract terms. What we find in literature are the connections and conceptual relationships that shape various forms of life, so that engaging with literature “is an education in how to picture and understand human situations” (Murdoch 1999: 326). Being responsive to these different conceptual articulations, revealed in the use of language in specific and concrete contexts explored by literature, means being morally responsive to the expressive variety of human life. It is important to emphasize the aspect of plurality because the idea is precisely that literature provides an exploration of the interrelations among concepts, in all their variety, shaping different moral forms of life. For example, one could consider Wittgenstein’s fascination with Dostoevsky’s novels, where complex relationships among concepts such as sin, guilt, redemption, happiness, and love are explored, even when not explicitly mentioned. Recognizing and appreciating this interplay activates a moral attention that is “unsentimental, detached, unselfish” (Murdoch 1999: 64), allowing us to unfold a vision of reality that is clear, lucid, and ethically charged. In a specific sense, as Murdoch says, “the greatest art is ‘impersonal’ because it shows us the world, our world and not another one, with a clarity which startles and delights us” (Murdoch 1999: 352). In this respect, one may think of the *sub specie aeterni* perspective (TLP 1972: 6.45) that Wittgenstein envisioned for the artistic viewpoint, insofar as it distances us from our more practical and self-interested concerns, while still engaging the reader’s ethical disposition.

The reader and their moral life are directly involved in their engagement with the narrative material and characters, with such force and intensity that it opens the possibility for a change in their moral vision and personality, which is always susceptible to ethical growth and transformation. Furthermore, this ties into the idea of the moral value of the “apprehension of the unique” (Murdoch 1999: 46) or the particular. Wittgenstein was precisely interested in the concrete and particular nature of the situations and scenarios explored in literature, as opposed to the abstraction that traditional philosophy often falls victim to. In this sense, far from there being a divide, there is a profound continuity between our attitude toward the particular in ordinary life and the attitude we adopt toward what we encounter in works of literary fiction. Paying attention to all the particular aspects of a person that shape the relationship between their use of words and their personal view of things amounts to grasping what Murdoch calls “the texture of a man’s being” (Murdoch 1999: 39), which involves active engagement with our moral life. This form of attention may find resonance with Wittgenstein’s idea of “an attitude towards a soul” (PI 2009: 187). It is an attitude that perceives as immediately expressive the various particular traits constituting a person’s conceptual and linguistic life, such that we do not feel the need to infer that we are dealing with another person, as though they were an object to decipher. In relation to literature, this form of attention seems to be directly activated, enabling us to be responsive to all the particular details that animate the moral lives of the characters and the work itself (in the sense of the vision it conveys as a whole) with which we engage, often moving us profoundly. The clarity that can result is an achievement of the reader, calling into play their personal moral vision. In Murdoch’s words: “clear vision is a result of moral imagination and moral effort” (Murdoch 1999: 329).

Given these points of contact, it is no surprise to see that, for example, Cora Diamond, while developing a framework inspired by Wittgenstein on the relationship between moral philosophy and literature and commenting on ideas drawn from Martha Nussbaum’s ethical approach, refers explicitly to Murdoch. Diamond praises the general approach that moral philosophy should not be confined to discussions of action and choice, as well as the idea that attention to the particular holds great moral value, a principle observable both in ordinary life and in literary fiction. More specifically, Diamond values a

phrasing we have already encountered above when she writes: “‘Texture of being’ is a useful expression for this area; it also allows for an extension of Miss Murdoch’s general point. Moral reflection may be directed not just towards individual human beings but towards forms of social life” (Diamond 2004: 140). In this way, moral significance is not restricted to the individual sphere or interiority but permeates our reality in a global sense, consistent with the idea that literature often confronts us with a variety of moral forms of life, calling for our moral attention.

Similarly, Toril Moi, in her relatively recent work *Revolution of the Ordinary: Literary Studies after Wittgenstein, Austin, and Cavell*, also draws on ideas from Murdoch, combining them with Simone Weil and Diamond herself, to develop an approach to literature that shows how it can enable us to see reality with greater clarity (Moi 2017: 223). The notion she uses as a reference is precisely that of ‘attention’, adopted programmatically following Murdoch’s definition: it “express[es] the idea of a just and loving gaze directed upon an individual reality” (Murdoch 1999: 327). This type of gaze is what allows us to realize that there is much more to see beyond our personal perspective, and literature does just that: it opens our eyes to the world, inviting our attention to focus, without restrictions, on the particular. Fundamentally, Murdoch supports the idea of “moral attitudes which emphasise the inexhaustible detail of the world, the endlessness of the task of understanding” (Murdoch 1999: 46), which seems to find a direct application in our experience of engaging with literary works. From this, Moi gives substance to this idea with various literary examples, testing her proposal to understand reading as a practice of acknowledgment (in Cavell’s sense) and as an adventure in discovering the new (de Beauvoir, Nussbaum, Diamond). She combines the idea of attention as a “just and loving gaze”, seen as a form of realism in an ordinary sense, with the ‘realistic spirit’ Diamond discusses in relation to Wittgenstein, which is distinct from philosophical or metaphysical realism (Diamond 1995). Interestingly, the result is always one of seeing the world clearly, which has nothing to do with objectivity in the scientific sense but is instead a matter of a ‘moral response’ that involves our entire personality and ethical attitude. The analogy with the ambition of the *Tractatus* is strikingly evident.

3. Conclusions

How should we understand these observations on the relationship between literature and the moral dimension? Some clarifications are necessary. Caution must be exercised against the temptation to derive a general approach to literature. That is, we should avoid rushing to make universally valid claims: as obvious as it may sound, literature is infinitely varied and functions in many different ways. The case of Wittgenstein shows that he, after all, was particularly interested in the philosophical value of his readings, in the sense that he found in them profound reflections and representations of themes of great importance to life in general. The ethical sense of his relationship with literature was undoubtedly paramount and was experienced in continuity with the ethical commitment of his work in philosophy and his efforts to lead an ethical life. However, this is only one aspect of the relationship we can have with literature, even if Wittgenstein considered it probably the most important and significant. With this in mind, if one wishes to systematically explore the moral value that literature can convey, following the path indicated by Murdoch in a way broadly consistent with Wittgenstein's insights, it is essential to remember that this will be one aspect of our grammar of the concept of 'literature', so to speak, and not a universal definition or account of it. From this perspective, the lesson from Wittgenstein and Murdoch is of great importance: it shows us how to consider the deep continuity between life, philosophy, and literature, and to approach the latter without expecting to receive a neatly packaged moral message. Instead, it encourages us to actively engage our attention and moral imagination, ready to undergo a transformation in our vision and personality.

References

- Diamond, Cora (1995) *The Realistic Spirit: Wittgenstein, Philosophy, and the Mind*, Cambridge, MA: MIT Press.
- Diamond, Cora (2004) "Having a Rough Story about what Moral Philosophy is", in: John Gibson and Wolfgang Humer (eds.) *The Literary Wittgenstein*, London: Routledge, 133-145.
- Dooley, Gillian (ed.) (2003) *From a Tiny Corner in the House of Fiction: Conversations with Iris Murdoch*, Columbia: University of South Carolina Press.
- Malcolm, Norman (1984) *Ludwig Wittgenstein: A Memoir*, Oxford: Oxford University Press.

McGuinness, Brian (2013) "Wittgenstein and Dostoevsky", in: Sascha Bru, Wolfgang Huemer and Daniel Steuer (eds.) *Wittgenstein Reading*, Berlin: De Gruyter, 227-242.

Moi, Toril (2017) *Revolution of the Ordinary: Literary Studies after Wittgenstein, Austin, and Cavell*, Chicago: Chicago University Press.

Monk, Ray (1990) *Ludwig Wittgenstein: The Duty of a Genius*, London: Penguin.

Murdoch, Iris (1999) *Existentialists and Mystics: Writings on Philosophy and Literature*, Edited by Peter Conradi, New York: Penguin.

Rhees, Rush (ed.) (1984) *Recollections of Wittgenstein*, Oxford: Oxford University Press.

Wittgenstein, Ludwig (1972): *Tractatus Logico-Philosophicus*. Übersetzt von D. F. Pears und B. F. McGuinness. Reprinted with corrections in the light of the publication of Wittgenstein's correspondence with C. K. Ogden about the first translation. London: Routledge and Kegan Paul.

Wittgenstein, Ludwig (1991): *Geheime Tagebücher*. Hg. von Wilhelm Baum. Zweite Auflage. Wien: Turia und Kant.

Wittgenstein, Ludwig (1998): *Ludwig Wittgenstein Vermischte Bemerkungen. Eine Auswahl aus dem Nachlaß / Culture and Value. A Selection from the Posthumous Remains*. Hg. von Georg Henrik von Wright unter Mitarbeit von Heikki Nyman. Neubearbeitung des Textes durch Alois Pichler. Übersetzt von Peter Winch. Revised second edition. Blackwell: Oxford.

Wittgenstein, Ludwig (2009): *Philosophical Investigations / Philosophische Untersuchungen*. Hg. von P. M. S. Hacker und Joachim Schulte, übersetzt von G. E. M. Anscombe, P. M. S. Hacker und Joachim Schulte. New York: Wiley.

Realism, Wittgenstein, and the Formative Years of the Croatian Analytic Tradition

Dušan Dožudić (Zagreb, Croatia)

Abstract

This paper discusses the character and development of philosophy in Croatia over the past hundred and fifty years. The emphasis is on the beginning and early development of the Croatian analytic philosophy tradition as well as on its domestic predecessors (from the late 19th century onwards). In the proposed reconstruction of the development of that tradition (its proponents, themes of interest, and institutional support), the emphasis is on women philosophers' role in its early phase. In particular, the emphasis is on their role during the 1960s and 1970s and the philosophical background that helped shape their interests. The conclusion is that the general realist position, a particular philosophy of science, and an interest in Wittgenstein and Wittgenstein-related themes played an important role in the formative period of the Croatian analytic philosophy tradition. Given its motivation and direction, this tradition opposed the then-dominant philosophical orientation in Croatia, namely, Marxism combined with German continental thought.

1. Introduction

Exploring the work of women involved in the paths of early analytic philosophy primarily leads one to philosophers of the second half of the 19th century and the first decades of the 20th – the formative years of the analytic tradition. It leads one to women philosophers such as Victoria Welby (1837–1912), E. E. Constance Jones (1848–1922), and Susan Stebbing (1885–1943). Women philosophers who began working at the end of that period but continued their engagement afterwards also come to mind: Susanne Langer (1895–1985), Alice Ambrose (1906–2001), G. Elizabeth M. Anscombe (1919–2001), or Ruth Barcan Marcus (1921–2012). All these are just a few recently noted in the literature (e.g. van der Schaar and Schliesser 2017; Verhaegh and Peijnenburg 2022). Besides the mentioned British and American analytically minded women philosophers, there are those from that period coming from the continental part of Europe that should also be added to the tradition. Most notably, those involved in the analytically oriented groups at the continent (the Vienna Circle and the Lvov-Warsaw School), with representatives such as Olga Hahn-Neurath (1882–1937), Rozalia Rand (1903–1980), Maria Kokoszyńska-Lutmanowa (1905–1981), and Maria Ossowska (1896–1974) (cf. Brożek 2017; 2022; Verhaegh and Peijnenburg 2022). Unsurprisingly, women philosophers of the time adopted the analytic mindset primarily in places where analytic

philosophy had strong representation, such as Cambridge, Oxford, Vienna, Warsaw, and some American universities.

The outlined framework allows one to rethink, supplement, and reconstruct the history of analytic philosophy, particularly its early (or earlier) period, roughly until the middle of the 20th century, acknowledging the place and contributions of the above-mentioned (as well as other) women philosophers in it. Subsequent developments in these places and among associated groups and their representatives are less plausibly categorised as *early* analytic philosophy. In that later period, with the already established framework and support, more and more women became engaged in the field and adopted the analytic mindset. But, in a sense, such a reconstruction, with its unrestricted understanding of the phrase “early analytic philosophy”, might leave out some elements important for telling a more complete story. And that is particularly true if one focuses on the continental part of Europe.

The “early” of the “early analytic philosophy” phrase is a context-sensitive term whose content changes with parameter changes, including geographic location. Therefore, what is early relative to one geographic location may not be early relative to another, and if placed at a single timeline, the distance between the two agents or events might be significant. The phrase “analytic philosophy” itself may thus also signal different things in different contexts. The phrase might mean analytic philosophy *simpliciter* (no matter how one may characterise it conceptually or frame it geographically and chronologically), or its meaning might be more restricted. And one restrictive parameter is a particular location that one has in mind when considering representatives of early analytic philosophy. In many places across the globe, where analytic philosophy is nowadays well-present (if not the dominant orientation), it was the latecomer at the local philosophical scene, surely a later arrival compared to the places mentioned above. Or latecomer compared to other philosophical orientations represented there, such as phenomenology, existentialism, Hegelianism, or Marxism, many of which one typically and loosely subsumes under the heading “continental”.

Building on the above-outlined conception, in this paper, I present a particular tradition of analytic philosophy that began to develop later than in many other places, such as those mentioned above. Nonetheless, it also had its early

period, and women philosophers played a significant role in it. In the following section, I will briefly reflect on some predecessors of analytic philosophy in Croatia. Then, I will explore the role of women philosophers in shaping analytic philosophy in Croatia during its formative years (1960s–1980s) when Croatia was part of socialist Yugoslavia and other philosophical trends dominated its scene.

2. A hundred years of philosophy, Zagreb 1874–1974

It would be incorrect to say that Croatia had no traces of analytic philosophy until the second half of the 20th century. It would be equally incorrect to say that there was no interest in the style and themes particularly close to the analytic orientation in Croatia even before a more general awareness of the orientation characterised and perceived as *analytic* philosophy. However, what was missing in that period was that until the end of the 1960s and the early 1970s, there were no *systematic* engagements with analytic philosophy and no *systematic* attempts to do philosophy in an analytic fashion. Before that time, no one identified themselves as an *analytic* philosopher.

The Zagreb University was re-established in 1874, and in the next ninety years or so, one can recognise several philosophical attempts that were close in style and content to the analytic tradition. For example, Franjo Marković (1845–1914), the first philosophy professor at the re-established Zagreb University, adopted a Herbartian line of doing philosophy. He was strongly anti-idealistically oriented, promoting a realist and antipsychologistic empiricist view, often based on conceptual and linguistic analyses, a clear argumentative style, and a problem- rather than history-based approach, with a high appreciation of logic, mathematics, and natural sciences. Marković was also influenced by a Bolzanian line of thought, as well as by several Neo-Kantians. Thus, unsurprisingly, he developed a style of philosophy easily related to Frege, who worked in the same period and was often stimulated by the same 19th-century philosophers, such as J. S. Mill, H. Lotze, and F. A. Trendelenburg (for details concerning Frege, see Gabriel 2013; concerning Marković, Kovač 1996 and Marotti 2020: 13–27; and for Herbart's relation to analytic philosophy, Beiser 2022).

Although Marković made a significant contribution to the development of modern Croatian philosophy in several respects (cf. Kovač 1996), his way of

conceiving and doing philosophy was considered outdated and thus never adopted by his successors in the period before World War II, such as Gjuro Arnold (1853–1941), Albert Bazala (1877–1947), and Pavao Vuk-Pavlović (1894–1976). They all formed their philosophical conceptions by relying on (primarily) German-speaking philosophers of different orientations, more distant from the analytic tradition (Lazović and Pavković 1998: 51; Mišćević 1995: 172). Thus, until the end of World War II, various philosophical traditions dominated Croatia, aligning more with phenomenology, idealism, anti-scientism, or anti-intellectualism than with analytic philosophy. Realistically minded Neo-Scholastic philosophers of the period, such as Stjepan Zimmermann (1884–1963), were closest in style and orientation to the analytic tradition but without any engagements with it, focused mainly on Kant's theoretic philosophy and the Scholastic tradition. An exception was a group of philosophically minded natural scientists, particularly Zvonimir Richtmann (1901–1941) and Rikard Podhorsky (1902–1994), who, in the 1930s, adopted and promoted doctrines of logical positivism and confronted the orthodox Marxist line of thought (for details, see Kovačević 1989; cf. also Lazović and Pavković 1998: 53; Pavković 1988: 6–7).

After the War, being ideologically and politically close to the Eastern Communist Block, the Marxist line of thought prevailed in Croatia (and the rest of Yugoslavia), just as in other countries of the Block. During the 1950s, after Yugoslavia's 1948 parting with the USSR, the prevailing orthodox form of Marxism loosened gradually, allowing for (primarily) the German continental philosophical influence to interfere, first of all its Hegelian and Heideggerian orientation. The prospects for the analytic orientation in the period were clearly dim (for elaboration, see Lazović and Pavković 1998: 53; Mišćević 1995: 172; Pavković 1988). That is, for example, reflected in the institutional undermining of the plans of the leading Croatian Marxist philosopher of the subsequent period, Gajo Petrović (1927–1993), to engage with analytic philosophy and continue in the late 1950s his PhD study with Alfred J. Ayer in London (Dožudić 2024a: 819). Despite that, Petrović did continue his engagements with the analytic tradition and in 1960, he made one of the first-ever translations of Wittgenstein's *Tractatus*. He was (allegedly) also the first in the former Yugoslavia to use the term “analytic philosophy” (Pavković 1988: 10; cf. Petrović 1964).

In close connection to Petrović, at least as far as their academic positions are concerned, in the 1970s, two teaching assistants at the Zagreb philosophy department, Goran Švob (1947–2013) and Neven Sesardić (b. 1949), started to promote analytic philosophy. The former of them was primarily oriented on Frege's work and themes in the philosophy of logic in general, and the latter on the philosophy of science, of mind, and epistemology (cf. Dožudić 2024a: 802; Lazović and Pavković 1998: 53; Miščević 1995: 172; Pavković 1988: 20–21; Sesardić 1978). But as far as the Zagreb department was concerned, despite the well-recorded later influence of the two, analytic philosophy remained at the academic margins, apparently more so than at other Yugoslav departments in Belgrade, Ljubljana, and – Zadar.

3. A women-managed philosophy department and its analytic turn

By the time logical positivism became hard to overlook in the European philosophy scene, Croatian philosophers employed at the Zagreb University had already adopted different orientations. For almost ninety years (from 1874 onwards), the Zagreb philosophy department was the only one in Croatia. Several women philosophers graduated with a PhD in that period, all of them before World War II. After the War, the first woman completed her PhD program in philosophy in 1965, and the first woman became a member of the philosophy department (as an assistant) in 1976 (the first with a higher academic rank, an associate professor, was employed in 1984). For over a century, the Zagreb philosophy department (for years, the only philosophy department in Croatia) was exclusively male-employed. However, in the early 1960s, another philosophy department was established in Croatia at the Zadar Faculty of Philosophy.

The first head of the Zadar philosophy department was Marija Brida (1912–1993). She graduated with a PhD in 1935 and worked within a realist phenomenological tradition, different from Husserl's. The position was primarily developed by her mentor, Pavao Vuk-Pavlović, who adopted the orientation in the 1920s, influenced significantly by Meinong's theory of objects (e.g., Vuk-Pavlović 1928; 1930; for details, see Lazović and Pavković 1998: 51; Mađarević 2005: 96–98; Miščević 1995; 1999). Brida had no direct interest in analytic philosophy; her writings were anything but analytic. Nevertheless, as a long-term head of the department (until her retirement), she did play a role in establishing the analytic tradition in Zadar in the 1970s and

1980s. Before turning to philosophy, Brida studied mathematics and physics, eventually adopting a science-oriented anti-Heideggerian perspective with which the analytic orientation matched well (Dožudić 2024a: 809–812; 2024b: 912–913). Brida was the only professor permanently employed at the Zadar philosophy department for seven years. In 1967, another woman philosopher, Heda Festini (1928–2018), permanently joined the department as an associate professor, and for the next eight years, Brida and Festini were the department's core. Brida retired in 1982, and for the following ten years, Festini served as the head of the department.

By 1971, Festini adopted the analytic orientation, prompting a departmental reorganisation according to it. In 1971, the department's initial, more traditional arrangement was replaced by its sub-departmental arrangement. Now, there was one sub-department for the history of philosophy and the other for logic and methodology of science. The latter subdepartment consisted of analytically oriented courses, such as logic, philosophy of science, epistemology, and philosophy of language. In 1975, Nenad Mišćević (1950–2024) joined the department as a teaching assistant and was permanently employed as an associate professor in the early 1980s. In the meantime, he abandoned his early anti-realistically oriented continental philosophical approach and preoccupation with contemporary French philosophy, embracing an analytic, realistically oriented approach (Dožudić 2024a: 788–792; 2024b: 896–900). Soon after, in 1979, another woman philosopher joined the department, Vanda Božičević (b. 1952), whose first book was on Susanne Langer's aesthetics (Božičević 1983). Thus, for some fifteen or twenty years, the Zadar philosophy department was exclusively a centre managed by women philosophers and, by the 1980s, primarily analytically oriented.

According to Festini, her first interest in analytic philosophy concerned the philosophy of science. Next, in 1971, she was enrolled as a researcher at the Institute of Philosophy in Zagreb, and for several years, she worked on the logical writings of the Croatian philosopher Albino Nagy (1866–1901). She placed Nagy's work in a broader context, associating it with the authors working on the foundations of mathematics (e.g., Peano, Frege, Dedekind, Hilbert, Brouwer, and Gödel) and a number of analytic authors (e.g., Russell, Popper, Reichenbach, Quine, Strawson, and Sellars). In that period, her interests changed, and for the next fifteen years, Wittgenstein and the themes

related to his philosophy preoccupied her research (e.g., Festini 1978; 1982; 1985; 1988; 1992; for discussion, see Dožudić 2024a: 793–796, 800–806).

Philosophers generally adopted the analytic mindset in institutions where analytic philosophy already had established proponents, but this was not the case for Festini. During her student and graduate years (the mid-1940s to mid-1960s), her academic environment in Zagreb was all but analytic. So, why had she adopted the analytic orientation and started promoting it? Festini wrote her PhD thesis on Nicola Abbagnano's positive existentialism. Abbagnano was an atypical existentialist who, with time, distanced himself from German and French existentialists, transforming his position into a "scientific methodology", declaring "his affinities with the neo-positivistic and neo-naturalistic movements in the Anglo-American world" (Romanell 1967: 2). Thus, during the 1960s, Festini dived deeper into the analytic tradition and started to incorporate it in her subsequent work.

Interestingly, in addition to Abbagnano's writings directing Festini's interest towards analytic philosophy, they also defined, as it seems, her interests within it. Initially, she developed an interest in the philosophy of science, promoting it when the Heideggerian anti-science perspective still dominated the region. Then she took seriously Abbagnano's view that later Wittgenstein's idea of a use-dependent meaning is plausible, but that Wittgenstein failed to provide an adequate analysis of the concept of use (Romanell 1967: 2). As a consequence, most of Festini's Wittgenstein-related writings in the period 1977–1992 are concerned with the problem of use, particularly in connection to linguistic meaning and (anti)realist positions. After her 1992 book on Wittgenstein, Festini's writings moved away from the Wittgenstein themes (Dožudić 2024a: 812–816). Nevertheless, Wittgenstein continued to be of interest to prominent members of the Zadar department from the 1980s in the following period (cf. Božičević 1995; Mišćević 1996), relocated in the meantime from Zadar elsewhere. Festini continued to write on analytic themes, now preoccupied with the Croatian philosophical heritage, especially the Renaissance and 19th-century Dalmatian philosophers.

4. Concluding remarks

In the 1961–1991 period, the Zadar philosophy department was an interesting phenomenon. In the 1980s, it was almost exclusively an analytically oriented

philosophy department; before that, it was predominantly analytic. – A unique phenomenon in Croatia during that period, probably unmatched by other similar departments in the former Yugoslavia and surely by many philosophy departments across Continental Europe. In addition, for over a decade, the department was run solely by two women philosophers who jointly, even if at times unintentionally, contributed to the department's subsequent analytic turn. And even after Mišćević became its leading analytic representative, Festini continued to occupy her vital role as the head of the department until her retirement in 1991, and after that, she remained supportive of the Croatian now-rearranged analytic scene (Dožudić 2024a; 2024b; Festini 1991; 1998; Jutrović 1994).

In his overview of the Yugoslav analytic tradition, Aleksandar Pavković posed an interesting question: “Why in the early seventies did the Marxists both in the academia and outside it, condone the appointment of analytic philosophers to a series of academic and research posts in Beograd, Zagreb, Zadar and Ljubljana?” (1988: 5). He does not provide an answer but suggest it would go beyond “the scope of history of philosophy proper”. As far as Croatia of the period is concerned, part of the reason is that some then-prominent and influential philosophers at the university, such as Festini's mentor Vladimir Filipović (1906–1984), (the already mentioned) Gajo Petrović and a member of Festini's doctorate committee Danilo Pejović (1928–2007), developed an interest (or at least a tolerance) for the analytic and other non-Marxist orientations (cf. Pejović 1967; Petrović 1964). Pavković also observes that, without the appointments, “only a few philosophers in Yugoslavia would [in the late 1980s] consider themselves analytic philosophers”. Looking from today's perspective, it is probably more likely that, without the appointments, more and more philosophers in the region would eventually consider themselves to be analytic, only that would be noticeable after 1991, in the period defined by deep social, cultural, and political changes in ex-Yugoslav republics.

In his overview, Zadar is mentioned once, and Pavković apparently associates it only with Mišćević (and Zagreb primarily with Sesardić). That is partly expected since, in the late 1970s and early 1980s, the two became the leading representatives of the analytic tradition in Croatia (cf. Lazović and Pavković 1998: 53; Mišćević 1995: 172; Pavković 1988: 20–22; Dožudić 2024a: 788–792).

However, neglecting the rest of the Zadar philosophy department, including the period before Mišćević's arrival and his subsequent analytic turn, would be wrong. Although less influential than Mišćević and not as widely recognised as him, Festini played a vital role in the Zadar analytic story. Indeed, reconstructing the introduction, promotion, and development of analytic philosophy in Croatia in general without taking Festini into account would be flawed. Mišćević's remark that Festini "stands at the beginning of the only institutionalized home analytic tradition; if the tradition goes on successfully, her name will be written on it in golden letters" (2019: 585) supports that. And now, one might add to the earlier question the following one: How come that over a decade, a Croatian philosophy department was dominantly, at times even exclusively, a women-manged department (unlike other similar departments in Croatia and across the rest of former Yugoslavia)? That question is also worth untangling, and the answer to it would also likely go beyond the scope of the history of philosophy proper.

References

- Beiser, Frederick C. (2022) *Johann Friedrich Herbart: Grandfather of Analytic Philosophy*, Oxford and New York: Oxford University Press.
- Božičević, Vanda (1983) *Estetika Susanne K. Langer* [Aesthetics of Sussane K. Langer], Zagreb: Zavod za filozofiju Filozofskog fakulteta Sveučilišta u Zagrebu.
- Božičević, Vanda (1995) "Tractarian Support for a Nonpropositional Paradigm", *Acta Analytica* 10(13), 191–207.
- Brožek, Anna (2017) "Maria Kokoszyńska: Between the Lvov-Warsaw School and the Vienna Circle", *Journal for the History of Analytical Philosophy* 5(2), 19–36.
- Brožek, Anna (2022) "Two Papers by Maria Ossowska", *Filozofia Nauki* 30(1), 131–133.
- Dožudić, Dušan (2024a) "Heda Festini i počeci analitičke filozofije u Hrvatskoj" [Heda Festini and the Beginings of Analytic Philosophy in Croatia], *Prilozi za istraživanje hrvatske filozofske baštine* 50(2), 893–919.
- Dožudić, Dušan (2024b) "Predanost, ustrajnost i analitički duh: Nenad Mišćević, 1965–2024" [Commitment, Persistence, and an Analytic Spirit: Nenad Mišćević, 1965–2024], *Prilozi za istraživanje hrvatske filozofske baštine* 50(2), 787–821.

Festini, Heda (1978) "One Semantical Consequence of Wittgenstein's Concept of Proposition as Logical Picture", in: Elisabeth Leinfellner et al. (eds.), *Wittgenstein and His Impact on Contemporary Thought: Proceedings of the Second International Wittgenstein Symposium 29th August to 4th September 1977*, Wien: Hölder-Pichler-Tempsky, 139–142.

Festini, Heda (1982) "Is Wittgenstein's Concept of Proposition in *Philosophical Grammar* Intensional, Extensional, or Something Else?", in: Werner Leinfellner et al. (eds.), *Language and Ontology: Proceedings of the Sixth International Wittgenstein Symposium 23rd to 30th August 1981*, Wien: Hölder-Pichler-Tempsky, 498–500.

Festini, Heda (1985) "Dummett's Conception as Theory of Meaning for Hintikka's Type of Game-theoretical Semantics (I) ('Use' and 'Language-game' in Wittgenstein and Dummett)", in: Georg Dorn and Paul Weingartner (eds.), *Foundations of Logic and Linguistics: Problems and Their Solutions*, New York: Springer, 639–665.

Festini, Heda (1988) "Antirealism/Realism of Wittgenstein's Language-Game Idea", *Radovi* 27(4), 5–20.

Festini, Heda (1991) "Analitička filozofija zadarskog kruga" [Analytic Philosophy of the Zadar Circle], *Filozofska istraživanja* 11(1), 87–89,

Festini, Heda (1992) *Uvod u čitanje Ludwiga Wittgensteina* [Introduction to a reading of Ludwig Wittgenstein], Zagreb: Hrvatsko filozofsko društvo.

Festini, Heda (1998) "Analitička filozofija u nas" [Analytic Philosophy in Our Country], *Scopus* 3(1), 93–97.

Gabriel, Gottfried (2013) "Frege and the German Background to Analytic Philosophy", in: Michael Beaney (ed.), *The History of Analytic Philosophy*, Oxford and New York: Oxford University Press, 280–297.

Jutrović Tihomirović, Dunja (1994) "Novija analitička filozofija u Hrvatskoj" [Recent Analytic Philosophy in Croatia], *Filozofska istraživanja* 14(2-3), 585–591.

Kovač, Srećko (1996) "Franjo pl. Marković: On the Hundred and Fiftieth Anniversary of His Birth", *Studia Historiae Philosophiae Croaticae* 3(3), 169–188.

Kovačević, Božo (1989) *Slučaj zagrebačkih revizionista* [The Case of the Zagreb Revisionists], Zagreb: Grafički zavod Hrvatske.

Lazović, Živan; Pavković, Aleksandar (1998) "South Slavs, Philosophy of", in: Edward Craig (ed.), *Routledge Encyclopedia of Philosophy* Vol. 9, London and New York: Routledge, 48–56.

Mađarević, Lovorka (2005) "Croatian Philosophers III: Pavao Vuk-Pavlović (1894–1976)", *Prolegomena* 4(1), 93–105.

Marotti, Bojan (2020) *Marković's Critique of Hegel's Logic: Franjo Marković and the so-called "Austrian Realism"*, Zagreb: ArTresor.

Miščević, Nenad (1995) "Croatian Philosophy", in: Ted Honderich (ed.), *The Oxford Companion to Philosophy*, Oxford and New York: Oxford University Press, 171–172.

Miščević, Nenad (1996) "Computationalism and the Kripke-Wittgenstein Paradox", *Proceedings of the Aristotelian Society* XCVI, 215–230.

Miščević, Nenad (1999) "Ontologija Pavla Vuk-Pavlovića: Vukova kritika spoznajnoteorijskog polazišta" [The Ontology of Pavao Vuk-Pavlović: Vuk's Criticism of the Standpoint Founded on the Theory of Knowledge], *Filozofske istraživanja* 74(3), 445–451.

Miščević Nenad (2019) "Constructing a Happy City-State. In memoriam Heda Festini", *Croatian Journal of Philosophy* XIX(57), 583–596.

Pavković, Aleksandar (1988) "Introduction", in: Aleksandar Pavković (ed.), *Contemporary Yugoslav Philosophy: The Analytic Approach*, Dordrecht, Boston, and London: Kluwer Academic Publishers, 5–26.

Pejović, Danilo (1967) *Suvremena filozofija Zapada* [Contemporary Western Philosophy], Zagreb: Matica hrvatska.

Petrović, Gajo (1964) *Od Locke do Ayera* [From Locke to Ayer], Beograd: Kultura.

Romanell, Patrick (1967) "Abbagnano, Nicola", in: Paul Edwards (ed.), *The Encyclopedia of Philosophy* Vol. 1, New York: Macmillan Publishing Co. Inc. & The Free Press, 1–2.

Sesardić, Neven (1978) "O analitičkoj filozofiji" [On Analytic Philosophy], *Pitanja* 10(7-8), 46–53.

Van der Schaar, Maria; Schliesser, Eric (2017) "Women in Early Analytic Philosophy: Volume Introduction", *Journal for the History of Analytical Philosophy* 5(2), 1–15.

Verhaegh, Sander; Peijnenburg, Jeanne (2022) "Introduction: Women in the History of Analytic Philosophy", in: Jeanne Peijnenburg and Sander Verhaegh (eds.), *Women in the History of Analytic Philosophy: Selected Papers of the Tilburg–Groningen Conference, 2019*, Cham: Springer, 1–21.

Vuk-Pavlović, Pavao (1928) "Spoznajna teorija i metafizika" [Theory of Knowledge and Metaphysics], *RAD Jugoslavenske akademije znanosti i umjetnosti* 235, 50–58.

Vuk-Pavlović, Pavao (1930) "Erkenntnistheorie und Metaphysik", *Bulletin International de l'Académie des sciences et des arts des Slaves du Sud de Zagreb* 3, 107–112.

Acknowledgement

This paper was produced within the Zagreb Institute of Philosophy projects *Antipsychologistic Conceptions of Logic and their Reception in Croatian Philosophy* (APsiH, reviewed by the Croatian

Ministry of Science and Education and financed through the National Recovery and Resilience Plan by the European Union – NextGenerationEU) and *Intentionality and Modes of Existence* (IP-2022-10-5915, financed by the Croatian Science Foundation). Thanks to Tomislav Bracanović and Neven Sesardić for comments on the manuscript.

Vom Denken zum Handeln: Verbindungen vom Einschleichen der Kunst nach Wittgenstein zu feministischer Kunst

Lilli Förster (Stuttgart, Deutschland)

Abstract

Ludwig Wittgensteins Sprachphilosophie ist in der feministischen Auseinandersetzung mit den Zusammenhängen zwischen Denken, Sprache und gesellschaftlichen Praktiken und Strukturen fruchtbar gewesen – warum sollte dies nicht auch auf seine Gedanken und Ausführungen zur Kunst zutreffen? Genauer, wie lässt sich der von Wittgenstein in seinen späteren Schriften formulierte Begriff des Einschleichens der Kunst in unser Leben anwendbar machen für feministische Ansätze in der Kunst?

Während der Autor des *Tractatus* lediglich der Sprache Mitteilungswert zuschreibt, erweiterte Wittgenstein seine Auffassung später zugunsten der Mitteilungsfähigkeit von Kunst – auch wenn die Art der Mitteilung eine andere ist: Kunstwerke können etwas über das Leben selbst mitteilen. Darüber hinaus sind sie in der Lage, sich in unser Leben einzuschleichen. Wittgenstein beschreibt am Beispiel von Musik: „Sie schleicht sich in mein Leben ein. Ich mache sie mir zu eigen.“ (VB 553f.) Damit bergen sie die Möglichkeit in sich, uns von innen heraus zu verändern – unseren Denkstil und damit unseren Lebensstil, unsere Sichtweisen und Verhaltensweisen zu transformieren. Und das ist wichtig, wie Wittgenstein schon im *Tractatus* festgestellt hat: Wie wir denken und sprechen gestaltet unsere Welt und ihre Grenzen.

Diese transformatorische Kraft kann in Beispielen der feministischen Kunst wiedergefunden werden, um so neue Perspektiven und Zugänge zu diesen aufzuzeigen. Dies soll über zwei Zugänge erfolgen: Feministische Kunst, die ihre (i) Inhalte beispielsweise aus der kritischen Auseinandersetzung mit gesellschaftlichen Machtstrukturen, Geschlechterverhältnissen und Identitätsfragen zieht, setzt häufig, in Bezug auf die (ii) Methode, auf subtile, indirekte und transformative Strategien.

Ludwig Wittgensteins Sprachphilosophie ist in der feministischen Auseinandersetzung mit den Zusammenhängen zwischen Denken, Sprache und gesellschaftlichen Praktiken und Strukturen fruchtbar gewesen – warum sollte dies nicht auch auf seine Gedanken und Ausführungen zur Kunst zutreffen? Genauer, wie lässt sich der von Wittgenstein in seinen späteren Schriften formulierte Begriff des Einschleichens der Kunst in unser Leben anwendbar machen für feministische Ansätze in der Kunst?

Während der Autor des *Tractatus* lediglich der Sprache Mitteilungswert zuschreibt, erweiterte Wittgenstein seine Auffassung später zugunsten der Mitteilungsfähigkeit von Kunst – auch wenn die Art der Mitteilung eine andere ist: Kunstwerke können etwas über das Leben selbst mitteilen. Darüber hinaus

sind sie in der Lage, sich in unser Leben einzuschleichen. Wittgenstein beschreibt am Beispiel von Musik: „Sie schleicht sich in mein Leben ein. Ich mache sie mir zu eigen.“ (VB 553f.) Damit bergen sie die Möglichkeit in sich, uns von innen heraus zu verändern – unseren Denkstil und damit unseren Lebensstil, unsere Sichtweisen und Verhaltensweisen zu transformieren. Und das ist wichtig, wie Wittgenstein schon im *Tractatus* festgestellt hat: Wie wir denken und sprechen gestaltet unsere Welt und ihre Grenzen.

Diese transformatorische Kraft kann in Beispielen der feministischen Kunst wiedergefunden werden, um so neue Perspektiven und Zugänge zu diesen aufzuzeigen. Dies soll über zwei Zugänge erfolgen: Feministische Kunst, die ihre (i) Inhalte beispielsweise aus der kritischen Auseinandersetzung mit gesellschaftlichen Machtstrukturen, Geschlechterverhältnissen und Identitätsfragen zieht, setzt häufig, in Bezug auf die (ii) Methode, auf subtile, indirekte und transformative Strategien.

Bevor sich diesem Vorhaben gewidmet wird, ein paar Vorabbemerkungen: der vorliegende Text dient als erste Skizze neue Zugänge zu feministischen Kunstwerken ausgehend von einer offenen Interpretation Wittgensteins Einschleichens von Kunstwerken in unser Leben aufzuzeigen. Die Wahl der nachfolgenden Beispiele feministischer Kunst beruht auf rein persönlichem Interesse und ist keiner kunsthistorischen Fachexpertise geschuldet. Letztlich soll betont werden, dass die Methode des Einschleichens in feministischen Auseinandersetzungen auch außerhalb des Kunstbereichs eine und nicht die Methode der Mitteilung dargestellt.

1. Wittgenstein zu Sprache und Kunst

Bevor auf die im frühen und späten Werk Wittgensteins geäußerte Bemerkungen eingegangen werden kann, möchte ich eine kurze Einordnung der folgenden Gedanken des Aufsatzes in die fortlaufende Debatte zwischen den verschiedenen Lesarten seiner Werke und der Einschätzungen ob und wann es Übergänge zwischen einem frühen, mittleren und späten Wittgenstein gab, vornehmen.

Die Unterschiede im Denken Wittgensteins, beispielsweise über die Funktion bzw. Funktionen der Sprache, zwischen seinen frühen Äußerungen im *Tractatus* und seinen späteren, sollen nicht vernachlässigt werden. Trotzdem erscheint es seltsam aus einer Person zwei Autoren zu machen. Wittgenstein

ist ein Mensch, dessen Denken sich mit den Jahren weiterentwickelt hat. Damit gehen Verfeinerungen, Verwerfungen und Korrekturen vergangener Ansichten einher – ein natürlicher Prozess. Das Nachvollziehen dieser Gedankengänge ist für die folgende Thematik dieses Aufsatzes wertvoll. Daher sollen seine Auffassungen mit Ernst gelesen und mitgedacht werden. Um die Abfolge der Denkprozesse einzuhalten, werden die Begrifflichkeiten, die sich erst in seinen späteren Schriften wiederfinden, nachfolgend auch nur in diesen Kontexten verwendet.

Der Autor des *Tractatus* setzte mit seinem Werk harte Grenzen zwischen dem, was klar gesagt werden kann und dem, was nicht auf diese Art und Weise mitgeteilt werden kann. Er schreibt im Vorwort:

Was sich überhaupt sagen läßt, läßt sich klar sagen; und wovon man nicht reden kann, darüber muß man schweigen. Das Buch will also dem Denken eine Grenze ziehen, oder vielmehr - nicht dem Denken, sondern dem Ausdruck der Gedanken: Denn um dem Denken eine Grenze zu ziehen, müßten wir beide Seiten dieser Grenze denken können (wir müßten also denken können, was sich nicht denken läßt). (T Vorwort).

Um dem Anliegen dieses Aufsatzes nachgehen zu können, ist zunächst ein Blick auf genau diese Grenzziehung notwendig.

„Wir machen uns Bilder der Tatsachen“ (T 2.1) und aus diesen Tatsachen besteht die Welt (T 1.1) Im Ernst gesprochene Sätze, wie beispielsweise die in den Naturwissenschaften, sind sinnlich wahrnehmbare, logische Bilder der Wirklichkeit. (T 4.01) Denken und Sprache sind, wie hier deutlich wird, unmittelbar verbunden. (T 3) Wir sind in der Lage Sätze ohne weitere Hilfsmittel oder Erklärungen verstehen zu können. Nur dadurch sind wir, dem Verfasser des *Tractatus* zu Folge, im Alltag in der Lage miteinander zu kommunizieren – Sprache „ist Teil des menschlichen Organismus und nicht weniger kompliziert als dieser.“ (T 4.002) Wir sind so in der Lage unsere Gedanken, unsere lebendigen Bilder von Tatsachen, anderen Menschen direkt mitteilen zu können. Um dies gewährleisten zu können, muss die Sprache den gesamten Bereich des Möglichen darstellen können. Die Welt wird durch die Sprache geordnet und strukturiert: „Wie ich meine Sprache durch ihren Gebrauch gestalte, das bestimmt zugleich den Aufbau meiner Welt.“ (Schulte

1989, 90) Das, was also von Sprache mitgeteilt werden kann, betrifft die Tatsachen der Welt. Logik, als Gerüst der Welt, lässt sich daraus folgend ebenso wenig sprachlich fassen, wie Ethik oder Ästhetik. (T 6.124, 6.42, 6.421) Was bedeutet dies für die Kunst?

Zunächst vertrat Wittgenstein, seinem Lehrer und Mentor Gottlob Frege folgend, die Auffassung, dass Kunst und ihre Werke den Menschen nichts über die Welt und die Tatsachen in ihr mitteilen können. Am 20.10.1916 notierte er in sein Tagebuch den bekannten Satz aus Schillers *Wallensteins Lager*: „Ernst ist das Leben, heiter ist die Kunst.“ Es soll jedoch betont werden, dass Kunst und ihre Werke durch diese Einschätzung zu ihrer Vermittlungsfähigkeit nicht wertlos sind; sie stehen aber auf einer vollkommen anderen Ebene als die Sprache.

Der Autor der *Philosophischen Untersuchungen*, im Gegensatz zu seinen Ausführungen im *Tractatus*, erweitert seine Ansichten in Bezug auf die Zusammenhänge von Denken, Sprache und der Welt. Keine Rede ist mehr von Elementarsätzen aus denen alle weiteren Sätze zusammengesetzt werden können. Hingegen aber von Sprachspielen, die nicht isoliert betrachtet werden können, sondern eingebettet in die menschlichen Lebensformen sind. (PU 23) Die Sprache wird zu einem Werkzeugkasten (PU 11) womit ihr mehr als eine Aufgabe zuteilwird. Und mit der sogenannten therapeutischen Methode versucht Wittgenstein den unter Verhexungen der Sprache leidenden PhilosophInnen einen Ausweg aus dem Fliegenglas zu zeigen. (PU 133, 109, 309)

Diese Veränderung des Denkens und Herantretens an Sprache schlägt sich auch in seinen erweiterten Ansichten in Bezug auf Kunst nieder. Wittgenstein kam zu der Erkenntnis, dass Kunst näher an der Sprache ist, als sein früheres Ich zugeben würde. (Milkov 2020, 507) Er erkannte: wie Sprache uns etwas mitteilen kann, so können es auch Kunstwerke. „Die Menschen heute glauben, die Wissenschaftler seien da, sie zu belehren, die Dichter und Musiker etc., sie zu erfreuen. Daß diese sie etwas zu lehren haben; kommt ihnen nicht in den Sinn.“ (VB 501) Das, was Kunstwerke mitzuteilen haben, bezieht sich dabei auf das Leben selbst, die Lebensführung, Denkweisen und Lebensstile. Wie kann so eine Mitteilung durch Kunst erfolgen?

2. Das Einschleichen der Kunst in unser Leben

Wittgenstein spricht davon, dass sich Musik in das Leben des Menschen „einschleichen“ kann, (VB 553f.) sich also den Weg in das Leben des Hörenden bahnen. Ähnlich wie das Verstehen von Sprache ist dieses Phänomen für uns natürlich. Im alltäglichen Sprachgebrauch findet sich dies ebenfalls wieder: „einschleichen“ wird in Verbindung gesetzt mit einschleichen in Gedanken oder in das Vertrauen eines Anderen. Medikamente schleichen sich ein und aus; Diebe schleichen sich ein. Dies passiert unbemerkt, heimlich, teils über einen längeren Zeitraum.

Bevor sich der Konsequenzen dieses Einschleichens in Bezug auf Kunstwerke gewidmet werden kann, soll skizziert werden, warum eine von Wittgenstein über Musik getroffene Aussage auch auf andere Bereiche in der Kunst übertragen werden kann:

Für Wittgenstein ist die Musik „in gewissem Sinne die raffinierteste aller Künste“ (Wittgenstein, 2022, 180). Mit folgender Einordnung soll betont werden, dass die zwischen den Kunstformen bestehenden Unterschiede unangetastet bleiben. Nichtsdestotrotz spricht Wittgenstein in seinen Bemerkungen immer wieder über mehrere Kunstformen in einem Atemzug – so folgen Gedanken über Melodien hören im nächsten Satz eine Bemerkung zum Wirkkonzept von der Kunst im Allgemeinen in Zusammenhang mit Gefühlsäußerungen (Wittgenstein, 2022, 131). Gesondert betont er: „Ja, auch [einem Menschen] Verständnis für Gedichte oder Malerei beibringen, kann zur Erklärung dessen gehören, was Verständnis für Musik sei.“ (Wittgenstein, 2022, 143) Besonders in diesem Zusammenhang hervorgehoben werden soll Wittgensteins bereits zitierte Äußerung, dass „[d]ie Menschen heute glauben, die Wissenschaftler sein da, sie zu belehren, die Dichter und Musiker etc., sie zu erfreuen. Daß diese sie etwas zu lehren haben; kommt ihnen nicht in den Sinn.“ (VB 501) Unabhängig der Kunstform, Dichtung oder Musikstück in diesem Falle, kann Kunst uns etwas lehren, etwas mitteilen.

Die Besonderheit dieser Art der Mitteilung durch Kunstwerke liegt in ihrem passiven, nicht unmittelbar beeinflussbaren Moment. Ähnlich wie Leo Tolstoi den Begriff des „Infizierens“ versteht (Tolstoi 1993), erfolgt das Einschleichen von Kunstwerken zunächst unwillentlich und unmittelbar. Man kann sich dem Einschleichen des Kunstwerkes in das eigene Leben nicht entziehen. Die

Ebenen, auf die sich die Kunst einzuschleichen vermag, liegen tief in uns. Andere, beispielsweise auf wissenschaftlicherer Ebene basierende Ansätze, erreichen diese tiefen Ebenen nicht, denn hier liegen die „Probleme des Lebens,“ und diese sind „an der Oberfläche unlösbar.“ (VB 555) Wittgenstein betont bereits im Tractatus: „Wir fühlen, daß selbst, wenn alle möglichen wissenschaftlichen Fragen beantwortet sind, unsere Lebensprobleme noch gar nicht berührt sind.“ (T 6.52) Was Wissenschaften für den Menschen leisten können, betrifft die Tatsachen der Welt. Kunstwerken hingegen gelingt der Zugang zu Fragen des Lebens selbst.

Nach dem Einschleichen der Kunstwerke, bleibt unsere Auseinandersetzung mit ihnen nicht auf einer passiven Ebene stehen – wir können und sollten über das, was uns passiert reflektieren. Wir können entscheiden, wie mit dem, was das Kunstwerk in uns hineingetragen hat, umzugehen ist, wie wir es bewerten, was wir davon annehmen und was wir ablehnen. Deutlich wird dies, wenn Wittgenstein am Beispiel der Musik gleich nach: „Sie schleicht sich in mein Leben ein.“, betont: „Ich mache sie mir zu eigen.“ (VB 553f.) Wichtig an dieser Stelle ebenfalls zu betonen: das Kunstwerk überträgt nur sich selbst (VB 533). Alles, was das Kunstwerk, hier das Thema eines Musikstücks, ausmacht ist in ihm enthalten und das bezieht sich auch auf „de[n] Rhythmus unserer Sprache, unseres Denkens und Empfindens.“ (VB 523)

In dieser beschriebenen Kombination aus passiver Art und Weise des Einschleichens in unser Leben und die aktive und bewusste Auseinandersetzung und des Zu-eigen-Machens liegt die Stärke der Künste. Es folgt, dass wenn Kunstwerke also uns über das Leben selbst etwas mitteilen und durch das Einschleichen Teil von uns werden, sie in der Lage sind unser Leben in allen Ebenen zu beeinflussen und zu verändern. Diese Beeinflussung betrifft beispielsweise unsere Art und Weise Dinge und bestimmte Sachverhalte zu sehen, wahrzunehmen und zu verstehen, aus welcher Perspektive wir ihnen gegenüber treten und wie wir uns in Bezug auf sie fühlen. Kurzum – sie beeinflusst unseren Denkstil richtungsweisend, was wiederum gesamtgesellschaftliche Auswirkungen hat.

Erwähnenswert ist an dieser Stelle, dass Wittgenstein in seinen *Vorlesungen über Ästhetik* genau an dieser Stelle eine seiner wichtigsten Aufgaben sieht:

Wieviel von dem, was wir tun, besteht darin, den Stil des Denkens zu ändern, und wieviel von dem, was ich tue, besteht darin, den Stil des Denkens zu ändern, und wieviel tue ich, um andere davon zu überzeugen, ihren Denkstil zu ändern. (Vieles von dem, was wir tun, ist eine Frage der Änderung des Denkstils.) (VÄ 45)

Dieser prägt unsere Sichtweise auf, unsere Einstellung zum Leben selbst und somit auch unseren Lebensstil – unser alltägliches Denken und Handeln gleichermaßen. Es sei an den Autor des *Tractatus* erinnert: Denken, Sprache und Welt sind direkt miteinander verbunden. Wie wir denken und sprechen gestaltet unsere Welt und ihre Grenzen. Über Kunst und den Prozess ihres Einschleichens in uns diese zu verändern, unterstreicht den Stellenwert, den wir Kunstwerken in unserem Leben einräumen können und sollten.

3. Feministische Kunst und Wittgenstein

Kunsthistorisch bezieht sich „Feministische Kunst“ auf die Künstlerinnen ab den 70er Jahren, die, im Zuge der zweiten großen feministischen Welle des 20. Jahrhundert, sich mit der Erfahrung des Frau-Seins, weiblich-gelesen-Werdens, weiblicher Identität und Geschlechterrollenverhältnisse und deren Konstruktion, bzw. Dekonstruktion auseinandersetzen. Der Kunstbetrieb selbst wird in diesem Zuge ebenfalls auf bestehende Normen und Werte kritisch hinterfragt. (Zimmermann 2008)

In diesem künstlerischen Prozess werden bestehende kulturelle Praktiken, ebenso wie die dahinter liegenden gesellschaftlichen Strukturen und Machtverhältnisse aufgezeigt und kritisch infrage gestellt. Damit kann das Aufzeigen alternativer Sichtweisen und Handlungsweisen einhergehen. Heutige Strömungen finden zwar ihre Wurzeln in diesen Bewegungen, sind jedoch vielseitiger in ihren Ausgestaltungen. Zudem sei erwähnt, dass die Auseinandersetzung mit den angesprochenen Themen auch vor den 70er Jahren stattfand.

Nach den im *Tractatus* herausgestellten Zusammenhängen von Denken, Sprache und Welt, betont Wittgenstein später die Bedeutung der Veränderung der Sprache und die damit einhergehenden Veränderungen von Lebenspraxen. Dieser Basis für gesellschaftliche Veränderungen und Transformationsprozesse gilt es im feministischen Sinne abzuleiten. Das Kunstwerk sensibilisiert für eben die Lebenspraxen in denen es eingebettet

liegt, auch außerhalb der Werklogik. So wollen wir im Folgenden nachzeichnen, wie das eben skizzierte Phänomen des Einschleichens von Kunstwerken und deren zu Eigen-machen fruchtbar gemacht werden kann für die Ziele der feministischen Kunst. Zwei größeren Zugängen soll in dieser Skizze nachgegangen werden. Betont werden soll zuvor jedoch, dass die hier gegenwärtigen Persönlichkeiten keine StellvertreterInnenrollen einnehmen, sondern beispielhaft mögliche Zugänge eröffnen sollen.

4. Das Einschleichen als Methodik

Unter Sneaky Feminism oder Stealth Feminism sind bereits Methoden bekannt, die beispielsweise feministische Ideale mit anderen Themen kombinieren, um diese breitenwirksam in den gesellschaftlichen Diskurs zu integrieren. Angewendet auf den Bereich der Kunst unter Berücksichtigung des hier vorgestellten Prozesses des Einschleichens zeigt sich eine weitere Stärke dieses Ansatzes. Wie Ausführungen zum Status-Quo-Bias, Systemrechtfertigungstheorie und psychologischer Reaktanz andeuten, können radikale, abrupte Veränderungen gesamtgesellschaftlich als Bedrohungen wahrgenommen werden, was subtile, weniger direkt auf Konfrontation gehende Herangehensweisen vielversprechender erscheinen lässt. Bestehende Strukturen zu unterwandern, von innen heraus zu transformieren stellt so eine Möglichkeit des Versuchs des gesellschaftlichen Wandels dar.

Einen subtilen und persönlichen Ansatz wählt die Performance-Künstlerin Adrian Piper in ihrer Auseinandersetzung mit Identität, Rassismus und Geschlecht. Mit ihren „Calling Cards“ unterbricht sie Alltagssituationen, ohne in offene Konfrontationen zu gehen. Ziel ist es, Menschen dazu zu bewegen sich ihrer Vorurteile bewusst zu werden und durch bewusste Anstöße neue Perspektiven und Gedankengänge aufzuzeigen. Dies tut sie, indem sie kleine Kärtchen in Situationen verteilt, in denen jemand in der Öffentlichkeit sexistische oder rassistische Kommentare von sich gibt.

In „My Calling Card #2“ (adrianpiper.weebly) wird gesondert dem Vorurteil nachgegangen, dass weiblich gelesene Personen sich in Bars nur aufhalten, um von männlich gelesenen Individuen angesprochen werden zu wollen. Sie thematisiert damit sowohl die notwendige Akzeptanz vom Allein-Sein von weiblich gelesenen Personen als auch das in Anspruch Nehmens von eigenem

Raum. Still, aber nicht wortlos wird der eigene Standpunkt aufgezeigt und so eine alltägliche, soziale Interaktion zur künstlerischen Intervention, die soziale Normen in Frage stellt.

5. Das Einschleichen als thematischer Zugang

Neben dem Einschleichen als methodischer Zugang zu feministischer Kunst, stellen die Inhalte und Themen dieser einen weiteren Anknüpfungspunkt dar. Das Alltägliche, das Leben selbst steht oftmals im Zentrum der künstlerischen Auseinandersetzung. Themen wie Care-Arbeit, das Häusliche, persönliche Erfahrungen mit Diskriminierung, Körpererfahrungen, um nur ein paar zu nennen. Das Einschleichen der Kunst ermöglicht subtile Zugänge auf die Ebenen, auf denen diese Themen verhandelt werden.

Die Künstlerin Louise Bourgeois ist bekannt für ihre an verschiedenen Orten auf der Welt stehenden, riesigen Spinnen-Skulpturen („Maman“), die neben Ur-Angst auch Mutterschaft, Schutz und Wiederinstandsetzung verkörpern. Mit „The Destruction of the Father“ (1974) hingegen setzt sie sich mit familiären Hierarchien und darin auftretenden Geschlechterdynamiken auseinander. Die Befreiung von männlicher Dominanz auf persönlicher und gesamtgesellschaftlicher Ebene steht im Zentrum.

Wie diese beiden Werke erkennen lassen, setzt sich die Künstlerin in ihren Skulpturen mit Weiblichkeit, Familie, Trauma, Identität und Sexualität auseinander. Ihre Zugänge sind dabei oftmals emotional und autobiografisch. Intime Innenräume baut sie groß im öffentlichen Raum und in Museen auf, um die dahinter liegenden Geschlechterstrukturen zu dekonstruieren. Das Beobachten und Erkennen eigener sozialer und familiärer Dynamiken sollen so angeregt werden.

Gesondert in ihrer Reihe „The Cells“ (1991-2010) kann die Verbindung von Persönlichem mit Gesellschaftlichem, Allgemeinem in Bezug auf die Themen Körper, Geschlecht, Schmerz und Ängste wahrgenommen werden. Intimität und Isolation werden Gegenstände und Stilmittel der Werke. Architektonisch begrenzte Räume stellen emotionale Mikrokosmen dar, in denen mit verschiedenen Materialien, wie etwa Kleidungsstücke, Möbel, Skulpturen aus Säcken oder gefundenen Gegenständen theaterartige Inszenierungen aufgebaut werden. Die Künstlerin selbst vermerkt: „Die 'Zellen' repräsentieren verschiedene Arten von Schmerz: physischen, emotionalen, psychologischen,

geistigen und intellektuellen Schmerz ... Jede 'Zelle' befasst sich mit dem Genuss des Voyeurs, mit dem Reiz des Sehens und Gesehenwerdens. Die 'Zellen' ziehen sich entweder an oder stoßen einander ab. Es gibt diesen Drang, sich zu verbinden, zu verschmelzen oder zu zerfallen." (HausderKunst)

6. Exkurs: Perspektivwechsel durch Aspekt-Sehen

Eine weitere Verbindung zwischen Wittgensteins Bemerkungen in den Philosophischen Untersuchungen und feministischer Kunst, leuchtete mir beim Kennenlernen von Mierle Laderman Ukeles künstlerischer Tätigkeit auf und soll hier als Exkurs angedeutet werden: die Aspektwahrnehmung.

Die Künstlerin nutzt in ihren Werken soziale Interaktion als künstlerisches Medium und wird dadurch vom New York Queens Museum als „Mutter“ der Maintenance Art bezeichnet. (DeutschlandFunkKultur) In Performences hebt sie die sonst im Verborgenen stattfindenden Akte der Pflege, Fürsorge und der Instandhaltung ins Rampenlicht. Diese meist von Frauen ausgeführten Care-Arbeiten erhebt sie in einem Manifest 1969 zu Kunst und fordert somit Anerkennung für diese ein. Ausgehend von ihren Vorbildern Pollock und Duchamp, die den schöpferischen, schaffenden Akt ins Zentrum rücken, möchte sie den Fokus auf die pflegende, erhaltende Seite verschieben. Alles am Laufen zu halten, soll als Tätigkeit sichtbar gemacht und gewürdigt werden. Mit dieser Veränderung der Wahrnehmung und des damit einhergehenden Perspektivwechsels möchte sie zu einem Umdenken in Bezug auf Instandhaltungsarbeit bewegen. In ihrer Arbeit „Touch Sanitation“ (1977-1980) schüttelte sie beispielsweise über 8500 Instandhaltungsarbeitern und Servicekräften bei Müllabfuhr, Winterdienst oder U-Bahn die Hände und dankte ihnen mit „Thank you for keeping NYC alive.“ Im Rahmen der COVID19-Pandemie fand dieser Akt der Würdigung in digitaler Version erneut statt.

Wittgenstein widmet sich der Aspektwahrnehmung im zweiten Teil der Philosophischen Untersuchungen und beschreibt dieses als „Sehen als“ – ein plötzliches Aufleuchten eines Aspekts, das nicht bloß eine Eigenschaft des Objekts ist, sondern aus der Relation zwischen Wahrnehmung, Erfahrung und Kontext entsteht: „Ich betrachte ein Gesicht, auf einmal bemerke ich seine Ähnlichkeit mit einem anderen. Ich sehe, daß es sich nicht verändert hat; und sehe es doch anders. Diese Erfahrung nenne ich ‚das Bemerkens eines Aspektes‘.“ (PU II 518). Auch das berühmte Beispiel des Enten-Hasen-Bildes

fällt hierrunter: Wer mit solchen Bildern nicht vertraut ist, sieht zunächst nur eine Ente oder nur einen Hasen, während eine geübte Person den Wechsel der Aspekte bewusst wahrnehmen kann. Dies zeigt, dass Aspektsehen nicht nur eine passive Wahrnehmung, sondern eine aktive Deutung ist.

In der ästhetischen Wahrnehmung kann Aspektwahrnehmung ebenfalls eine zentrale Rolle einnehmen. Dies kann zum einen im Werk selbst geschehen – jedoch auch außerhalb der Werklogik wirken: Kunstwerke lenken unsere Aufmerksamkeit auf bestimmte Merkmale und ermöglichen es uns, verschiedene Aspekte zu erkennen. In diesem Sinne ist Aspektwahrnehmung nicht nur eine Frage der individuellen Wahrnehmung, sondern auch eine kulturell und kontextuell geprägte Fähigkeit, die durch Erziehung und Erfahrung geschult werden kann. Hier teilt sich Wittgenstein seinen Aufruf, die Dinge anders zu sehen, mit den Kunschtchaffenden der feministischen Kunst: Aufzeigen andere Perspektiven, anderer Blickwinkel und Sichtweisen auf Etwas – ein Wechsel des Wahrnehmens bestimmter Aspekte.

Fazit

Die Werke von Künstlerinnen wie Adrian Piper, Louise Bourgeois und Mierle Laderman Ukeles greifen Themen auf, die tief in den alltäglichen Erfahrungen verwurzelt sind, und machen sie durch ihre künstlerische Auseinandersetzung mit diesen zugänglich und sichtbar. Ob durch die Dekonstruktion vertrauter Sehgewohnheiten, die ästhetische Kritik an patriarchalen Strukturen oder die Aufwertung unsichtbarer Carearbeit – feministische Kunstwerke fordern uns auf, unseren Denk- und somit Lebensstil kritisch zu hinterfragen. Die Veränderung unseres Denkstils zu bewirken kann Kunstwerken auf Grund ihrer Fähigkeit zugeschrieben werden, sich in uns einzuschleichen und von uns zu eigen gemacht zu werden. Wittgenstein hebt mit dem Phänomen des Einschleichens von Kunstwerken in unser Leben genau diese Kraft hervor. Anders als noch der Autor des *Tractatus*, sieht er später Kunstwerke durchaus in der Lage, den Menschen etwas mitzuteilen. Und dieses Etwas betrifft das Leben selbst. Kunstwerke als Teil gesellschaftlicher Transformationsprozesse wahrzunehmen und anzuerkennen ist die Konsequenz dessen.

References

Milkov, Nikolay (2020) „Aesthetic Gestures: Elements of a Philosophy of Art in Frege and Wittgenstein,“ in: Shyam Wuppuluri & Newton da Costa (Hrsg.), *Wittgensteinian (adj.) Looking at the World from the Viewpoint of Wittgenstein's Philosophy*. Berlin: Springer, S. 505–517.

Schulte, Joachim (2016) *Wittgenstein. Eine Einführung*, Stuttgart: Reclam, 1. Aufl. 1989.

Tolstoi, Leo (1993) *Was ist Kunst?*, München: Diederichs.

Wittgenstein, Ludwig (1984) *Tractatus logico-philosophicus, Tagebücher 1914–1916*, [u.a], Werkausgabe in 8 Bänden, Bd. 1, Frankfurt a.M.: Suhrkamp.

Wittgenstein, Ludwig (1984) *Vermischte Bemerkungen*. Werkausgabe in 8 Bänden, Bd. 8, Frankfurt am Main: Suhrkamp.

Wittgenstein, Ludwig (2005) *Vorlesungen und Gespräche über Ästhetik, Psychoanalyse und religiösen Glauben*, Frankfurt am Main: Fischer.

Wittenstein, Ludwig (2022) *Betrachtungen zur Musik*, Berlin: Suhrkamp.

Zimmermann, Anja (2008) „Was war feministische Kunst? Sechs Historisierungen im Lexionformat,“ in: Mostegl, S., Ratzinger, G. (Hrsg.) *Matrix*, Wien: Springer.

Onlinequellen:

<https://adrianpiper.weebly.com/>

<https://www.deutschlandfunkkultur.de/mierle-laderman-ukeles-die-begruenderin-der-maintenance-art-100.html>

<https://www.hausderkunst.de/eintauchen/louise-bourgeois-cells>

A Moral Inquiry into Unintentional Speech

Christina Fritz (Graz, Austria)

Abstract

In this contribution, I will explore the moral responsibility of individuals for spontaneous acts of speech, focusing on cases where speech occurs automatically without conscious intent. Using the example of an agent called Paul, who unintentionally makes a sexist remark in a moment of frustration, the discussion examines whether individuals should be held accountable for such speech despite lacking deliberate intent. The analysis considers the nature of automatic actions and speech, the distinction between mere utterances and genuine speech acts, and the moral weight of unintentional harm. I argue that even acts of speech performed without attending to a decision can carry moral significance, prompting a reassessment of moral responsibility and the validity of excuses in such contexts. My thesis is that determining whether someone is blameworthy depends on the *causal history* of the respective action. Notably, while actions performed automatically are not driven by present deliberation, their causal history may include past deliberations, actions, and beliefs, which can play a crucial role. Regarding the given example, I want to emphasize that agents like Paul may still be responsible for their words, even if they do not genuinely mean them, depending on the causal history of the act of speech.

1. Introduction: Situating the Problem

In life, we generally distinguish between “doing something and having something happen to one” (Mayr 2011: 1. Originally, this distinction appears in Frankfurt 1978: 157). Yet, there are moments when we find ourselves doing something we did not consciously intend. For the purposes of this discussion, let us stipulate that many of us are familiar with the embarrassing feeling that follows when words slip out unintentionally. I’m addressing situations in which we may have said too much—perhaps by revealing a secret or saying something that is (morally) wrong. In such situations, we experience a lack of control over our actions. These spontaneous acts of speech represent actions that lie between actively doing something and merely having something happen to us, demanding careful consideration when it comes to moral responsibility.

Let us consider the following case:

Paul is a kind and fair individual. He identifies as a feminist, firmly believing in the equal value of women, men, and non-binary individuals. In his work, he treats his male and female colleagues with equal respect and does not consider his female colleagues to be less competent. Yet, after

a long, stressful day and a major setback in a project involving only female colleagues, Paul becomes overwhelmed with frustration. In the heat of the moment, he mutters, “Women, always the same”. Immediately, he regrets his words, realizing they caused offense among his colleagues. He insists he didn’t mean them and feels confused about where the statement came from.

This raises an important question: Can someone be held morally responsible for such acts of speech?

In this contribution, I set aside inquiries into implicit bias and assume that Paul does not exhibit any. However, although Paul did not intend to insult anyone and felt a lack of control over his words, I argue that he remains morally responsible for the utterance. Furthermore, I intend to demonstrate that determining Paul’s moral responsibility requires addressing complex issues. Many of the actions we perform daily occur automatically and without deliberate attention (see, e.g., Lumer 2017, Bargh; Chartrand 1999, Custers; Aarts 2010, Schlosser 2019). This is especially true for speech. Most of us speak fluently without consciously planning every word, particularly when proficient in a language.

While certainly not all speech qualifies as a genuine speech act, even automatic utterances can harm others and carry moral weight. Paul’s words demonstrate how speech performed without conscious attention can still have significant consequences.

2. Unintentional Actions and Moral Responsibility

When we perform actions unintentionally, questions of responsibility arise. Whether it’s absentmindedly placing your keys in the fridge while distracted, unintentionally bumping into someone while stretching, or, perhaps more seriously, saying something hurtful in the heat of the moment that you would not have uttered had you been paying closer attention—these are actions, not mere events. Someone must bear responsibility for these actions, though the degree of blameworthiness may differ from that for deliberate actions.

Let us stipulate that someone deliberately shouts, “The police suck!”. In this case, they have chosen the words deliberately, and their utterance reflects a conscious, intentional decision to express a particular sentiment. Moreover,

the person is likely to be aware of the moral and social implications of their words—they know that such a statement could offend or provoke others. By contrast, Paul’s statement was not a product of deliberate thought; it was an unintentional utterance that escaped his lips without his conscious involvement. Paul himself was taken aback by the words he had just spoken. Upon realizing what he had said, he immediately felt regret and distress. He did not mean to express such a thought, nor did he intend to endorse the harmful stereotype that the phrase implied. In fact, the very nature of his reaction—surprise and remorse—demonstrates that he did not consciously intentionally convey the message. His words slipped out automatically rather than deliberately.

When an action appears to be unintentional, it is tempting to reduce or even eliminate the degree of responsibility attributed to the individual. However, the situation is often more nuanced, as even unintentional actions can still carry significant moral weight, depending on their consequences and the individual’s capacity to reflect on and learn from them. In light of this, it is also crucial to consider the role of excuses in such cases, as they can significantly influence how responsibility is assigned when actions occur without conscious intention.

3. The Role of Excuses

In situations like Paul’s, we frequently hear excuses. People may argue they weren’t paying attention or didn’t fully control their actions. These excuses influence how we evaluate moral responsibility. For example, someone who intentionally insults their audience is typically held more accountable than someone whose words slip out unintentionally. Similarly, we distinguish between deliberate and accidental physical harm: intentionally punching someone is judged more harshly than accidentally bumping into them. It is also worth noting that in Paul’s case, we might have to consider different excuses if he had used a slur instead of the phrase “Women, always the same.” While philosophical examinations of slurs generally classify a slur as a bad word (see, e.g., Sosa 2018: 1), Paul’s statement does not include such a bad word per se. His statement heavily relies on context, which leaves room for debate. In a different context, “Women, always the same” could be interpreted positively, suggesting that women are consistently great.

In the following, I draw on the extensive research of Paulina Sliwa regarding excuses. She begins by showing that excuses are commonplace in our daily practice of holding each other accountable for our actions, and she claims that excuses do not absolve agents of their moral responsibility. Here is her argument:

I argue that excuses do not negate but modify an agent's moral responsibility. They bear on the way in which an agent is responsible for what she has done. [...] My argument that excuses leave moral responsibility intact goes as follows: excuses leave moral residue; they leave something for the agent to apologize, to make amends for, to feel distress about. The best explanation for why excuses leave such moral residue is that the agent is morally responsible for the wrong done. (Sliwa 2019: 64)

In Sliwa's view, excuses importantly serve to acknowledge the wrongdoing of an action. I stipulate that we all share the intuition that Paul is guilty of some wrongdoing for making that remark. Nevertheless, Paul will likely feel the urge to explain himself. His excuse might be based on the fact that he did not intend to make the remark. Performing an action unintentionally or without conscious intention can serve as an excuse. Sliwa points out that when we say we did something unintentional, we do not simply want to say that the respective action "reflect[s] the lack of a morally adequate intention. Rather, something went wrong in putting my morally adequate intention into action: I slipped, tripped, lost my balance, or I was ignorant about some crucial feature of the situation" (Sliwa 2019: 49).

However, this analysis only applies to *deliberate* actions. To illustrate this, consider an example discussed by Sliwa: When I deliberately put what I believe to be sugar in my friend's tea, motivated by the morally adequate intention of sweetening their beverage because I know they like their tea sweet, I might unknowingly use a white substance labeled 'sugar' that is actually poison. In this case, my action of poisoning my friend can be excused in the sense that I am not morally blameworthy for it. However, in our example, Paul did not slip or trip in this sense; his utterance is not caused by

any current deliberation. My objective in this contribution is to analyze whether an automatic action, such as Paul's utterance, is morally blameworthy or not.

Although not paying attention to an action can, in principle, serve as an excuse, this does not necessarily absolve someone of all responsibility. When an action appears automatic or not consciously intentional, we must consider whether the agent's past decisions, habits, or beliefs contributed to the event. This now brings me to my central thesis: that in the context of moral responsibility, the causal histories of actions play a significant role.

4. The Importance of Causal History

My argument is that determining moral responsibility for automatic actions, such as Paul's utterance, depends on the causal history of the action (see, e.g., Sartorio 2016 for a substantial discussion of causal histories in the context of free will). By 'automatic', I do not imply that these actions are necessarily performed repeatedly. It is possible that Paul made this statement for the first time, yet he expresses it automatically. By 'automatic', I only imply that the action is performed without consciously attending to a decision and the action itself. I emphasize that, while automatic actions are not directly caused by current deliberations, they often stem from prior decisions, habits, and beliefs.

We can distinguish several scenarios to illustrate this point:

Paul* (the sexist): In this version of the case, Paul holds sexist beliefs, and these beliefs are part of the causal history of his utterance. While he may not have intended to say those words at that moment because he did not want his colleagues to know about his sexist attitudes, they nonetheless reflect his underlying beliefs. Paul* is clearly morally blameworthy, as his utterance stems from his sexist worldview.

Paul's past mindset: Paul does not currently hold sexist beliefs but did so in the past. Years ago, he often used the phrase "Women, always the same" when upset. Although he no longer endorses such views, his prior habits influence his automatic utterance. In this case, Paul bears some responsibility for failing to thoroughly unlearn harmful patterns of speech.

External influence: Paul has never held sexist beliefs, but his utterance echoes something his father used to say in similar contexts. While this reduces Paul's responsibility, the causal history still connects the utterance to external influences that shaped his habits.

The puzzling case: Paul has never held sexist beliefs, never used the phrase before, and cannot trace the utterance to any particular prior experience or influence. The words simply 'slipped out', raising questions about whether actions can ever be fully disconnected from an agent's history.

5. Analyzing Responsibility in Automatic Actions

Causal history plays a critical role in evaluating moral responsibility. In cases like Paul's, even seemingly automatic utterances are often connected to prior deliberations or established habits. Paul's current claim that he acted unintentionally, even if completely honest, does not fully absolve him of responsibility. While he may not have consciously intended to utter the words, his failure to address past habits or influences may have contributed to the incident.

Automatic speech and actions reveal how past choices and patterns can resurface unexpectedly, especially in moments of stress or frustration. These moments challenge traditional notions of responsibility, prompting us to consider not only the immediate context but also the deeper histories that shape our actions.

Paul's case highlights the moral complexities of automatic speech. While he did not consciously choose to say, "Women, always the same", the utterance still caused harm and cannot be dismissed as mere accident. By examining the causal history of such actions, we can better understand the extent to which individuals are responsible for their automatic behaviors.

Ultimately, Paul's responsibility lies in his ongoing efforts to confront and correct past habits, biases, or influences that might resurface in his actions. Even when words slip out unintentionally, they are rarely disconnected from the choices, beliefs, and patterns that define us over time.

6. Conclusion

In this contribution, I have examined how an agent can be morally responsible for an action, even when that action does not seem entirely under their control. By emphasizing that many of our actions are performed automatically rather than deliberately, I argue that acts of speech, in particular, are prone to unintentionally causing harm when executed automatically. In such cases, we frequently encounter familiar patterns of excuses. Drawing on Paulina Sliwa's extensive work on excuses, I contend that while excuses may mitigate the perception of wrongdoing, they do not absolve an agent of moral responsibility. This is evident in Paul's case: although he can offer an excuse for his behavior, he remains morally accountable for it.

Furthermore, I argue that when an agent is aware of their susceptibility to certain language habits, they bear the responsibility to exercise caution in contexts where self-control may falter. While Sliwa's analysis primarily addresses deliberate actions, my contribution extends this discussion to actions performed automatically. I propose that moral responsibility hinges on the causal history of an action, even when the agent's control over the action is limited at the moment of its performance.

References

- Bargh, John A.; Chartrand, Tanya L. (1999): "The Unbearable Automaticity of Being", *American Psychologist*, 54 (7): 462-479.
- Custers, Ruud; Aarts, Henk (2010): "The Unconscious Will: How the Pursuit of Goals Operates Outside of Conscious Awareness", *Science*, 329 (5987): 47-50.
- Frankfurt, Harry G. (1978): "The Problem of Action", *American Philosophical Quarterly*, 15 (2): 157-162.
- Mayr, Erasmus (2011): *Understanding Human Agency*. OUP.
- Lumer, Christoph (2017): "Automatic actions: Agency, Intentionality, and Responsibility", *Philosophical Psychology*, 30 (5): 616-644.
- Sartorio, Carolina (2016): *Causation and Free Will*. OUP.
- Schlosser, Markus (2019): "Agency", in: Zalta, Edward N. (ed.): *The Stanford Encyclopedia of Philosophy*. Winter 2019 Edition. URL: <<https://plato.stanford.edu/archives/win2019/entries/agency/>>.

Sliwa, Paulina (2019): "The Power of Excuses", *Philosophy and Public Affairs*, 47 (1): 37-71.

Sosa, David (2018): *Bad Words. Philosophical Perspectives on Slurs*. OUP.

A (quasi)-Mathematical Theory of Public Language

Giovanni Galli (Teramo, Italy)

Niccolò Covoni (Urbino, Italy/Lugano, Switzerland)

Abstract

This paper proposes an interdisciplinary framework that bridges Claude Shannon's information theory with Ludwig Wittgenstein's work on language. By interpreting entropy as a measure of ignorance within communication systems, we offer a novel perspective on the longstanding debate surrounding the nature of public and private language. Shannon's formalization of information transmission, rooted in entropy and probabilistic structures, is employed to model linguistic exchanges. In particular, we argue that public language, as conceptualized by Wittgenstein, functions as a system that minimizes entropy through shared rules, communal validation, and socially grounded language games. Conversely, Wittgenstein's critique of private language—exemplified by his diary case and beetle-in-the-box thought experiments—is reinterpreted through the lens of information theory: a private language operates at maximum entropy due to the absence of external criteria and mutual intelligibility, making consistent communication unattainable. The paper suggests that Shannon's concept of mutual information parallels Wittgenstein's notion of family resemblance, as both describe the reduction of uncertainty through structural interdependence between sender and receiver. Ultimately, this (quasi)-mathematical approach provides a fresh lens to view Wittgenstein's assertion that language is inherently public and opens future avenues for formalizing communicative noise through entropy-based models in diverse linguistic and social contexts.

1. Introduction

In the field of communication studies, language has traditionally been seen as a complex system for the exchange of meanings. Claude Shannon's seminal work on information theory, which provides a mathematical framework for analyzing information transmission across diverse systems, is of particular interest. A key concept in Shannon's theory is entropy, which measures the uncertainty or ignorance about the state of a system. This concept has been widely applied in fields such as telecommunications and computer science, serving as a powerful tool for measuring the flow and loss of information (Shannon 1948). However, its relevance extends beyond these areas, offering a fresh perspective for understanding the dynamics of human language (Wittgenstein 2009). This paper explores the intersection of Shannon's information theory with Ludwig Wittgenstein's philosophical inquiry on language, particularly his distinction between public and private language. Wittgenstein's work in *Philosophical Investigations* challenges the notion of a purely private language, asserting that meaning is inherently linked to public,

shared practices. The current study suggests, in section 2, that entropy, as a measure of ignorance, can act as a conceptual link between the structure of communication in public language and the constraints of private language. By viewing language as a system of signals subject to entropy, public language can be seen as a means of reducing uncertainty through shared rules and communal validation, as described in section 3. In contrast, private language, lacking such external references, would be subject to maximum entropy, making communication impossible. This interdisciplinary framework is used to explore the role of entropy in language as both a mechanism for organizing communication and a tool for understanding the boundaries of meaningful discourse. The current study aims to present an original perspective on the fundamental nature of language, communication, and the transmission of meaning by connecting Shannon's mathematical theories with Wittgenstein's stance on language.

2. Entropy as a Measure of Ignorance

The theory of information has its roots in the work of Claude Elwood Shannon. In his paper *A Mathematical Theory of Communication* he formulates a definition of information. The author states: "The fundamental problem of communication is that of reproducing at one point either exactly or approximately a message selected at another point" (Shannon 1948: 379). For Shannon, how information is transmitted through messages is not linear: information is, from his point of view, a search in a set of possible good messages. In his view, the core structure of a communication system has a specific structure with some crucial elements:

- The source is what produces the message.
- The transmitter modifies the message into a signal, ready to be transmitted.
- The channel is the medium that conveys the message.
- The receiver encodes (transforms) the message received from the transmitter.
- The destination is the recipient of the message.

With this simple structure that makes up a communication channel, we can build different models, not only those related to computer devices or telephones, but it is also possible to model communication systems made up of

people, objects and very different kinds of actors. In Shannon's approach, the measure of information is formalized by considering the number of possible outcomes of the communication and evaluating their probabilities using the normalization of the probability distribution (the sum of all the probabilities must be one).

This approach is very similar to our understanding of measurement, where we tend to weigh things by comparison, and the result that we obtained is a map of the communication processes through discrete values. Shannon's interpretation gives us an equation that is the measure of ignorance about a system, the Shannon's Entropy:

$$H = - \sum p(x) \log_2 p(x)$$

The mathematician formalized this concept to describe the knowledge of a system from the perspective of the "destination". An important part of the equation is the logarithm and the k constant. Indeed, they are used to describe information in bits. What is obtained here is a quantized definition of information, which is helpful to apply to classical digital systems, where the unit information is represented by the bit.

The function has two crucial points: at its minimum, it represents a highly predictable set of data, where the "destination" predicts future transmission with some confidence, like, for example, what happened for twins that can often complete sentences for one another. Instead, "the equilibrium state of a system maximizes the entropy because we have lost all information about the initial condition except for the conserved quantities; maximizing the entropy maximizes our ignorance about the details of the system" (Sethna 2023: 109).

Shannon's entropy has three key properties that need a "natural" description of ignorance. Moreover, it is the unique (continuous) function to satisfy them. (i) Entropy is maximum for equal probabilities, (ii) entropy is unaffected by extra states of zero probability, and (iii) entropy changes for conditional probabilities.

These three conditions define information as a measure of ignorance, (i) explains us the content of Shannon's equation. It says that entropy is a measure of ignorance through the degree of uncertainty of a probability

distribution. It can also be seen from the form of the function indeed the fact that Shannon's entropy is a concave imposed that the average value of the function over a set of points is less than or equal to the function evaluated at the average of the points. These considerations tell us that, in a certain sense, it is impossible to extract information from the system because all the possibilities have an equal likelihood. (ii) is a boundary of all the possible states of the system taken into consideration, and the property clearly says that impossible events do not give us information. Once again, it is easy to see this from a mathematical point of view because $p(x)$ goes to zero and does not sum with the rest. The last property (iii) claims that the definition of ignorance requires a concept used in statistics: conditional probability.

"A conditional probability is a probability of some event, conditional on another event occurring. The conditional probability of p , conditional on q , is written $\Pr(p \mid q)$, which may be read as the probability of p , given q . These are related to unconditional probabilities by $\Pr(p \& q) = \Pr(p \mid q) \Pr(q)$ " (Myrvold 2021: 242).

Taking this concept into consideration, what is obtained is that the ignorance of a system is related to correlated events. The amount of knowledge is given by the amount of ignorance remaining given the knowledge of another event:

$$\langle S_i(A \mid B) \rangle = S_i(AB) - S_i(B).$$

The equation explains that taking two events, the conditional entropy of an event given another one is given by the joint ignorance of both events minus the original ignorance. The conditional property is satisfied by Shannon's entropy. Indeed, the equation $\langle S_i(A \mid B) \rangle$ could be rewritten as the probability distribution of an element given another one, and the result is the Shannon entropy equation.

So, what is the picture of this property, and what is the definition of entropy? Information is the average amount of information generated by a source and received at the destination. This amount is zero if the two systems are entirely independent; otherwise, it is at maximum when the dependence between the two systems is total. What changes the value of entropy is "the equivocation

(E): the average amount of information generated at source but not received at destination. The mutual information can be computed as: $H(S; D) = H - E$ (Lombardi & al. 2014: 1986).

The Shannon entropy represents the average uncertainty level about possible outcomes. Using this definition and the common structure (in Shannon's view) of each communication system, information processes can be considered, such as a message successfully delivered from a transmitter to a receiver, with a correlation of data, acquired and missed, in the form of entropy. Shannon's overview of the analysis scratches only the surface of information because "the mathematical theory of communication deals with the transmission of information, not the information itself" (Weaver 1949: 12). But this is not a sloppy result; what is obtained is a formal tool to deal with processes where it is possible to recognize the 'common structure' made by Shannon that helps to clarify what happens when communication is involved, like in the case of languages.

3. Public and Private Language

Wittgenstein's exploration of language in the *Philosophical Investigations* invites us to reconsider the very nature of meaning and communication by contrasting what we might call public language with its purported private counterpart, a contrast that becomes especially vivid when one considers the role of entropy in a system of signs. As has been already recognized, the private language argument (PLA) is one of Wittgenstein's most debated notions (Fogelin 1976: 153). We consider PLA as the demarcation of the communicability of meanings in public language. Language is a public phenomenon: "From the very nature of a language itself, we find language means only public language" (Panda and Nath 2020: 1940). Public language, according to Wittgenstein's language game concept, is inherently a communal activity in which words acquire meaning through shared use, mutual correction, and the observable practices of a community. In these language games, meaning is not fixed by some abstract, inner essence but emerges from the practical, rule-governed activities that define our interactions. The public language is characterized by a network of conventions that are open to all; its rules are externally visible and subject to testing and revision, thereby ensuring that any word used in a conversation can be understood, critiqued, and refined through collective agreement or disagreement. For instance, when

we use the word "game" in everyday discourse, its meaning is not derived from a private mental image but from a host of public practices and activities that provide a context for understanding what "game" entails. This system of shared practices effectively reduces the entropy of language by imposing a structure that all participants can recognize and follow, thus allowing for a stable, coherent medium of communication.

In stark contrast stands the notion of a private language, as exposed in the PLA, a language that exists solely within the confines of an individual's inner experiences. Wittgenstein challenges this idea through thought experiments such as the diary case (PI 2009: §258) and the beetle-in-the-box scenario (PI 2009: §293). Imagine a diary written in a private code, where each symbol is supposed to correspond to an inner sensation known only to the writer. Without any external criteria for correctness or shared standards for interpretation, such a diary would be like a system at maximum entropy, where each entry drifts into disorder because there is no common framework to check, validate, or even recognize the meaning of those signs. In this private language, every internal sensation could be as fleeting and variable as the next, leading to a state where no stable, consistent mapping between signs and experiences could be maintained. The beetle-in-the-box scenario further dramatizes this point by positing that if everyone claimed to have a "beetle" in a private, inaccessible inner space, then the word "beetle" would lose all substantive meaning. Since no one could ever inspect another's beetle, there would be no way to ensure that everyone's use of "beetle" referred to the same kind of thing, rendering the word entirely arbitrary and communication effectively impossible.

The argument that the maximum of entropy corresponds to the private language thus hinges on the absence of external, public criteria. In a well-regulated public language, entropy is minimized because words are continually aligned with shared practices and the mutual expectations of a linguistic community. Here, the regularity and structure of language games act as a kind of constraint, preventing the meanings from drifting into chaos. In a private language, by contrast, there is no such stabilizing mechanism; each individual's internal signs are subject only to their own fluctuating interpretations, and without any external reference point, these interpretations can diverge wildly over time, ultimately reaching a state of

maximum entropy where communication, the very purpose of language, is no longer feasible.

Wittgenstein's insistence on the public nature of language is not merely a methodological preference but a profound insight into how human understanding is grounded in social interaction. The language game concept shows that language belongs to a form of life, embedded in the practices and forms of behaviour that constitute our social coexistence. Every utterance is part of a larger context in which the rules of the game are negotiated and enforced, thereby creating a stable arena for the exchange of ideas. This communal aspect is what renders language capable of conveying not just information, but also subtle nuances of meaning that arise from shared experiences and cultural backgrounds. Public language thrives on the common ground that individuals build together, ensuring that every term is subject to public scrutiny, correction, and refinement. This is why the public dimension of language is seen as essential: without it, the possibility of genuine understanding evaporates, leaving behind a kind of linguistic solipsism where each individual is trapped in a private universe of meanings that no one else can access or verify.

The diary case and the beetle-in-the-box scenario serve as powerful illustrations of the inherent difficulties in conceiving of a truly private language. In the diary case, the idea of recording one's inner sensations in a private code illustrates that without a common external channel of communication, it is impossible to determine whether the same sign consistently corresponds to the same inner experience over time. The variability and subjectivity of internal sensations mean that any attempt to codify them privately would be doomed to instability, as there would be no reliable standard to ensure consistency. Similarly, the beetle-in-the-box scenario underlines the futility of speaking about private inner objects when there is no way to compare them across different individuals. Even if everyone were to claim that a particular sensation or inner object is a "beetle", the absence of any public criteria for what a beetle actually is would mean that the term "beetle" is devoid of any concrete referent. In both cases, the lack of public verification leads to a kind of maximum entropy, a state where the variability of internal signs is so high that no stable meaning can be maintained, thus undermining the very possibility of communication.

By insisting that language must be public, Wittgenstein effectively challenges the notion that our inner experiences can ever be fully captured or expressed in a self-contained, private system. The maximum entropy associated with a private language is not merely a metaphor for disorder; it is a reflection of the fundamental impossibility of establishing a reliable correspondence between private sensations symbols and the public symbols we use to communicate. Without a shared framework, every word becomes an isolated island, disconnected from the collective fabric of meaning that is woven through our everyday interactions. Communication, in this sense, is not a matter of transmitting pre-existing private contents from one mind to another, but of engaging in a public process of negotiation, correction, and mutual understanding. It is precisely this process that imbues language with its power and resilience, ensuring that despite the inherent variability of individual experiences, we can nonetheless arrive at a common ground of understanding.

In summary, Wittgenstein's *Philosophical Investigations* compels us to recognize that the notion of a private language, a language so idiosyncratic and isolated that it verges on a state of maximum entropy, is antithetical to the very possibility of communication. Public language, with its language games and shared criteria, provides the structure necessary to reduce entropy and maintain a coherent system of meaning. Through the diary case and the beetle scenario, Wittgenstein vividly illustrates that without public benchmarks and communal practices, language would devolve into an unmanageable chaos of individual, unverified signs, rendering meaningful communication impossible. This insight underscores the idea that language is not a private possession but a public resource, essential for the intersubjective exchange that lies at the heart of human life.

4. Conclusions

In this paper, we have introduced a way to connect the mathematical formulation of information through the use of the concept of Shannon entropy and the analysis of public and private languages made by Wittgenstein in PI. By focusing on the concept of mutual information and the interpretation of entropy as a measure of ignorance, the picture proposed that private language could exist within the limits of the entropy function when the two systems are entirely independent. In particular, this case is represented by the lack of a shared framework between the source and the destination. In all the other

cases we are "playing" in the ground of the public language, all possible values proposed by the formulation of entropy represent cases where the social interactions and the numbers of standard rules are presented, the value of $H(S;D)$ represents the family resemblance that is present between the systems of the communication.

More broadly, Shannon's information turns out to be an interpretative key to understanding the structure and function of language; this sort of theorisation of a formalised language does not go against the Wittgenstein formulation of the language as a language like fundamentally a public phenomenon, rooted in social practices and regulated by shared criteria that guarantee its comprehensibility and reliability. Moreover, the theory proposed by Shannon provides a mathematical framework to describe this dynamic: public language reduces entropy through the establishment of rules and the collective verification of meaning, while the lack of external references in a private language leads to increased uncertainty and a loss of information. A future area of work is the possibility of formulating an algorithm capable of calculating the noise through the common rules present in various communication settings; this could be done by reinterpreting the notion of relative information where the physical restriction in the number of states the two systems together can be as the possible common rules, in this way what can be obtained is a quantity "which quantifies how much each system is constrained by the other" (Rovelli 2024: 34).

References

- Fogelin, Robert J. (1976) *Wittgenstein*, London and Boston: Routledge.
- Lombardi, Olimpia; Holik, Federico; Vanni, Leonardo (2016) "What is Shannon information?", *Synthese* 193, 7, 1983-2012.
- Myrvold, Wayne C. (2021) *Beyond Chance and Credence: A Theory of Hybrid Probabilities*, Oxford: Oxford University Press.
- Panda, Mamata M., Nath, Rajakishore (2020) "Wittgenstein on Public Language About Personal Experiences", *Philosophia* 48, 1939-1960.
- Rovelli, Carlo (2024) "Princeton seminars on physics and philosophy", <https://arxiv.org/abs/2407.01989>

Shannon, Claude E. (1948) "A mathematical theory of communication", *The Bell System Technical Journal* 27, 4, 379-423.

Sethna, James P. (2023) *Statistical Mechanics: Entropy, Order Parameters, and Complexity*, Oxford: Oxford University Press.

Weaver, Warren (1949) "The Mathematics of Communication", *Scientific American* 181, 1, 11-15.

Wittgenstein, Ludwig (2009), *Philosophical Investigations*, Oxford: Blackwell Publishing.

From Meaning-Facts to Form of Life: Dissolving Kripke's Puzzle with Wittgenstein

Ruosen Gao (Beijing, China)

Abstract

This paper argues that Kripke's skeptical paradox about rule-following is rooted in a mistaken metaphysical demand: the search for meaning-facts that justify linguistic practices. Drawing on Wittgenstein's later philosophy, I show that the demand for such justification misconstrues the nature of normativity. Where Kripke sees a lack of grounding, Wittgenstein sees the end of justification—in our shared form of life. By contrasting Kripke's metaphysical question with Wittgenstein's therapeutic response, I propose that rule-following should be understood not through the discovery of hidden facts, but through our ordinary practices. Wittgenstein does not solve the paradox; he dissolves it.

1. Introduction

Kripke (1982), in his *Wittgenstein on Rules and Private Language*, presents a skeptical challenge to explain how linguistic expressions can have determinate meaning. His famous “quus” example leads to the conclusion that nothing in a speaker's prior use of “+” determines that they meant addition rather than a deviant function. He suggests that in the *Philosophical Investigations*, Wittgenstein abandons the truth-conditional theory of meaning advocated in the *Tractatus*, replacing it with an account in terms of assertability conditions. Kripke's Wittgenstein (hereafter KW) goes further, arguing that rule-following and meaning only make sense relative to community practices, rendering solitary rule-following impossible.

KW uses this thesis to claim that there are no facts—whether reductionist or not—that could constitute a speaker's meaning. This has motivated attempts to restore a factualist account. Some theorists propose reductionist strategies (e.g., Horwich 1990; 2012), grounding meaning in use or dispositions. Others resist reduction, defending normative accounts of rule-following (e.g., Wright 1984; McDowell 1984). The core question is whether facts about meaning can carry the normative force needed to guide judgments.

In this paper, I argue that Kripke and Wittgenstein address fundamentally different problems. Kripke's puzzle presupposes that rule-following requires metaphysical grounding: that an agent's action must be justified by an antecedent understanding of the rule's meaning. Accordingly, his solution

must identify a meaning-fact that constrains use. But Wittgenstein's approach is epistemological and therapeutic. He shows that the demand for a meaning-fact misunderstands how normativity works in language. It is not facts that justify our going on, but reasons—and those reasons come to an end.

2. Ginsborg's Solution to Kripke

Kripke describes the skeptical challenge as “taking two forms”:

“First, [the skeptic] questions whether there is any fact that I meant plus, not quus, that will answer his sceptical challenge. Second, he questions whether I have any reason to be so confident that now I should answer ‘125’ rather than ‘5.’ The two forms of the challenge are related. I am confident that I should answer ‘125’ because I am confident that this answer also accords with what I meant.” (Kripke 1982: 11)

He stipulates that the second challenge rests on the first, leading to four steps in the skeptical reasoning:

- A. There is a specific rule in play.
- B. The agent understands the meaning of that rule.
- C. This understanding is based on a fact that determines the rule's meaning.
- D. The agent acts in accordance with that meaning, and the fact justifies the action.

Step B is crucial: Kripke insists that one cannot follow a rule without grasping its meaning. Hence, rule-following is not merely behavioral; it requires justification grounded in meaning-facts. Steps A, B, and D reflect the epistemological aspect of the paradox, while B and C form its metaphysical core.

In response, Ginsborg (2011b) proposes a “partially reductive” account of meaning. She writes:

“While my approach is partly reductionist, in that it aims to reduce facts about meaning to facts that are in a sense more primitive, it does not attempt a reduction of meaning to facts conceived purely naturalistically.” (Ginsborg 2011b: 230)

Her key move is to introduce a notion of primitive normativity—a normative sense of how one ought to proceed, even in the absence of prior understanding. For Ginsborg, we ought to say “125” without needing to assume that we, or anyone else, meant “plus” previously (2011b: 231). What underlies this ought is not grasp of a rule, but a kind of felt appropriateness in the act itself.

But this is precisely where her proposal fails to meet Kripke's challenge. If the agent does not understand the rule, how can the act count as rule-following? Ginsborg calls this special “ought” a primitive ought, grounded in a normative attitude that precedes conceptual content. However, her account blurs the very distinction that Kripke's paradox brings into relief: between following a rule and merely continuing a pattern.

To clarify this point, I propose distinguishing two kinds of action:

Rule-following action: guided by an explicit rule and requiring conceptual grasp.

Going-on action: guided by past usage or habit, with only an implicit sense of how to proceed.

This distinction is not just terminological. It tracks two very different kinds of normativity: one requiring justification (Kripke's target), and the other embedded in our training and behavior (Wittgenstein's focus). Their difference mirrors that between knowing-that and knowing-how, or between justification and natural responsiveness.

Finally, all explicit rules ultimately rest on something more basic—what Wittgenstein calls our form of life. Whether that is best conceived as natural or normative is a further question, but it is clear that trying to ground rule-following in “primitive normativity” without an account of meaning leaves Ginsborg unable to satisfy the normative constraint Kripke demands.

3. Kripke's Skepticism and Wittgenstein's Going-On

Kripke's challenge has both epistemological and metaphysical dimensions. On the epistemic side, he asks how we can know the correct application of a rule. On the metaphysical side, he demands a fact—independent of our behavior or mental states—that constitutes what we meant. For Kripke, these two

challenges are linked: if any fact counts as your meaning addition by "+", it must also justify your confidence in answering "125" rather than "5."

Ginsborg tries to decouple these dimensions. She maintains that one can follow a pattern in line with past usage, even without grasping its meaning (Ginsborg 2011b: 232). But this elides a crucial distinction. Conforming to past usage is not the same as conforming to past meaning. Meaning involves a rule that constrains future applications, and that requires understanding.

Return to Kripke's example. Suppose someone's use of "+" coincides with the plus function up to 56, but afterward they say "5" instead of "125." The quus function fits this pattern just as well. If both rules are extensionally identical up to a point, no behavior can distinguish which one is being followed. The only way to break the symmetry is by invoking what the speaker meant—but that is precisely what's under dispute. Kripke concludes that no natural fact can do this normative work (Kripke 1982: 69–71).

So Ginsborg's appeal to primitive normativity—without reference to meaning—falls short. The "ought" she invokes cannot justify an action as correct unless it is grounded in the agent's understanding of the rule. Normativity without contrast between right and wrong reduces to mere inclination.

Wittgenstein, however, approaches the issue differently. His examples rarely depend on explicit grasp of rules. Instead, he shows how rule-following arises within broader patterns of training and participation in practices. "I shall also call the whole, consisting of language and the actions into which it is woven, the 'language-game'" (PI 2009: §7).

Consider a numerical series: "2, 4, 6, 8, ..." There are many ways to continue it:

"2, 4, 6, 8, 10, 12, ..." (standard +2 rule)

"2, 4, 6, 8, 2, 4, ..." (cyclical repetition)

"2, 4, 6, 8, 8, 8, ..." (stagnant repetition)

Each continuation fits a possible rule. Before a rule is stated, any of them might seem correct. In such cases, the learner does not proceed from an articulated rule but from a practice. As Wittgenstein writes:

“Now we get the pupil to continue a series, say ‘+2’, and he writes 1000, 1004, 1008, 1012. We say to him ‘Look what you’re doing!’ He doesn’t understand. We say: ‘You should have [sollen] added two: look how you began the series!’ He answers: ‘Yes, isn’t it right? I thought that was how I should [sollen] do it.’” (PI 2009: §185)

Here, the pupil’s behavior appears reasonable from his own perspective. But once we have said, “Think what ‘+2’ means!”, further justification gives out. The only way forward is to return to examples—to training. We hit what Wittgenstein calls bedrock.

This notion of “going on” does not invoke meaning-facts. In *Zettel*, Wittgenstein remarks:

“You must remember that there may be such a language-game as ‘continuing a series of digits’ in which no rule, no expression of a rule is ever given, but learning happens only through examples.” (Z 1981: §295)

In such language-games, there is no hidden fact that determines the next step. The idea that each step requires justification by an inner rule or grasped pattern is alien to these forms of life. The practice itself is the rule.

Thus, Wittgenstein and Kripke are responding to different demands. Kripke asks for causes, for metaphysical facts. Wittgenstein redirects our gaze to reasons, which operate within practices—and which come to an end. Their disagreement is not merely about the nature of normativity, but about what counts as a philosophical problem in the first place.

4. Wittgenstein’s therapeutic

Wittgenstein does not share Kripke’s concern with whether “meaning-facts” exist. He raises no skeptical paradox of the Kripkean sort. His later philosophy is not a theory of meaning but a redirection of our attention—from metaphysical demands to everyday practices.

While the term “disposition” appears only once in the *Philosophical Investigations* (PI 2009: §149), the notion plays a background role throughout Wittgenstein’s reflections. A child trained in arithmetic is disposed to continue “2, 4, 6...” with “8,” not “7.” This disposition is not what grounds meaning, but reflects how rule-following is inculcated. In everyday language-games, the

problems of infinitude, error, and normativity do not arise. These problems appear only when philosophers ask what makes such practices correct—as if some fact must underwrite them.

Wittgenstein diagnoses this urge to locate meaning in a metaphysical object. In the *Blue Book*, he writes:

“We are looking for the use of a sign, but we look for it as though it were an object co-existing with the sign... we are looking for a ‘thing corresponding to a substantive’.” (BB 1958: 5)

Meaning, he insists, is not hidden behind use. Nor is it a separate entity that makes our practices legitimate. The demand for such an entity is a philosophical illusion, born of grammatical confusion.

His alternative is to show how our practices justify themselves—until they no longer do. In PI §217, he writes:

“How am I able to obey a rule?”—if this is not a question about causes, then it is about the justification for my following the rule in the way I do. If I have exhausted the justifications, I have reached bedrock, and my spade is turned. Then I am inclined to say: ‘This is simply what I do.’” (PI 2009: §217)

This is not a refusal to justify, but a recognition that justification ends within life itself. What Kripke takes as a failure—no further fact to cite—Wittgenstein sees as the natural termination of philosophical inquiry.

He illustrates this with the image of a gesture:

“Such a case would present similarities with one in which a person naturally reacted to the gesture of pointing with the hand by looking in the direction of the line from fingertip to wrist, not from wrist to fingertip.” (PI 2009: §185)

Why do we respond one way and not another? There is no deeper answer than this: it is what we do. This is not to reduce normativity to brute behavior but to embed it in forms of life. We are trained to respond in certain ways, and these shared responses form the background against which any justification makes sense.

Kripke's skeptic asks for a cause—some objective fact that determines meaning and action. But such questions, Wittgenstein shows, presuppose a mistaken picture of how language works. There is no metaphysical anchoring beyond our form of life, and seeking one only leads us astray.

5. Conclusion

Kripke and Wittgenstein pose fundamentally different questions about rule-following. Kripke's challenge is metaphysical: what kind of fact could determine that one means addition rather than some deviant function? He demands an answer that explains how meaning-facts can ground correct applications. This requirement gives rise to the skeptical paradox. In seeking a normative constraint in the form of a fact, Kripke finds none—and concludes that no such facts exist.

Wittgenstein, by contrast, dissolves the puzzle by challenging its very framing. He does not argue that meaning-facts are unknowable or inaccessible; he argues that we do not need them. The demand for metaphysical grounding mistakes how rule-following works. In our ordinary practices, justification comes to an end—not in a foundational fact, but in our form of life.

In Kripke's picture, normativity must be externally grounded—by something beyond behavior, intention, or community. In Wittgenstein's, normativity is internally embedded in our practices. What Kripke sees as a failure to justify, Wittgenstein sees as the natural termination of justification. The paradox vanishes once we see that the request for a meaning-fact is a philosophical illusion.

Thus, the contrast is not merely between two answers, but between two conceptions of what counts as a problem. Kripke constructs a metaphysical demand that cannot be satisfied. Wittgenstein invites us to give up the demand altogether—and to return from theorizing to life.

References

- Boghossian, Paul (2005) "Is Meaning Normative?", in: Beckermann, Ansgar; Nimtz, Christian (eds.) *Philosophy – Science – Scientific Philosophy*, Paderborn: Mentis.
- Ginsborg, Hannah (2011b) "Primitive Normativity and Skepticism about Rules", *The Journal of Philosophy* 108(5), 227–254.

Horwich, Paul (1990) "Wittgenstein and Kripke on the Nature of Meaning", *Mind & Language* 5(2), 105–121.

Horwich, Paul (2012) *Wittgenstein's Metaphilosophy*, Oxford: Oxford University Press.

Kripke, Saul A. (1982) *Wittgenstein on Rules and Private Language*, Cambridge, MA: Harvard University Press.

McDowell, John (1984) "Wittgenstein on Following a Rule", *Synthese* 58, 325–364.

Wittgenstein, Ludwig (1958) *The Blue and Brown Books*, Oxford: Basil Blackwell.

Wittgenstein, Ludwig (1981) *Zettel*, Oxford: Blackwell.

Wittgenstein, Ludwig (2009) *Philosophical Investigations*, Rev. ed., Wiley-Blackwell.

Wright, Crispin (1984) "Kripke's Account of the Argument Against Private Language", *The Journal of Philosophy* 81(12), 759–778.

Acknowledgment as Feminist Praxis: Reframing Skepticism and Empathy through Wittgenstein, Cavell, and Aesthetic Media

Freya Katharina Gerz (Berlin, Germany)

Abstract

This paper bridges feminist philosophy, Ordinary Language Philosophy, and film theory to reframe the classic problem of other minds as an ethical and political crisis of acknowledgment. Drawing on Wittgenstein's *Philosophical Investigations* and Stanley Cavell's ethics of acknowledgment, it challenges traditional epistemology's obsession with certainty and instead foregrounds relational, emotional, and embodied ways of knowing. The core contribution is the introduction of the "Soul Bridge": a feminist-epistemic model of empathy that resists appropriation, embraces vulnerability, and attends to systemic asymmetries in recognition. Combining Cavell's moral vision with Martha Nussbaum's theory of emotions and feminist critiques of abstraction, the Soul Bridge expands acknowledgment into an affective and aesthetic practice—one that extends beyond language into perception, mood, and gesture. Film emerges as a central site of ethical rehearsal: not merely representing but performing acknowledgment. Through cinematic techniques such as rhythm, framing, and silence, film cultivates a mode of seeing that is attuned, tentative, and ethically charged. Ultimately, the paper argues that skepticism is not just a failure of knowledge but a failure of care. The Soul Bridge offers a feminist imperative: to respond, to see, and to build conditions under which genuine acknowledgment is possible—in theory, in art, and in life.

1. Feminism, Skepticism, and the Crisis of Connection

Feminist philosophy has long challenged the abstraction and emotional detachment of canonical epistemologies. Whether through critiques of Cartesian dualism, the revaluation of embodied knowledge, or the insistence on the political stakes of language, feminist thinkers have reframed what counts as knowing and who counts as a knower. Yet one problem persists in subtle but insidious ways: the philosophical skepticism about other minds. Though seemingly abstract, this skepticism reflects a deeper estrangement—a refusal or failure to acknowledge others as subjects. This paper intervenes at the intersection of feminist epistemology and Ordinary Language Philosophy, proposing an expanded conception of acknowledgment as a feminist praxis. Building on Wittgenstein's later philosophy, Stanley Cavell's ethics of acknowledgment, and feminist theories of relationality and emotion, I introduce the "Soul Bridge" as a conceptual model of empathy that foregrounds the moral, affective, and aesthetic dimensions of understanding

others. Film, I argue, is not merely illustrative of this model but constitutive of it: cinema becomes a space where acknowledgment is enacted, negotiated, and felt.

2. Wittgenstein and the Feminist Potential of Ordinary Language Philosophy

Wittgenstein's *Philosophical Investigations* does not provide a solution to skepticism in the traditional sense. Instead, it redirects our attention: away from metaphysical doubt and toward our forms of life, our practices, and our words in context. His famous remarks on understanding others—"If I see someone writhing in pain with evident cause I do not think: all the same, his feelings are hidden from me" (2009:225)—challenge the epistemic framing of other minds. Here, knowledge is not a matter of inference but of participation in shared practices of expression and response. This insight bears deep affinity with feminist critiques of abstraction. Feminist epistemologists such as Miranda Fricker (2007) and Lorraine Code (1991) have emphasized the ethical dimension of knowing, framing it as situated, relational, and responsive. In this light, Wittgenstein's refusal to "solve" skepticism becomes itself a feminist gesture: it resists the closure of rational certainty and insists on the primacy of everyday human responsiveness. Cavell radicalizes this orientation. For him, the problem of other minds is not an epistemological error but a moral failure—a withdrawal from acknowledgment. The very demand for proof of another's inner life, he suggests, already displaces us from the shared forms of life in which responsiveness is possible. Skepticism, then, is not refuted but endured; it becomes a condition of our vulnerability to one another. This shift transforms the grammar of knowledge claims into a grammar of responsibility. As Cavell puts it: "But why is sympathy expressed in this way? Because your suffering makes a claim upon me. It is not enough that I know – am certain – that you suffer; I must do or reveal something (whatever can be done). In a word, I must acknowledge it, otherwise I do not know what '(your or his) being in pain' means." (1976: 263) In this view, acknowledgment is not reducible to epistemic certainty—it is enacted in the moral and expressive responsiveness that renders the other intelligible in the first place.

3. Acknowledgment Beyond Language: Toward a Feminist Recasting

Yet Cavell's notion of acknowledgment, while ethically potent, remains primarily linguistic. He explores acknowledgment through Shakespearean dialogue, through confession, silence, and refusals to speak—but still within a

linguistic paradigm. Feminist theory presses further. It asks: what of the body? What of emotion? What of those whose language is systematically excluded or devalued? Here, I propose a feminist expansion of Cavell's idea: acknowledgment must be reconceived as a multimodal, affective, and aesthetic practice. Drawing on Martha Nussbaum's cognitive theory of emotions (2001)—which treats emotions as value judgments that disclose what matters—I argue that acknowledgment is not only something we say but something we feel, perceive, and perform. Emotions are not merely inner states, but forms of attentiveness to others, ways of being ethically attuned. This expansion resonates with feminist insistence on the epistemic value of affect. As bell hooks reminds us, genuine love is rarely an emotional space where needs are instantly gratified; rather, "to know love we have to invest time and commitment... dreaming that love will save us, solve all our problems or provide a steady state of bliss or security only keeps us stuck in wishful fantasy, undermining the real power of the love – which is to transform us" (2000:114). This knowing is not reducible to rational assent. It is participatory, embodied, and often aesthetic. In this sense, acknowledgment becomes not just a speech act, but an act of ethical imagination—a reaching toward the other through forms that exceed propositional discourse.

4. The Soul Bridge: Empathy as Feminist-Epistemic Practice

To conceptualize this expanded practice of acknowledgment, I introduce the "Soul Bridge"—a term for empathy understood not as passive identification or emotional contagion, but as a dynamic, relational, and transformative practice. The Soul Bridge connects Cavell's notion of acknowledgment with Nussbaum's theory of emotions, and enriches both through feminist philosophy's attentiveness to embodied, affective, and marginalized experience. The Soul Bridge challenges the individualism of traditional accounts. Rather than treating empathy as an inner faculty or skill, it treats it as a shared practice—a crossing over that requires risk, vulnerability, and responsiveness. It is not about fully knowing the other, but about making oneself available to their reality. In this sense, it is feminist in its refusal of mastery, its emphasis on interdependence, and its insistence that knowing is always situated and partial. Crucially, the Soul Bridge is also aesthetic. It operates not only through language and rational argument, but through perception, mood, atmosphere, rhythm, and tone. Here, Wittgenstein's remarks on "seeing-as" become

relevant (2009). Understanding others is not just a matter of evidence but of how we see, what forms of expression we attend to, and what we count as meaningful. Feminist philosophy, especially in its engagements with literature, visual culture, and film, insists on the moral significance of this kind of seeing. From a feminist perspective, the Soul Bridge is not only a model for ethical recognition—it is also a response to the political and epistemic asymmetries that structure who is seen, heard, and understood in our social world. As scholars like Miranda Fricker (2007), Kristie Dotson (2011), and bell hooks (2000) have emphasized, epistemic injustice is not a matter of individual misunderstanding but of systemic exclusion: certain voices are routinely dismissed, misheard, or denied uptake. Empathy, then, must not be conceived as a neutral disposition but as a situated, accountable, and reparative practice. The Soul Bridge foregrounds this by reframing empathy as an effortful form of recognition that acknowledges the power differentials and emotional labour involved in cross-positional understanding. It aligns with feminist critiques of epistemic violence, where knowledge practices reproduce forms of marginalization, and with calls to attend to gendered, racialized, and affectively coded experiences that often remain illegible within dominant epistemic frameworks. In this way, the Soul Bridge becomes not only an epistemological intervention but a feminist practice of acknowledgment, oriented toward justice as much as understanding.

5. Film as Feminist-Epistemic Medium

To explore how the Soul Bridge is enacted, I turn to cinema. Film, unlike philosophy or literature, does not merely represent emotions—it structures them through visual and auditory form. The play of light and shadow, the pacing of scenes, the cadence of silence, the embodied presence of actors—all these shape our perception of others in ways that resist paraphrase. Cavell's own writings on film, especially *The World Viewed* and *Pursuits of Happiness*, provide rich resources for this claim. He treats cinema as a medium of philosophical reflection, one that renders visible our conditions of connection and estrangement. Yet he does not fully connect his film theory to his ethics of acknowledgment. Nevertheless, certain moments in his writing reveal the latent continuity between the two. As he writes, "These two simply appreciate one another more than either of them appreciates anyone else, and they would rather be appreciated by one another more than by anyone else. They just are

at home with one another, whether they can ever live together under the same roof—that is, ever find a roof they can live together under” (1981:167). Here, acknowledgment takes the form of mutual attunement—a shared inhabitation of emotional space that exceeds propositional knowledge.

While Cavell’s reflections remain largely confined to heterosexual romantic narratives, feminist film theory confronts more directly the politics of visibility, exclusion, and power. My proposal is that film does not only depict acknowledgment; it performs it. This performance has feminist stakes. Feminist film theory—from Laura Mulvey’s critique of the male gaze (1975) to bell hooks’ ethics of looking (2000)—has long emphasized how film shapes subjectivity and power. Building on this, I suggest that certain films cultivate not voyeurism but recognition—a practice of attunement that enacts the Soul Bridge.

Films such as *Ikiru* (1952) and *The Double Life of Véronique* (1991) illustrate how cinema can become a space of ethical responsiveness rather than psychological transparency. In *Ikiru*, the terminally ill bureaucrat Watanabe undergoes a quiet, existential transformation, not through didactic dialogue but through gestures, silences, and sequences of repetition. The camera lingers on his face, his solitary movements through the city, the snow falling on a playground he helped create. These images cultivate a space of mourning and moral reflection, inviting the viewer not to solve or interpret his interior state, but to dwell in the ambiguous register of his becoming. Empathy here arises not from narrative certainty but from aesthetic attention—from remaining with his vulnerability without resolving it.

Similarly, *The Double Life of Véronique* constructs a mirrored subjectivity between two women—Weronika in Poland and Véronique in France—whose lives never directly intersect but are bound by uncanny visual, emotional, and musical resonances. The film resists explanatory closure and instead envelops the viewer in a texture of glances, colors, and moods that suggest a shared interiority without asserting it. We do not come to know the characters in the epistemic sense; rather, we are drawn into an aesthetic mode of receptivity, of ‘seeing-with,’ that suspends our desire for mastery and affirms the opacity of the other.

Returning to the notion of aspect perception ("seeing-as") previously introduced, we find in Part II of Wittgenstein's *Philosophical Investigations* a further elaboration on how perceptual shifts occur—not by way of new stimuli, but through a transformation in how the same figure is seen, as exemplified by the duck-rabbit. This transformation is not merely visual; it marks a change in understanding, a reconfiguration of salience and meaning. "I see that it has not changed; and yet I see it differently. I call this experience 'noticing an aspect'" (2009:193), he writes, and more broadly: "My visual impression has changed and now I recognize that it has not only shape and colour but also a quite particular 'organization'" (2009:196). Perception, in this sense, is structured by our forms of life and attuned to contexts of significance.

This idea is particularly productive in relation to cinematic experience. Film conveys meaning not only through narrative or dialogue, but also through formal techniques—framing, rhythm, silence, movement, and light—that shape how we see. The viewer's perception becomes an active, cultivated responsiveness—a shift in what is foregrounded, what counts as expressive, what calls for acknowledgment—so that film does not merely represent acknowledgment, but enacts it. Certain films stage relationality not as mutual comprehension but as attuned co-presence in the face of uncertainty. They invite us to remain with what is opaque, incomplete, or emotionally unresolved—echoing the Wittgensteinian insight that some of the most significant forms of understanding cannot be said, only shown.

In this light, cinema becomes a feminist-philosophical practice: it trains us in how to see, feel, and acknowledge without appropriating or dissolving the other into ourselves. It does not simulate empathy through identification but shapes perception itself—cultivating an ethical orientation grounded not in certainty but in openness, attentiveness, and restraint. These films do not simply depict the Soul Bridge as a concept; they embody its logic through rhythm, silence, and visual composition. They offer a perceptual and emotional education in how to remain with difference and vulnerability without retreating into detachment or control.

When cinema enables empathetic engagement, it does so not by revealing inner states through exposition, but by transforming our mode of perception. This shift occurs because film reconfigures how we attend to others

—not through propositional content, but through formal cues that guide our sensibility. Formal techniques of film orient the viewer toward a particular mode of responsiveness: hesitant, open, attuned. In this way, cinema trains perception as an ethical capacity. Film teaches us to see something as expressing pain, hesitation, or intimacy. Such understanding is not argumentative or propositional; it is gestalt-bound, embodied, and affectively mediated. It is not about reaching a conclusion but about experiencing a shift in how the world is presented to us. In this light, the *Soul Bridge* expands beyond linguistic and emotional dimensions to encompass the visual and perceptual as sites of ethical relation. Empathy, then, becomes a way of seeing—one that is cultivated through aesthetic form and grounded in our shared capacity for aspectual transformation.

6. From Skepticism to Solidarity: Feminist Implications

What, then, are the feminist implications of this reframing? First, it challenges the privileging of epistemic certainty in philosophical accounts of other minds. Feminist philosophy has long argued that such certainty is a masculinist ideal, one that neglects the relational, affective, and embodied dimensions of knowing. The *Soul Bridge*, by contrast, embraces partiality, vulnerability, and mutual recognition as epistemic virtues. Second, it expands the domain of legitimate philosophical tools. Emotions, film, aesthetic experience—these are not mere supplements to "real" knowledge, but central to how we understand and engage with others. This affirms the feminist claim that the boundaries of philosophy must be porous, attentive to lived experience, and open to interdisciplinary resonance. Third, it reframes skepticism not as a failure to know, but as a failure to care—as an ethical breakdown that feminist theory is uniquely positioned to address. When we refuse to acknowledge the other, we do not merely commit a cognitive error—we perpetuate social and emotional forms of injustice. In this sense, the *Soul Bridge* becomes not just a philosophical model but a feminist imperative: to respond, to attend, to build forms of life where acknowledgment is possible. One might reasonably ask whether film is truly a reliable site for ethical acknowledgment. Does cinema risk aestheticizing suffering or simulating empathy without demanding real-world responsibility? Moreover, does the concept of empathy—as framed by the *Soul Bridge*—run the risk of romanticizing mutual understanding, especially across axes of difference marked by structural inequality? Might it

obscure the asymmetries of power that shape whose experiences are legible, whose pain is recognized, and whose voice is granted space? These concerns are not only legitimate but essential. However, the Soul Bridge does not present empathy as immediate transparency or perfect understanding. On the contrary, it emphasizes empathy as a deliberate, situated, and fallible practice, one that must be attuned to the complexities of power and difference. It also reframes film not as a mirror of real life but as a mediated space of ethical rehearsal—a domain in which perceptual and affective habits are formed, challenged, and transformed. In this sense, film does not guarantee recognition—but it can, under the right conditions, help cultivate the dispositions necessary for it.

7. Conclusion: Toward a Feminist Philosophy of Acknowledgment

By integrating Wittgenstein's philosophy of language, Cavell's ethics of acknowledgment, and feminist critiques of abstraction, this paper offers a new perspective on the problem of other minds. Through the Soul Bridge, empathy becomes a shared, aesthetic, and ethical practice—one that resists the solipsism of skepticism and affirms the feminist commitment to relationality and justice. Film, in this account, is not merely an object of analysis. It is a feminist-philosophical ally—a medium through which we practice the kind of acknowledgment that Cavell theorized and that Wittgenstein gestured toward. By enacting the Soul Bridge, cinema does not resolve skepticism but transforms it—into a question not of knowing, but of being with. In a time marked by political polarization, epistemic mistrust, and emotional numbness, the call to acknowledge others is more urgent than ever. Feminist philosophy, Wittgensteinian reflection, and cinematic experience converge to answer this call—not with certainty, but with openness, humility, and care.

References

- Cavell, Stanley (1971) *The World Viewed: Reflections on the Ontology of Film*, Cambridge, MA: Harvard University Press.
- Cavell, Stanley (1976) *Must We Mean What We Say? A Book of Essays*, Cambridge: Cambridge University Press.
- Cavell, Stanley (1979) *The Claim of Reason: Wittgenstein, Skepticism, Morality, and Tragedy*, Oxford: Oxford University Press.

Cavell, Stanley (1981) *Pursuits of Happiness: The Hollywood Comedy of Remarriage*, Cambridge, MA: Harvard University Press.

Code, Lorraine (1991) *What Can She Know? Feminist Theory and the Construction of Knowledge*, Ithaca: Cornell University Press.

Dotson, Kristie (2011) "Tracking Epistemic Violence, Tracking Practices of Silencing", *Hypatia* 26(2), 236–257.

Fricker, Miranda (2007) *Epistemic Injustice: Power and the Ethics of Knowing*, Oxford: Oxford University Press.

Hooks, Bell (2000) *All About Love: New Visions*, New York: William Morrow.

Kieślowski, Krzysztof (1991) *The Double Life of Véronique*. France/Poland: Sidéral Productions / Studio Filmowe TOR.

Kurosawa, Akira (1952) *Ikiru*. Japan: Toho Company Ltd.

Mulvey, Laura (1975) "Visual Pleasure and Narrative Cinema", *Screen* 16(3), 6–18.

Nussbaum, Martha C. (2001) *Upheavals of Thought: The Intelligence of Emotions*, Cambridge: Cambridge University Press.

Wittgenstein, Ludwig (2009) *Philosophical Investigations*, revised 4th edn., trans. G. E. M. Anscombe, P. M. S. Hacker, and Joachim Schulte, Malden, MA: Wiley-Blackwell.

Wittgenstein on concepts and forms

Nicola Gianola (Rome, Italy)

Abstract

This paper aims to analyze the evolution of the distinction between genuine concepts and formal concepts in Wittgenstein's thought. In the first part, I will examine this distinction as it is presented in the *Tractatus*, highlighting the reasons why formal concepts are regarded as pseudo-concepts. I will then show how Wittgenstein gradually softened this separation in his post-*Tractatus* reflections, attempting to clarify the motivations behind this shift. I will argue that, as emerges from his later works, the difference between genuine and formal concepts is not intrinsic but purely functional. It depends on how a concept is employed: whether as part of a hypothesis to be tested, or as a linguistic tool that provides a grammatical rule for describing reality. Based on this distinction, I will conclude by demonstrating that this thesis does not contradict Wittgenstein's view that grammatical rules cannot be justified empirically, but rather offers an alternative understanding of what it means to be empirical.

In the debate between the resolute readers of the *Tractatus* and the proponents of the traditional interpretation, one of the central points of contention concerns the correct understanding of a distinction that Wittgenstein introduces between two conceptual categories: genuine concepts and what he refers to as “formal concepts” (Hacker, 2000; Diamond, 1991, 1995). With this contribution, I do not intend to take a position in the dispute between these two interpretations of the *Tractatus*, but rather to clarify the distinction in question by analyzing the evolution of Wittgenstein's reflections on this point. I believe that such a clarification can offer useful tools for approaching the debate from a renewed perspective.

In sections 4.126–4.1272 of the *Tractatus*, Wittgenstein briefly introduces the distinction between formal concepts and genuine concepts. The latter are described in terms consistent with Frege's account (Frege, 1891): he offers a dual characterization of them, which I will provisionally label *functional* (TLP 1972: §4.126) and *intensional* (TLP 1972: §4.1241). According to the functional definition, concepts are understood as functions: unsaturated expressions that contain an empty place (a free variable) and, when saturated, yield a proposition. Wittgenstein, as is well known, consistently defines propositions in terms of the principle of bipolarity: a meaningful proposition must be capable of being both true and false. The sense of a proposition is independent of and precedes its truth-value, which can only be determined through comparison with the world. Therefore, for a proposition to be meaningful, it

must be conceivable both in a situation where it is true and one where it is false. In accordance with this principle, a proposition derived from the saturation of a concept's variables must, in order to qualify as a genuine proposition, be capable of being either true or false. As a function that must be capable of giving rise to a verifiable proposition and therefore of describing a determinate state of affairs, a concept must also contain within itself the criteria by which the description can be compared with experience. At this point, we shift to the intensional characterization of concepts, according to which a proposition resulting from the saturation of a concept is verifiable on the basis of the properties or set of properties associated with that concept, that is, the properties common to the objects that fall under the concept's extension. A conceptual function describes a state of affairs by attributing properties or relations to objects, and these properties serve as criteria for determining the truth or falsity of propositions. For this reason, a concept can be defined intensionally as a linguistic sign to which a set of features is associated, features that determine the properties of the objects that fall under it. By virtue of this view, different concepts can be distinguished on the basis of their respective characteristic features. One can, for instance, distinguish the concept of "birch" from that of "oak" by listing the properties possessed by one and not the other, and *vice versa*. However, due to the theoretical framework of the *Tractatus*, in particular its account of propositional bipolarity, Wittgenstein maintains that any description is meaningful only if the opposite situation is also conceivable. If it is a fact that birches have triangular leaves and firs have needle-like leaves, then the description distinguishing birches from firs based on their characteristic properties must also accommodate the possibility of the reverse, that is, a conceivable scenario in which the properties of these concepts are swapped.

To understand the reasoning behind this view, it is helpful to recall that, in addition to the "static" description of concepts, Wittgenstein shares with Frege the idea that the intensional content of a concept derives from the generalization of common properties observed in a set of similar objects (WWK 1979: 224). The genesis of the concept of "birch," for example, begins with the empirical recognition of a plurality of similar objects, which are grouped together by virtue of shared properties. These properties are then abstracted from individual cases and become associated with the general

concept of that class of objects, thus constituting the intensional content of the concept. This thesis on the empirical genesis of concepts implies that their intensional content is a posteriori: the characteristic features that define a concept are determined and confirmed (or falsified) through experience.

It is precisely this conception that leads Wittgenstein to assert that the distinction between concepts can be made only on the condition that the opposite situation is also conceivable. If it is true that birches are distinguished from firs by the shape of their leaves, it must nevertheless be imaginable (and meaningful) that birches might have needle-like leaves. Wittgenstein thus acknowledges that it is possible to formulate meaningful propositions concerning the properties of concepts precisely because those properties are grounded in experience. These, then, are the two characterizations that Wittgenstein offers regarding the nature of concepts within the theoretical framework of the *Tractatus*. On the basis of this dual definition, he distinguishes genuine concepts from what he calls “formal concepts”: linguistic signs with an apparently conceptual nature, which, however, cannot be considered genuine concepts (Wittgenstein in fact refers to them as pseudo-concepts), since they do not meet the criteria just outlined. Examples of such pseudo-concepts include: “number,” “function,” “proposition,” “color,” “object,” and so on. With regard to the first criterion, the functional one, Wittgenstein clarifies his position through examples. Consider the grammar of the word “color”: according to Wittgenstein, it represents a formal concept because, if treated as a function, the statements resulting from the saturation of its variables are not genuine propositions. This is because such statements are either explicitly nonsensical or violate the principle of bipolarity (LWL 1980: 10-11). For instance, the proposition “red is a color” seems to express an essential truth about the concept of red, but this very feature makes it unfit to be considered a genuine proposition, since its negation, “red is not a color,” is not conceivable (it is nonsensical). A genuine proposition, in fact, must be capable of being tested against experience, and thus its negation must also be conceivable. Since the statement cannot be empirically verified or falsified, it is neither true nor false, but simply nonsensical.

Turning now to the second criterion (the intensional one) Wittgenstein notes that formal concepts cannot be distinguished by listing the various features that define them (TLP 1972: §4.1241; LWL 1980: 46-47; BT 2005: 187e-188e). It is

not conceivable for an object falling under a formal concept to possess characteristics different from those that define it. For example, it is falsifiable to say that “birches have triangular leaves,” because we can imagine a situation in which birches have leaves different from those they actually have. But it makes no sense to distinguish sounds and colors by saying that the former have pitch and the latter have hue, because it is not conceivable for a sound to have hue or for a color to have pitch.

These are the reasons why Wittgenstein argues that, from a functional point of view, formal concepts cannot be represented by propositional functions, but must instead be expressed through propositional variables whose values are the objects that fall under the concept. While for genuine concepts it is possible to provide a linguistic description of their intensional content (i.e., by listing the properties that define them), no such description can exist for formal concepts, because their content does not consist of a set of properties but rather a form of expression within language. Formal concepts refer to the expressive form on which language itself rests and that it uses to represent the world. For this reason, language cannot express it through a description; any description in fact presupposes it.

The characteristics associated with formal concepts are not comparable to properties of the objects that fall under the concept, but rather to grammatical rules that govern the use of the linguistic signs referring to those objects. To clarify this point, Wittgenstein provides a telling example: the pseudo-proposition “x is a color” is not a proposition that attributes a property to an object, but a pseudo-proposition that places a certain sign among the values that a variable (the color variable) can assume (Black, 2007; Hamlyn, 1959). Here, the variable “color” represents a form of expression that determines a particular use of language, and the word “red” is one of the values of this variable (TLP 1972: 4.126-4.12712; PB 1975: 9). The sentence “red is a color,” therefore, does not say something about the nature of red, nor does it describe its essence, but rather signals that “red” is one of the values that the variable “color” can assume, thereby establishing the rules that must be followed in order to use it correctly. Wittgenstein writes that the so-called “characteristic features” of formal concepts, such as hue for colors, pitch for sounds, or extension for bodies, are not properties of the designated objects but grammatical traits of the linguistic symbols that designate them. Saying that

red is a color is equivalent to saying that it makes sense (not that it is true) to talk about the hue of red, its saturation, etc. While I can explain to someone what a birch is by listing its properties, I cannot explain what a color is in the same way: in order to understand it, the subject must already be familiar with at least one of the values that the variable “color”. This is why Wittgenstein states that it is not possible to describe separately a formal concept and an object that falls under it (TLP 1972: 4.1272). The moment an object falling under a formal concept is given, the formal concept itself is also given. So, while a genuine concept can have an empty extension (e.g., “unicorn”), since it is constructed through an aggregation of properties independent of the object they describe, a formal concept cannot have an empty extension: if it has no value, the sign of the formal concept cannot have meaning. It is no coincidence that Wittgenstein writes that the formal concept “implies its own existence” (BT 2005: 393e).

So far, these are Wittgenstein's ideas in the *Tractatus*. In this paper, I intend to show how Wittgenstein modifies his conception, gradually reducing the difference between genuine concepts and forms. First of all, Wittgenstein stops claiming that there are concepts with a special status. Instead, he believes that these concepts also have an ordinary meaning, just like all other concepts (YB 1982: 13; PI 2009: §97). Secondly, although he continues to maintain that concepts are functions, he abandons the idea that concepts should be understood as closed fences. Rather, he invites us to observe that the elements that make up the extension of a concept constitute a “family”, as their members are united by a broad network of similarities, not by a set of well-defined characteristics that all objects in the extension share (BBB 1991: 17; PI 2009: §67; BT 2005: 251e-256e). Wittgenstein instead contends that the similarity and unity among the objects falling under a concept are grounded in the use of examples and samples, which function as models for comparison. In this way, even genuine concepts take on traits of similarity with formal concepts: just as it is impossible to provide a description of the objects falling under a formal concept without providing a value for the variable, it also proves necessary to refer to a sample or an example in order to define a genuine concept. Not surprisingly, Wittgenstein seems to admit the possibility that even the concept of “man” could be understood as the sign of a form (a formal concept) rather than as a genuine concept (WWK 1979: 44). The reason

for this consideration stems from a problem in the translation of general propositions into Russellian language, where the concept of "man" begins to appear to Wittgenstein more like a variable than a function (Fogelin, 2002; McGinn, 2006).

I will now attempt to substantiate my argument, beginning with Wittgenstein's observation that one cannot distinguish between forms simply by listing the different properties they possess. Let us ask ourselves: to what extent is this statement true for genuine concepts? Indeed, it can be said that birches are distinguished from firs because birches have triangular leaves while firs have needle-like leaves. Modifying these characteristics would not lead us to think that birches have ceased to be birches and firs have ceased to be firs. But now let us try to compare the concept of "man" with the concept of "plant." One of the distinguishing features of plants is that they derive energy from the sun through photosynthesis, unlike humans, who nourish themselves through food. What should we say about the proposition "humans perform photosynthesis"? Is it simply false, or rather meaningless? Can we imagine humans performing photosynthesis? If one day we were to find humans capable of doing so, it might be a good reason to believe that these humans are not humans in the usual sense of the word. On the other hand, there are plants that do not perform photosynthesis but feed on living beings, such as carnivorous plants, and we certainly do not say that carnivorous plants are not plants. The crucial point, however, is that at some point, the distinction begins to become inconceivable. There is a moment when the attribution of unusual properties to a concept necessarily transforms it into a different concept. Could we hypothesize that cats grow on trees? Obviously not, and this idea is not even conceivable. If something resembling a cat were to grow on trees, we would have no reason to believe that these things are actually cats. The fact that cats do not grow on trees is not an empirical proposition but a grammatical one that expresses a criterion for how far something can be called a 'cat' in the ordinary sense. For this reason, each concept is always defined by a set of criteria (or internal properties) that limit the possibilities of producing meaningful propositions through it: one cannot only have symptoms but also criteria for concepts

The problem that arises when attempting to assimilate concepts and formal concepts is that this thesis appears to imply that even the grammar of formal

concepts could be verified through experience, thereby violating one of Wittgenstein's most fundamental claims: that grammatical rules cannot be justified empirically. In the conclusion of my contribution, I intend to explain why this interpretation does not create the problem it seems to suggest. As we have seen, both the meaning and the extension of a genuine concept depend on experience. Moreover, unlike formal concepts, a concept makes sense regardless of how things go in the world, and it continues to make sense even if its extension is empty (as in the case of unicorns or black holes, before their discovery). With formal concepts, the situation is reversed: whether a color has a hue does not depend on the world, because if the color didn't have a hue, we would not call it a color, just as in the case of cats. But if the meaning (the intension) of the formal concept does not depend on experience in this sense, since it is a variable defined by the set of its values, it can only have meaning by virtue of the existence of one of its values. The existence of an object that falls under the formal concept is the condition that makes it possible to use that object-(concept) as a paradigmatic example of an expressive form. Admitting this, however, does not mean asserting that formal concepts depend on experience in the same way that genuine concepts do. Wittgenstein claims that logic is independent of experience if by experience we mean the experience of what happens in the world, but it is not if by experience we mean the correlation that language structurally has with the world, and through which it makes sense (TLP 1972: 5.552; WWK 1979: 65, 217). Wittgenstein already states in the *Tractatus* that logic is independent of the experience of "how" but depends on the experience of "what." This kind of experience does not justify the correctness of grammar but is rather the experience that is reflected in the expressive forms of our language and makes the description of the world possible.

Having made this necessary clarification, Wittgenstein provides clear indications for distinguishing the two uses of concepts and forms. He distinguishes between hypotheses and postulates (LFM 1976: 129; YB 1982: 69-71). A hypothesis is a representation that can produce propositions verifiable or falsifiable by experience, like genuine concepts. A postulate, on the other hand, is a grammatical rule, whose validity is neither confirmed nor falsified by experience, as it is a tool for measuring reality that can be more or less practical. In my view, the difference between genuine concepts and formal

concepts is not intrinsic but depends on the use we make of a certain linguistic sign, depending on whether we use it to formulate hypotheses about the course of the world or grammatical rules to describe it. For example, the proposition "all men are mortal" can be seen as a proposition that predicates (either truly or falsely) a property to the concept of "man." If the word "man" is understood as a genuine concept, the proposition is a hypothesis that can be verified or falsified by experience. But if we consider "man" as a form, the proposition becomes a grammatical rule that specifies a formal property of the concept of man, providing a criterion for determining what is a man and what is not. Unlike hypotheses, which can be verified and falsified by experience, grammatical rules do not depend on experience but are rather linguistic tools through which we describe experience. On the other hand, Wittgenstein asserts that a grammatical rule can be said to be more or less practical, and the practicality or impracticality of grammatical rules is related to the application of such propositions, such as in the effective description of the world around us. This practicality, in turn, depends on experience, on how the world is made (LFM 1976: 129; LWL 1980: 17-18), but such dependence cannot serve as an empirical justification for the validity of concepts.

References

- Blank, Andreas (2007) "Material Points and Formal Concepts in the Early Wittgenstein", *Canadian Journal of Philosophy* 37, 245-261.
- Diamond, Cora (1991) "Throwing Away the Ladder: How to Read the *Tractatus*", *Philosophy* 63, 5-27.
- Diamond, Cora (1995) "Ethics, Imagination and The Method of Wittgenstein's *Tractatus*", *The New Wittgenstein*, 149-173.
- Frege, Gottlob (1891) *Function and Concept*, Jena: Hermann Pohle.
- Fogelin, Robert J. (2002) *Wittgenstein*, New York: Routledge & Kegan Paul Ltd.
- Hacker, P. M. S. (2000) "Was He Trying to Whistle it?", *The New Wittgenstein*, 353-388.
- Hamlyn, D. W. (1959) "Categories, Formal Concepts and Metaphysics", *Philosophy* 34, 111-124.
- McGinn, Marie (2006) *Elucidating the Tractatus. Wittgenstein's Early Philosophy of Logic & Language*, Oxford: Oxford University Press.
- Wittgenstein, Ludwig (1972), *Tractatus Logico-Philosophicus*, Übersetzt von D. F. Pears und B. F. McGuinness. Reprinted with corrections in the light of the publication of Wittgenstein's

correspondence with C. K. Ogden about the first translation, London: Routledge and Kegan Paul.

Wittgenstein, Ludwig (1975) *Philosophical Remarks*, Hg. von Rush Rhees, übersetzt von Raymond Hargreaves und Roger White, Oxford: Basil Blackwell.

Wittgenstein, Ludwig (1976), *Wittgenstein's Lectures on the Foundations of Mathematics, Cambridge, 1939*, Hg. von Cora Diamon, Ithaca, USA: Cornell University Press.

Wittgenstein, Ludwig (1979), *Wittgenstein and the Vienna Circle. Conversations recorded by Friedrich Waismann*, Hg. von B. F. McGuinness, übersetzt von Joachim Schulte und B. F. McGuinness, Oxford: Basil Blackwell.

Wittgenstein, Ludwig (1980), *Wittgenstein: Lectures, Cambridge 1930 -1932*, Hg. von Desmond Lee, Oxford: Basil Blackwell.

Wittgenstein, Ludwig (1982), *Wittgenstein's Lectures. Cambridge, 1932-1935*, S. 41–73. Hg. von Alice Ambrose, Oxford: Basil Blackwell.

Wittgenstein, Ludwig (1991), *Preliminary Studies for the "Philosophical Investigations". Generally Known as The Blue and Brown Books*, Oxford: Basil Blackwell.

Wittgenstein, Ludwig (2005), *The Big Typescript*, Hg. und übersetzt von C. Grant Luckhardt und Maximilian A. E. Aue. Malden (Mass.), Oxford, Carlton (Vic): Blackwell.

Wittgenstein, Ludwig (2009), *Philosophical Investigations*, Hg. von P. M. S. Hacker und Joachim Schulte, übersetzt von G. E. M. Anscombe, P. M. S. Hacker und Joachim Schulte, New York: Wiley.

Doing Justice to the Swansea School: An Appropriate Wittgensteinian Approach to the Philosophy of Religion

Jakub Gomułka, Jan Wawrzyniak (Kraków, Poland)

Abstract

The paper introduces a framework for understanding various forms of religious attitudes by employing Wittgenstein's method of objects of comparison. We develop a tree-structured set of models of religion, beginning with a baseline model and branching into three distinct approaches to religiosity.

The *baseline model* represents a form of religion that lacks a personal relationship with a deity. From this foundation three distinct models emerge: the *reciprocity model*, in which humans expect divine rewards, punishments, and answered prayers; the *quasi-aesthetic model*, in which humans merely contemplate God's perfection without expecting direct influence; finally, the *trust model*, in which faith transcends both transactional expectations and intellectual contemplation.

Each model faces specific criticisms: the *baseline model* may appear insufficiently religious; the *reciprocity model* can be seen as reducing religion to calculation; the *quasi-aesthetic model* portrays God as distant and abstract and the trust model is criticised as a form of ethical atheism in disguise. Nevertheless, each model captures important aspects of religious experience.

We argue that actual religious attitudes typically combine elements from multiple models rather than conforming exclusively to one. Our framework serves two purposes: first, it highlights religion's multifaceted nature, countering the tendency to assert a single "true" form of religion; second, it reveals that many philosophical debates on religion—including the debate concerning the Swansea School—arise from participants implicitly favouring different models and sensitivities. We conclude that philosophers, rather than advocating normative positions about the superiority of any particular model, should follow Wittgenstein's methodological approach and aim to clarify misunderstandings.

Introduction

The question of religion and humanity's relation to divinity can give rise to many misunderstandings. This text aims to demonstrate how such misunderstandings can be addressed using Wittgenstein's method of objects of comparison. In *The Emergence of Religious Narrative* (Gomułka, Wawrzyniak 2025), we applied the same approach to misconceptions in Wittgenstein's own thought and in that of his followers in the so-called Swansea School of the philosophy of religion. At the same time, we pointed out that this tradition offers important insights into religion. Here, we seek to explicate these insights. To do so, we will use a set of models of religiosity, structured as a tree

with one starting point and multiple branches. These models should be understood as objects of comparison in the sense of Wittgenstein's method of perspicuous representation.

The Method of Perspicuous Representation and Objects of Comparison

As Gordon Baker observed, Wittgenstein's method of objects of comparison serves to "liberate our thinking from enslavement to particular analogies by bringing to light other analogies which are equally well supported as the ones of which we unconsciously make use" (Baker 2004: 34). The author of the *Philosophical Investigations* is thus not concerned with developing a general theory of language; rather, his models serve as interventions in specific cases of misunderstandings (cf. Baker 2004: 27).

An object of comparison may take various forms: it can be a simplified fragment of real language, such as the apple-buying game in §1 of the *Investigations*; a more abstract construct, like the builders' game in §2; or even a fully fictional scenario, such as the "selling wood" example in §149 of the *Remarks on the Foundations of Mathematics*. In the latter case, the object of comparison highlights key differences in language use, revealing grammatical features through an example that ultimately diverges from actual practice.

As Oskari Kuusela emphasizes, this method enables Wittgenstein to avoid dogmatism: he does not treat his objects of comparison as exhaustive descriptions of phenomena, but rather as tools that highlight certain similarities or contrasts (cf. Kuusela 2008: 128–129). However, this does not imply that objects of comparison cannot offer adequate representations of reality. Of course, there is no absolutely adequate description—its adequacy depends upon our specific aims, that is, the particular philosophical problems we seek to dissolve.

The models of religiosity presented below serve as Wittgensteinian objects of comparison. Their purpose is to redirect debates surrounding key criticisms raised by the Swansea School. At the same time, they offer a set of adequate descriptions of possible forms of religiosity, aiming to resolve misunderstandings that have shaped—or even largely constituted—the debate. Some of these misunderstandings originate from Wittgensteinian philosophers

of religion, others arise from the limited religious imagination of certain critics of the Swansea School, leading to a restrictive view of possible relations between humans and God.

The Baseline Model

The collection of models we present has a tree structure, featuring a common root and three descendant nodes. Our aim is to illustrate a family of religious attitudes that emerge from a shared foundation—this foundation being the *baseline model*. It can be conceived either as a purely ritualistic stance devoid of theological reflection or as a highly intellectual construct.

In the first case, religion consists solely of emotional reactions and ritual acts associated with the sacred. As Rudolf Otto (1924) argued, initial responses to the *numinous* are inherently non-rational and may not require intellectual reflection within a cultural community. Crucially, this model allows for a religious attitude in which people neither expect nor desire anything from a deity—nor even recognize that such expectations are conceivable. One might say that believers simply believe, expressing their faith through adherence to taboos or the performance of rituals accompanied by certain emotions—but nothing more. This is, of course, an abstract conception, possibly never fully realised in practice.

The second version of our *baseline model* emerges from centuries of sophisticated theological reflection, resulting in a complete separation between the sacred and the profane. In this case, religious thinkers conclude that these realms are entirely incommensurable—that there can be no genuine contact between a human immersed in temporality and a transcendent god. Such a radical thesis entails paradoxical consequences, including the meaninglessness of ritual acts. This, too, is an abstract conception, but it is nonetheless conceivable. Indeed, it has already been imagined and developed into a coherent narrative of religion in a posthumanist world of self-aware machines by the science fiction writer Stanisław Lem. In *The Twenty-first Voyage* (written in the early 1970s; Lem 2016: 203–267), Lem explores a possible evolution of Catholic theology grounded in the principle of total separation between temporal and eternal orders. In this world, believing machines—computers and robots—hold no expectations of God and express no admiration for His works; their act of faith, they claim, is entirely *naked*.

In summary, the *baseline model* is characterised by the absence of a personal relationship with the sacred. This absence can manifest either as an unreflective engagement with religious experiences and rituals or, conversely, as a radical consequence of certain theological premises. Compared to the models introduced later in our text, this model is, in a sense, *limited*.

The Reciprocity Model

The first descendant node in our model tree is the *reciprocity model*, also known as the transactional model. In this model, a person's relationship with God is based upon expectations: one anticipates being rewarded or punished according to one's actions. Since God is just, He grants heavenly rewards for good deeds and condemns sinners to hell.

Another component of this model is the belief that God listens to human requests and may sometimes grant them—for instance, prayers for health or peace. However, this does not necessarily imply an automatic cause-and-effect relationship, as would be the case in magical thinking. In religions where gods are conceived as personal beings—whether monotheistic or polytheistic—these gods are typically viewed as autonomous in their decisions.

This understanding of religiosity is widespread. For instance, numerous biblical passages suggest that humans are justified in expecting rewards for good deeds and that God responds to prayers (cf. Mt 25:40–43; Mt 7:7–8).

The interpretation of this model depends on how reciprocity is understood. On the one hand, it can be viewed as transactional (Holland (1956) vividly describes this type of religiosity as the “greedy nephew approach”), where moral behavior, the avoidance of sin, and even belief in God itself are motivated by the expectation of reward and fear of punishment. Such religiosity is based on calculating one's future fate. Pascal's Wager provides a striking example of this transactional reasoning.

On the other hand, reciprocity can also be understood not in transactional terms, but rather as hope—the belief that God cares about human concerns. Here, God is perceived as intimately close, more like a friend than a distant, transcendent being. He is active in the world, sharing in human joys and struggles. The concept of an incarnate God fits well with this model; for instance, Jesus feasting at Cana and turning water into wine symbolises this

profound friendship between God and humanity. In this interpretation of the reciprocity model, what connects humans to God is not shared *interests* but shared *desires*.

The Quasi-Aesthetic Model

The second descendant model we consider is the *quasi-aesthetic model*. In this model, the human attitude toward God is primarily one of contemplation and admiration, grounded in a belief in God's perfection and immutability. God, as conceived here, does not possess human-like qualities. The extreme version of this approach—apophatic theology—rejects any attribution of properties to God, rendering Him entirely unknowable. Regardless of whether one adopts a moderate or radical version of this view, God is understood as distant. Within this model, neither human actions nor prayers can influence Him, and expecting prayer—understood as a request—to be granted constitutes a misunderstanding. One can only contemplate and admire God.

We call this *model quasi-aesthetic* (this is a concept used by Holland (1956), but we apply it here somewhat differently) because the relationship between humans and God is analogous to that between an observer and a work of art. Just as the thoughts, emotions, and actions of an audience (in the classical sense) do not affect the artwork itself, human engagement similarly does not alter God.

One source of this form of religiosity is the awe and admiration inspired by the vastness, beauty, and harmony of the universe—qualities through which the perfection of the Creator (or the Absolute, if identified with the universe) becomes visible. This attitude may also arise from reflection on and acceptance of apophatic theology or certain metaphysical systems, such as those of Descartes or Leibniz. Spinoza describes a similar stance, calling it the intellectual love of God. Traces of this attitude can also be found in Wittgenstein (e.g. TLP: 6.432).

This model differs significantly from the *reciprocity model*, as it does not assume a relationship of love between God and humans—a relationship central to many religious traditions. Despite its highly one-sided nature, however, it highlights key aspects of how both God and human religiosity can be understood. It emphasizes God's infinity, our inability to comprehend Him, and the entirely unselfish nature of religious admiration. The *quasi-aesthetic*

model thus serves as a counterbalance to the transactional version of the *reciprocity model*, which relies on a form of infantile anthropomorphism.

The Trust Model

The final model we present is the *trust model*, named after Wittgenstein's remark that true faith—contrasted with popular religion—is based on trust. Wittgensteinian fideists supported this model, regarding the *reciprocity* and *quasi-aesthetic models* as forms of the popular religion they criticised (cf. Holland 1956). Thus, the *trust model* is neither grounded in expectations that God will fulfill promises, nor in intellectual or aesthetic contemplation of the Creator's power or the grandeur of creation.

This model assumes a complete distinction between the temporal and the sacred realms, resembling the theological version of the *baseline model*. The key difference lies in the believer's position: in the *baseline model*, the believer remains entirely within temporality, whereas in the *trust model*, believers can participate in eternal reality. However, the term reality should be used cautiously, as it does not imply a replication of anything familiar from our world. The radical otherness of the non-temporal order means that none of our conceptual frameworks—based as they are on knowledge of the temporal world—can adequately apply to it. Terms like eternal life are merely pictures, whose true meaning can only be grasped from within the perspective of this non-temporal order (cf. Phillips 1996: 246).

The defining feature of this model is absolute trust—an experience Wittgenstein (1965) also described in his *Lecture on Ethics*. Here, goodness appears as a force beyond the world, opposing the brutal laws of temporality, which Simone Weil (2002) calls *gravity*. However, this force—*grace*—does not manifest itself as a miraculous supernatural intervention by an omnipotent God into nature. Its observable influence on physical reality operates solely through human moral sensitivity and the actions arising from it. In this view, the sacred is deeply connected to human life, yet is still understood as something not of this world.

Debates Over the Models

Each model is subject to criticism from within religious traditions. The *baseline model* may be dismissed as not genuinely religious, as it excludes personal engagement. Its ritual-experiential form presents a religiosity devoid of

reflection, grounded in rituals whose connection to communal life is unclear. The theological version, while intellectually rich, appears existentially barren, raising the question of why anyone would engage with it; notably, in Lem's story, only machines remain religious.

The *reciprocity model* is often criticised for its transactional aspect, which reduces religion to a utilitarian relationship with God. In its crudest form, it suggests that believers follow religious precepts primarily out of expectation of reward or fear of punishment—a calculated balance of gains and losses. Rush Rhees (1997: 36) mocked this attitude: "Think of your future, boy, and don't throw away your chances". His student, Dewi Z. Phillips, argued that petitionary prayer, if understood as an attempt to influence God toward a desired outcome, is superstitious and lacks genuine religious character (Phillips 1965: 118). More broadly, the *reciprocity model* is frequently regarded as indicative of spiritual immaturity.

The *quasi-aesthetic model* has also faced extensive criticism from figures such as Pascal, James, and Heidegger. Critics argue that this model portrays God as a detached philosophical construct rather than an object of existential commitment—the God of faith is not the "God of philosophers" but the "God of Abraham, Isaac, and Jacob". Within this framework, the sacred is remote and inaccessible, seemingly incompatible with the relationship of love between God and humanity central to the Judeo-Christian tradition.

Finally, the *trust model* has been criticized as a thinly veiled form of ethically sensitive atheism, in which God lacks any genuine transcendent referent. Critics argue that this model attempts to redefine religion and spirituality under the influence of an anti-metaphysical bias, thus erasing what is essential to these concepts (e.g. Hick 1977).

Each model, however, can be defended as highlighting a significant aspect of religion. Even the *baseline model* reveals the unmediated core of ritual and religious experience, while also emphasizing the transcendent nature of the sacred. The *reciprocity model* portrays God as engaged in human affairs—immanent, caring, and relational—rather than as a distant Absolute. The *quasi-aesthetic model*, like the theological version of the *baseline model*, purifies religion of infantile anthropomorphism, drawing attention to the unselfish

character of religious admiration expressed through awe and contemplation. Lastly, the *trust model* can be viewed as an attempt to articulate the radical otherness of the sacred without severing it from concrete human existence, seeking a balance between transcendence and immanence. It also offers a conception of spirituality that does not depend on metaphysical commitments.

Conclusion

A real relationship with God (or gods) does not necessarily conform exclusively to one of the models presented above. In fact, most believers' actual attitudes likely combine elements drawn from multiple models. In this sense, our models serve as ideal types—abstractions derived from real attitudes to construct a perspicuous representation of the variety of religious perspectives.

This collection of models serves two interconnected purposes. First, it highlights the multifaceted nature of religion. Many philosophers of religion—including Wittgenstein's followers and, to some extent, Wittgenstein himself—tend to theorize about religion in a substantialist manner, presuming that true religion possesses a specific, clearly defined nature. We argue that this approach is philosophically misguided. Following Wittgenstein's methodological advice, philosophers should adopt a descriptive rather than normative stance and, above all, recognise the diversity inherent in religious life.

Second, we aim to show that some philosophical debates about religion are not genuinely philosophical but rather religious disputes, rooted in the underlying models of religiosity adopted by the participants. For example, the disagreements between Phillips (1977; 2005: 49–94) and Hick (1977), Swinburne (1977), or Plantinga (1990) stem from a clash between Phillips's *trust model*—which he considers correct—and the *reciprocity model*, which his opponents regard as normative. Preferences for different models are undoubtedly shaped by differences in the religious sensibilities of those involved in these debates.

We do not claim that conflicts between different religious attitudes can—or should—be resolved, as that, too, would represent a normative stance. In our view, it is not the philosopher's role to dictate what people should believe or how they should act. Wittgenstein remarked that proper philosophy “leaves

everything as it is". At the same time, clarifying misunderstandings—the true task of philosophers—can foster greater mutual understanding and reveal that some disagreements are of a different nature than they initially appear.

References

- Baker, Gordon (2004), "Philosophical Investigations §122: Neglected Aspects", in: idem, *Wittgenstein's Method: Neglected Aspects*, edited by Katherine J. Morris, Oxford: Blackwell, 22–51.
- Gomułka, Jakub; Wawrzyniak, Jan (2025), "The Emergence of Religious Narrative", *Religions* 16 (3), 318; <https://doi.org/10.3390/rel16030318>.
- Hick, John (1977), "Remarks", in: Stuart C. Brown (ed.), *Reason and Religion*, Ithaca & London: Cornell University Press, 122–8.
- Holland, Roy F. (1956), "Religious Discourse and Theological Discourse", *Australasian Journal of Philosophy* 34 (3), 147–63.
- Kuusela, Oskari (2008), *The Struggle Against Dogmatism: Wittgenstein and the Concept of Philosophy*, Cambridge (MA): Harvard University Press.
- Lem, Stanisław (2016), *The Star Diaries*, translated by Michael Kandel, London: Penguin Books.
- Otto, Rudolf (1924), *The Idea of the Holy: An Inquiry into the Non-Rational Factor in the Idea of the Divine and its Relation to the Rational*, translated by John W. Harvey. Oxford: Oxford University Press.
- Phillips, Dewi Z. (1965), *The Concept of Prayer*, London: Routledge & Kegan Paul.
- Phillips, Dewi Z. (1977), "The Problem of Evil", in: Stuart C. Brown (ed.), *Reason and Religion*, Ithaca & London: Cornell University Press, 103–21.
- Phillips, Dewi Z. (1996), "Dislocating the soul", in idem (ed.), *Can religion be explained away?* New York: St. Martin's Press, 233–51.
- Phillips, Dewi Z. (2005), *The Problem of Evil and the Problem of God*, Minneapolis: Fortress Press.
- Plantinga, Alvin (1990), "God, Evil, and the Metaphysics of Freedom", in: Marilyn Adams, Robert Adams (eds.), *The Problem of Evil*, Oxford: Oxford University Press, 83–109.
- Rhees, Rush (1997), *Rush Rhees on Religion and Philosophy*, edited by Dewi Z. Phillips, Mario von der Ruhr, Oxford: Oxford University Press.
- Swinburne, Richard (1977), "The Problem of Evil", in: Stuart C. Brown (ed.), *Reason and Religion*, Ithaca & London: Cornell University Press, 81–102.

Winch, Peter G. (1964), "Understanding a Primitive Society", *American Philosophical Quarterly* 1 (4), 307–24.

Weil, Simone (2002), *Gravity and Grace*, translated by Emma Crawford and Mario von der Ruhr, London: Routledge.

Wittgenstein, Ludwig (1965), "A Lecture of Ethics", *The Philosophical Review* 74 (1), 3–12.

Wittgenstein, Ludwig (2023). *Tractatus Logico-Philosophicus*, translated with an Introduction and Notes by Michael Beaney, Oxford: Oxford University Press.

Frauen im Wiener Kreis: Zur epistemischen Ungerechtigkeit und Kanonkritik in der analytischen Philosophie

Patricia Grill (Wien, Österreich)

Abstract

Die weiblichen Mitglieder des Wiener Kreises stehen in der Philosophiegeschichte noch immer im Schatten ihrer männlichen Kollegen. Wenn sie überhaupt Erwähnung finden, dann meist nur in relationalen Rollen – als Ehefrauen, Studentinnen oder Assistentinnen. Besonders deutlich zeigt sich dies an den Fällen von Rose Rand und Olga Hahn-Neurath, deren intellektuelle Beiträge trotz fachlicher Relevanz marginalisiert wurden. Aufbauend auf aktuellen Arbeiten von Connell und Janssen-Lauret (2022), Beaney (2023) und Fricker (2023) analysiert dieser Beitrag ihre systematische Unsichtbarmachung als Ausdruck epistemischer Ungerechtigkeit und als Teil eines „feedback loop of exclusion“ (Connell/Janssen-Lauret 2022) in der Philosophiegeschichtsschreibung. Ziel ist es, nicht nur biografische Ergänzungen vorzunehmen, sondern die erkenntnistheoretischen und kanonischen Voraussetzungen philosophischer Relevanz kritisch zu reflektieren – und so zu einer feministischen Revision der analytischen Philosophie beizutragen.

Der Wiener Kreis war eine Gruppe von Akademiker*innen, die – unter Einbeziehung des sogenannten „ersten Wiener Kreises“ – von etwa 1907 bis 1936 zentrale Fragen der Philosophie und Wissenschaft diskutierte. Im Zentrum stand die Forderung, Aussagen nur dann als philosophisch oder wissenschaftlich relevant zu betrachten, wenn sie logisch beweisbar oder empirisch überprüfbar waren (Stadler 2015: 1–27). Die Mitglieder kritisierten die traditionelle Philosophie als esoterisch und setzten sich für eine Erneuerung zugunsten einer „wissenschaftlichen Weltauffassung“ ein, die keine „philosophische Methode“ anbot und sich ganz auf die Einzelwissenschaften stützte (Frank 2016: 23–30; Carnap 1931; Verein Ernst Mach 2013). In diesem Rahmen trugen sie maßgeblich zur Entwicklung des Logischen Empirismus bei.

In der öffentlichen Wahrnehmung wird der Wiener Kreis vor allem mit männlichen Philosophen wie Rudolf Carnap, Hans Hahn, Moritz Schlick und Otto Neurath verbunden – ihre Arbeiten gelten als kanonisch für den Logischen Empirismus. Weibliche Mitglieder wie Rose Rand und Olga Hahn-Neurath finden hingegen kaum Beachtung. So listet der Beitrag „Logical Empiricism“ in der *Stanford Encyclopedia of Philosophy* (2023) 28 Männer als zentrale Akteure, darunter 12 aus dem Umfeld des Wiener Kreises, aber keine einzige Frau.

Doch die geringe Sichtbarkeit in der Rezeption bedeutet keineswegs, dass Frauen im Wiener Kreis nicht präsent waren: An den Donnerstagabendtreffen am Mathematischen Institut in der Wiener Boltzmanngasse nahmen neben Waismann und Feigl auch Rand und Hahn-Neurath teil (Stadler 2015: 29, 48). Ihre Beiträge blieben jedoch im Schatten der männlichen Kollegen. Während das Werk fast aller Männer des Wiener Kreises seit den 1980er Jahren umfassend aufgearbeitet wurde, ist die Forschung zu den intellektuellen Leistungen der Philosophinnen bis heute lückenhaft. Viele von ihnen tauchen – wie zahlreiche Denkerinnen des 19. und 20. Jahrhunderts – allenfalls in Fußnoten oder Randbemerkungen auf (Connell/Janssen-Lauret 2023: 249).

Michael Beaney (2023) greift dieses Problem in seinem Artikel „Open-mindedness and ajar-mindedness in history of philosophy“ auf. Er plädiert dafür, bestehende philosophische Narrative kritisch zu hinterfragen und neue Perspektiven zuzulassen – ohne jedoch eine vollständige Umstrukturierung der Geschichtsschreibung zu erzwingen.

Diese Arbeit versteht sich als ein Beitrag zur Sichtbarmachung weiblicher Stimmen im Wiener Kreis – ohne den Anspruch auf Vollständigkeit zu erheben. Sie macht deutlich, dass Frauen in der Philosophiegeschichtsschreibung weder in gleichem Maße noch nach denselben Maßstäben berücksichtigt wurden wie Männer. Der Beitrag bewegt sich im Schnittfeld historischer Rekonstruktion und systematischer Kanonkritik. Aufbauend auf Fallstudien zu Olga Hahn-Neurath und Rose Rand wird argumentiert, dass deren Marginalisierung nicht bloß ein historiografisches Defizit darstellt, sondern als Form epistemischer Ungerechtigkeit im Sinne Miranda Frickers (2023) zu verstehen ist. Zugleich wird aufgezeigt, wie narrative Strukturen der Philosophiegeschichtsschreibung – etwa durch selektive Rezeption – zu einem „feedback loop of exclusion“ führen (Connell/Janssen-Lauret 2022: 201). Ziel des Beitrags ist es, historische Sichtbarkeit mit systematischer Reflexion über die Bedingungen philosophischer Relevanz zu verbinden – und so einen Beitrag zur Korrektur und Ergänzung des philosophischen Kanons zu leisten.

1. Olga Hahn inmitten der Anfänge des Wiener Kreises

1929 wurde die Diskussionsgruppe unter dem Namen Wiener Kreis der Öffentlichkeit bekannt (Verein Ernst Mach 2013). Dieses Jahr markiert jedoch

nicht den Beginn der philosophischen Debatten. Bereits 1907 hatte sich eine Gruppe von Wissenschaftler*innen zusammengeschlossen. Haller (1985) bezeichnet diese frühe Phase als den „ersten Wiener Kreis“ und verweist auf eine Gruppe um Philipp Frank, Hans Hahn und Otto Neurath, die eine wöchentliche Diskussionsrunde im Rahmen der Philosophischen Gesellschaft bildeten. Ihr Ziel war die Entwicklung eines „neuen Positivismus“, der unter anderem auf den Ideen von Ernst Mach beruhte. Schon damals beschäftigten sie sich mit Wissenschaftsphilosophie und Logik und suchten nach Möglichkeiten, die Philosophie mit den aktuellen Debatten der Naturwissenschaften zu verknüpfen (Limbeck-Lilienau 2022: 99–101).

Eine Person, die in diesem Kontext häufig unerwähnt bleibt, ist die Philosophin und Mathematikerin Olga Hahn. Obwohl sie – wie Kurt Reidemeister es ausdrückte – „a man’s intellect“ besaß (Neurath 1984, zitiert nach Burke/Sandner 2022: 110), wurde sie in der Historiografie der Philosophie nicht entsprechend gewürdigt. Stattdessen blieb sie zugunsten ihrer männlichen Diskussionskollegen weitgehend unsichtbar.

Das bedeutet jedoch nicht, dass sie bereits zu ihrer Zeit vollständig aus dem wissenschaftlichen Diskurs ausgeschlossen war: Trotz einer Sehnerventzündung, die sie im Alter von 22 Jahren erblinden ließ, promovierte Olga Hahn sieben Jahre später, im Jahr 1911, als dritte Frau an der Universität Wien im Fach Philosophie mit der Arbeit *Über die Koeffizienten einer logischen Gleichung und ihre Beziehungen zur Lehre von den Schlüssen* (Burke/Sandner 2022: 107–110). Bereits in den Jahren 1909 und 1910 publizierte sie gemeinsam mit Otto Neurath international beachtete Arbeiten (Stadler 2015: 8), darunter den Aufsatz „Zum Dualismus in der Logik“, in dem sie bedeutende Änderungen der Axiome und des Beweisverfahrens von Schröders System vorschlugen (Janssen-Lauret 2022: 128). C. I. Lewis zählte ihre frühen mathematischen Arbeiten in seinem *Survey of Symbolic Logic* zu den „most important contributions to symbolic logic“ (Stadler 2015: 442).

Auch nach ihrer Promotion blieb sie eine engagierte Wissenschaftlerin. Briefwechsel, unter anderem mit Rudolf Carnap, belegen ihre fortwährende Beteiligung am wissenschaftlichen Diskurs (Carnap/Neurath 2024: 201, 298, 307). Unterstützung erhielt sie insbesondere von ihrer Kreiskollegin Rose Rand, die ihr wissenschaftliche Texte über Logik vorlas (Carnap/Neurath 2024:

328, 329). Dennoch blieben ihre Handlungsmöglichkeiten stark eingeschränkt. Janssen-Lauret (2022) fasst ihre Situation treffend zusammen: „Hahn was triply marginalized as a woman, a Jewish scholar, and a person with a disability.“ (128).

Ein weiterer Grund für die Unterrepräsentation von Philosophinnen in der Geschichtsschreibung ist die fehlende Übersetzung ihrer Beiträge (Janssen-Lauret 2022: 129–132). Ein Beispiel für diese Vernachlässigung ist auch Hahn-Neuraths Dissertation, die erst 2024 von Richard Lawrence und Justin Vlasits ins Englische übersetzt wurde.

Trotz ihrer Leistungen wird Olga Hahn-Neurath in der Rezeption vor allem als von Otto Neurath unterstützte Wissenschaftlerin dargestellt. Diese Sichtweise hängt auch damit zusammen, dass Neurath sie nach seiner Rückkehr aus Berlin 1906 beim Fortsetzen ihrer akademischen Laufbahn unterstützte – unter anderem durch einen Lesekreis und Mathematikübungen zur Prüfungsvorbereitung.

Ihre Arbeiten wurden selten unabhängig von Neurath betrachtet, sondern meist im Kontext seiner Förderung rezipiert. Auch er selbst trug – bewusst oder unbewusst – zur geschlechtsspezifischen Rollenzuschreibung bei, die ihre wissenschaftliche Eigenständigkeit überlagerte: Nach ihrer Heirat im Jahr 1912 wurde erwartet, dass sie sich um seinen Sohn Paul kümmerte, dessen Mutter bei der Geburt verstorben war (Burke/Sandner 2022: 106–108). In der Ehe blieb Neuraths eigene wissenschaftliche Tätigkeit – insbesondere seine Museumsarbeit – zentral. Auch nach der Emigration 1934 in die Niederlande war Olga Hahn-Neurath häufig auf sich allein gestellt: blind und ohne nennenswerte Unterstützung ihres Mannes.

Obwohl Hahn-Neurath Beiträge zu den Diskussionen des Wiener Kreises leistete und mathematisch versiert war, wird sie in der Geschichtsschreibung meist nur als „Schwester von Hans Hahn“ oder „Ehefrau von Otto Neurath“ erwähnt (siehe etwa Reck 2022: 140). Ihre biografische Verbindung zu diesen Männern ist unbestritten – ihre Anerkennung als eigenständige Wissenschaftlerin hingegen bis heute unzureichend.

Das zeigt sich auch in den historiografischen Darstellungen des „ersten Wiener Kreises“: Olga Hahn wird darin kaum erwähnt. Stadler (2015: 436) weist

jedoch darauf hin, dass sie neben Neurath, Frank und ihrem Bruder Hans Hahn eine Protagonistin des Urkreises des Logischen Empirismus war. Daher kann sie – ebenso wie ihre männlichen Kollegen – als Pionierin des Wiener Kreises betrachtet werden. Spätestens mit Moritz Schlicks Berufung nach Wien 1922 und den regelmäßigen Donnerstagabendtreffen ab dem Wintersemester 1924/25 nahm der Wiener Kreis seine endgültige Form an (Stadler 2015: 29). In der Forschung wird Olga Hahn-Neurath zunehmend als Mitinitiatorin anerkannt (Stadler 2015: 48), bleibt aber im wissenschaftlichen Diskurs deutlich weniger rezipiert als ihre männlichen Kollegen.

Der Fall Hahn-Neurath verdeutlicht eine Form des sogenannten „Matilda-Effekts“ – ein Begriff, den Margaret W. Rossiter 1993 erstmals prägte. Er beschreibt das strukturelle Muster, dass die intellektuellen Beiträge von Frauen in der Wissenschaft systematisch übersehen, abgewertet oder männlichen Kollegen zugeschrieben werden. Rossiter (2003: 195–204) weist darauf hin, dass diese Form der Unsichtbarmachung nicht nur Einzelfälle betrifft, sondern durch Enzyklopädien, Zitationsgewohnheiten und wissenschaftshistorische Narrative institutionell verankert ist. Frauen erscheinen darin häufig nicht als eigenständige Denkerinnen, sondern in relationalen Rollen – etwa als Ehefrauen oder „Muse“ männlicher Wissenschaftler.

Auch Hahn-Neurath wurde postum vor allem in ihrer biografischen Beziehung zu Hans Hahn und Otto Neurath wahrgenommen, nicht jedoch als eigenständige Philosophin. Die Tatsache, dass sie trotz ihrer mathematisch-logischen Beiträge kaum unabhängig rezipiert wurde, verdeutlicht am Beispiel des Wiener Kreises, wie tief der Matilda-Effekt in die Geschichtsschreibung der Philosophie eingeschrieben ist.

2. Rose Rand: Die Chronistin des Wiener Kreises

Im Gegensatz zu Olga Hahn-Neurath war Rose Rand keine „Ehefrau von“ oder „Schwester von“, sondern eine „Studentin von“ Moritz Schlick und damit Teil der jüngeren Generation des Wiener Kreises (Stadler 2015: 59). Innerhalb der Diskussionsgruppe übernahm sie die Rolle der Protokollantin – eine Funktion, auf die ihr Name in der öffentlichen Rezeption meist reduziert wird, sofern er überhaupt Erwähnung findet. Tatsächlich war sie jedoch damit für die einzigen erhaltenen Quellen über die Diskussionen im Schlick-Zirkel in den

Jahren 1930–1933 verantwortlich (Stadler 2015: 497) und spielte somit eine zentrale Rolle für die Historisierung des Wiener Kreises. Dennoch wurde sie, ähnlich wie ihre Kollegin Olga Hahn-Neurath, in der Rezeption oft auf eine unterstützende Funktion reduziert – als „Assistentin“ statt als eigenständige Denkerin.

Rand hatte große Schwierigkeiten, sich nach ihrer Flucht vor dem Nationalsozialismus als Wissenschaftlerin zu etablieren. In ihrem akademischen Exil in England musste sie trotz fachlicher Kompetenz um Anerkennung kämpfen. Obwohl sie als versierte Logikerin galt, blieb ihre Arbeit zur mathematischen Logik weitgehend unbeachtet (Mihaljević 2023). Ihre Geschichte steht exemplarisch für die strukturellen Hürden, die Frauen im Zugang zur akademischen Welt begegneten.

Trotz der Wertschätzung durch einige ihrer Kreiskollegen, darunter Felix Kaufmann, konnten auch gelegentliche Unterstützungsversuche letztlich nicht verhindern, dass Rand aus dem wissenschaftlichen Diskurs verdrängt wurde (Mihaljević 2023: 297–316; Stadler 2015: 497).

Zudem sah sie sich – wie viele andere Frauen in der Wissenschaft – mit sexistischen Vorurteilen und abwertenden Einschätzungen konfrontiert (Connell/Janssen-Lauret 2022: 200–207). Obwohl die Forschung keinen Zweifel an Rands Kompetenz als mathematische Logikerin lässt (Mihaljević 2023: 307), waren zeitgenössische Urteile über ihren Charakter ambivalent. Während etwa Felix Kaufmann sowohl ihren Intellekt als auch ihre Persönlichkeit würdigte, betonten Rudolf Carnap und Ludwig Wittgenstein eher ihre als kompliziert wahrgenommene zwischenmenschliche Dynamik (Mihaljević 2023: 302–313). Auch Otto Neurath versuchte sie zwar finanziell zu unterstützen, äußerte jedoch Zweifel an ihrer fachlichen Eignung – insbesondere bei der Darstellung zentraler Positionen des Wiener Kreises. In einem Brief an Carnap schrieb er 1934:

Daß sie nicht sehr einfühlsam referiert und nicht sehr zu so etwas verwendbar ist, weiß ich leider seit langem, sonst würde ich ja auch längst daran gedacht haben, sie ständig zu beschäftigen. Sie ist nicht etwa ein Hempel im Kleinen, sondern ein schwieriger Mensch und ein ärmster Teufel, dem man schwer helfen kann. Was Du von Fragebogen sagst, ist selbstverständlich richtig. Das

wird aber alles erst gemacht, bis das Institut im Gang ist. Dann aber nicht durch Rand, sondern durch jemanden, der auch sinnvoll interviewen kann. (Carnap/Neurath 2024: 182).

Diese Passage zeigt deutlich, dass Neurath Rands fachliche Fähigkeiten nicht unabhängig von seiner persönlichen Einschätzung ihres Charakters beurteilte. Diese Tendenz, persönliche Einschätzungen über den Charakter einer Wissenschaftlerin in den Vordergrund zu stellen, hatte weitreichende Konsequenzen. Katarina Mihaljević (2023) zeigt in ihrer Untersuchung zu Rand, dass in Archivmaterialien oft Charakterbeschreibungen anstelle ihrer wissenschaftlichen Qualifikationen ausschlaggebend für Entscheidungen über finanzielle Unterstützung waren:

Interestingly, examination of the archival material shows that information concerning Rand's character, rather than her scholarly status, as assessed by the committee members themselves, became crucial in making award decisions between her arrival in the United Kingdom in 1939 and her departure in 1952. (Mihaljević 2023: 312)

Die systematische Abwertung von Rands wissenschaftlicher Kompetenz zugunsten von Charaktereinschätzungen kann im Sinne Miranda Frickers (2023) als Ausdruck epistemischer Ungerechtigkeit gedeutet werden – konkret als *testimonial injustice*. Fricker beschreibt damit Situationen, in denen Sprecher*innen aufgrund sozialer Vorurteile weniger epistemische Glaubwürdigkeit zugeschrieben wird. Rand wurde – trotz nachgewiesener Expertise – immer wieder als „schwierig“ beschrieben. Solche Zuschreibungen beeinflussten nicht nur ihre Position innerhalb des Wiener Kreises, sondern auch die Entscheidungen über finanzielle Förderung und wissenschaftliche Rezeption. Ihre Rolle als Protokollantin, die zentrale Quellen für die Historisierung des Schlick-Zirkels lieferte, wurde dabei ebenso unterschätzt wie ihre eigenen philosophischen Arbeiten zur Logik. Frickers Konzept erlaubt es, diese individuelle Abwertung als systemischen Ausschlussmechanismus zu verstehen, der strukturell mit anderen Formen der Diskriminierung – etwa Sexismus und Exil – verschränkt ist.

Diese Form epistemischer Ungerechtigkeit fällt umso stärker ins Gewicht, wenn man sie mit männlichen Philosophen derselben Epoche vergleicht. So

galt etwa Ludwig Wittgenstein ebenfalls als schwieriger Charakter (Edmonds 2021: 59–74; Stadler 2015: 230), doch seine wissenschaftliche Autorität blieb davon nicht nur unberührt, sondern seine Exzentrizität wurde ganz im Gegenteil als Ausdruck von Originalität und Tiefsinn gedeutet, was seine Stellung im Kanon eher stärkte als schwächte. Während Rands Persönlichkeit als Hindernis für ihre Anerkennung diente, wurde Wittgensteins Eigenwilligkeit als Teil seiner philosophischen Genialität gerahmt. Dieser Kontrast zeigt, dass ‚problematisches‘ Verhalten nicht per se zur Marginalisierung führt – sondern dass soziale Zuschreibungen eine zentrale Rolle dabei spielen, wer in der Wissenschaft als glaubwürdig und relevant gilt. In Rands Fall wurde die Bewertung ihrer Person zur dauerhaften Einschränkung ihrer epistemischen Autorität.

Zudem übernahmen Historiker*innen oft unkritisch die Perspektiven bekannter Philosophen, die Frauen wie Rand bereits zu ihrer Zeit als weniger bedeutend einstufte. Auf diese Weise verfestigt sich der philosophische Kanon: Frauen, deren Arbeiten zu Lebzeiten kaum Rezeption erfuhren, bleiben auch in der retrospektiven Forschung unsichtbar – sie gelten als „weniger relevant“, gerade weil sie zuvor bereits marginalisiert wurden (Connell/Janssen-Lauret 2022: 200–204). Connell und Janssen-Lauret (2022: 201–203) bezeichnen diesen zirkulären Ausschlussmechanismus als „feedback loop of exclusion“: Wer in der Vergangenheit ignoriert wurde, gilt auch in der Gegenwart als verzichtbar – und fehlt folglich auch in zukünftigen Narrativen.

Ein deutliches Beispiel für diese fortwährende Vernachlässigung zeigt sich im bereits erwähnten Beitrag „Logical Empiricism“ von Richard Creath (2023) in der *Stanford Encyclopedia of Philosophy*. Rand wird dort lediglich einmal erwähnt – und das nur in einer Randnotiz: „In 1939 Rose Rand, a less well-known member of the Vienna Circle, fled to England [...].“

Wie Janssen-Lauret (2022) betont, war Rand damit kein Einzelfall. Viele Frauen der analytischen Philosophie erlebten eine dreifache Marginalisierung: Sie waren institutionellem Sexismus ausgesetzt, litten unter den zusätzlichen Belastungen faschistischer Verfolgung und wurden schließlich von späteren Historiker*innen übersehen.

3. Die „Türen einen Spalt öffnen“

Michael Beaney (2023) Begriff der „ajar-mindedness“ liefert eine konzeptionelle Grundlage, um Denkerinnen wie Olga Hahn-Neurath und Rose Rand in den philosophischen Diskurs aufzunehmen. Die Metapher der „einen Spalt geöffneten Türen“ verweist auf eine spezifische Form epistemischer und institutioneller Offenheit, die sich bewusst zwischen zwei Extremen positioniert: Einerseits grenzt sie sich ab von geschlossenen Türen, die marginalisierte Perspektiven ausschließen; andererseits von vollständig geöffneten Türen, die zur Beliebigkeit oder Überforderung führen könnten. „Ajar-mindedness“ beschreibt eine Haltung und Praxis, die Türen bewusst (nur) einen Spalt offenlässt – gerade weit genug, um alternative Perspektiven sichtbar und zugänglich zu machen, ohne die Notwendigkeit epistemischer Struktur gänzlich aufzugeben. Beaney (2023) schreibt dazu: „Who knows what rewards might lie in the locked rooms, but we need the doors to be open to give scholars the chance of looking inside“ (218).

Diese Form der Offenheit zielt nicht nur auf die Erweiterung philosophischer Perspektiven, sondern auch auf die kritische Auseinandersetzung mit den strukturellen Bedingungen des Ausschlusses. Es geht darum, zu analysieren, warum bestimmte Personen und Positionen aus dem philosophischen Kanon ausgeschlossen wurden – und wie sich dies künftig verhindern lässt (Beaney 2023: 219). „Ajar-mindedness“ verweist somit auf die Notwendigkeit, institutionelle Rahmenbedingungen zu schaffen, die nicht nur neue Stimmen zulassen, sondern aktiv deren Sichtbarkeit und Anerkennung fördern.

In Bezug auf die Frauen des Wiener Kreises bedeutet das nicht, dass etwa Olga Hahn-Neurath oder Rose Rand die Bedeutung von Carnap oder Neurath schmälern sollen. Vielmehr geht es darum, ihre Beiträge überhaupt erst in den Blick zu nehmen – als wesentliche Ergänzungen, ohne die das historische und systematische Bild des Kreises verzerrt und fragmentarisch bleibt. Ziel ist es, ihre Leistungen sichtbar zu machen und anzuerkennen, dass nicht nur die männlichen Mitglieder den Kreis geprägt haben, sondern dass die kanonischen Texte des Kreises durch Philosophinnen inspiriert und herausgefordert wurden. Feministische Theoretikerinnen wie Helen Longino (1990) betonen, dass Wissenschaft ein kollektiver, sozial eingebetteter Prozess ist – ein Prozess, an dem Frauen nicht nur teilhaben, sondern den sie auch aktiv mitgestalten.

Die Fallstudien zu Olga Hahn-Neurath und Rose Rand bringen genau dies zum Ausdruck: Eine feministische Philosophiegeschichtsschreibung darf nicht bei einer bloß additiven Ergänzungsarbeit stehen bleiben. Vielmehr muss sie systematisch reflektieren, wie epistemische Machtverhältnisse die Grenzen des Sichtbaren und Anerkannten konstituieren – und wie Mechanismen der Kanonbildung zur Marginalisierung bestimmter Gruppen beigetragen haben. Eine feministische Relektüre der Philosophiegeschichte bedeutet daher nicht nur, bislang ausgeblendete Stimmen sichtbar zu machen, sondern auch die epistemischen Voraussetzungen philosophischer Sichtbarkeit grundsätzlich zu hinterfragen. Wer darf sprechen? Wer wird gehört? Und was gilt überhaupt als Philosophie?

4. Schlussfolgerung und Ausblick

Die Frauen des Wiener Kreises stehen exemplarisch für die strukturelle Marginalisierung von Philosophinnen in der analytischen Philosophie. Die Fälle von Rose Rand und Olga Hahn-Neurath zeigen, wie ihre Beiträge sowohl zeitgenössisch als auch historiografisch übergangen wurden. Eine kritische Revision der Philosophiegeschichtsschreibung ist daher notwendig, um weibliche Perspektiven systematisch zu integrieren.

Beaneys (2023) Konzept der „ajar-mindedness“ bietet hierfür einen produktiven Ansatz: Es fordert, vernachlässigte Themenfelder bewusst „einen Spalt weit“ zu öffnen – ohne dabei bestehende Strukturen vollständig aufzugeben. Dieser Perspektivwechsel darf jedoch nicht auf die historische Forschung beschränkt bleiben, sondern muss auch die gegenwärtige Gestaltung philosophischer Diskurse prägen – insbesondere in Schule und Hochschule. Es reicht nicht aus, Philosophinnen nur am Rande zu erwähnen; ihre Texte müssen aktiv in Seminare, Lektürekurse und Curricula integriert werden.

Im Fall des Wiener Kreises bedeutet das: Die Rolle der weiblichen Mitglieder sollte systematisch thematisiert und ihre Schriften eigenständig behandelt werden. Olga Hahn-Neuraths Arbeit „Zum Dualismus in der Logik“ eignet sich etwa für Einführungskurse zur Aussagenlogik oder Logikgeschichte. Auch Rose Rands Beiträge zur Metamathematik und Erkenntnistheorie sind relevante Quellen für Seminare zur Logik, Philosophie des 20. Jahrhunderts oder zur Exilforschung.

Eine solche Integration erlaubt es, Philosophiegeschichte als pluralen, dynamischen Prozess kollektiver Erkenntnisproduktion zu begreifen – mit bislang übersehenen Beiträgen. Damit verbunden ist eine epistemische Erweiterung im Sinne von Beaney (2023) „ajar-mindedness“ und Frickers (2023) Konzept epistemischer Gerechtigkeit.

Projekte wie *Bildersturm: Frauen in der Philosophie sichtbar machen und neue Vorbilder etablieren* an der Georg-August-Universität Göttingen oder die Arbeitsgruppe *Frauen in der Geschichte der Philosophie* der DGPhil zeigen, dass solche Erweiterungen des philosophischen Kanons sowohl möglich als auch notwendig sind. In diese Richtung weist auch eine virtuelle Ausstellung zur Geschichte des Logischen Empirismus, die derzeit am Institut Wiener Kreis entsteht: Sie verfolgt das Ziel, neben dem etablierten Kanon insbesondere auch die Rolle von Frauen in diesem philosophischen Umfeld sichtbar zu machen.

References

- Beaney, Michael (2023) “Open-mindedness and ajar-mindedness in history of philosophy”, *Metaphilosophy* 54, 208–222.
- Burke, Christopher; Sandner, Günther (2022) “Marie Reidemeister and Otto Neurath: interwoven lives and work”, *European Journal of Life Writing* 11, 103–129.
- Carnap, Rudolf (1931) „Überwindung der Metaphysik durch logische Analyse der Sprache“, *Erkenntnis* 2, 219–241.
- Carnap, Rudolf; Neurath, Otto (2024) *Briefwechsel*, hg. v. Christian Damböck, Johannes Friedl u. Ulf Höfer. Hamburg: Felix Meiner.
- Connell, Sophia M.; Janssen-Lauret Frederique (2022) “Lost voices: on counteracting exclusion of women from histories of contemporary philosophy”, *British Journal for the History of Philosophy* 30(2), 199–210.
- Connell, Sophia M.; Janssen-Lauret, Frederique (2023) ““Bad philosophy” and “derivative philosophy”: Labels that keep women out of the canon”, *Metaphilosophy* 54, 238–253.
- Creath, Richard (2023) “Logical Empiricism”, in: Edward N. Zalta & Uri Nodelman (eds.), *The Stanford Encyclopedia of Philosophy* (Winter 2023 Edition), Stanford University, <https://plato.stanford.edu/archives/win2023/entries/logical-empiricism/>
- Edmonds, David (2021) *Die Ermordung des Professor Schlick: Der Wiener Kreis und die dunklen Jahre der Philosophie*, Übers. v. Annabel Zettel. München: C.H. Beck.

Frank, Philipp (2016) *Das Kausalgesetz und seine Grenzen*, hg. v. Anne J. Kox. 2. Aufl. Frankfurt am Main: Suhrkamp.

Fricker, Miranda (2023) *Epidemische Ungerechtigkeit: Macht und die Ethik des Wissens*, Übers. v. Antje Korsmeier. München: C.H. Beck.

Haller, Rudolf (1985) „Der erste Wiener Kreis“, *Erkenntnis* 22(1/3), 341–358.

Janssen-Lauret, Frederique (2022) “Women in logical empiricism”, in: Christoph Limbeck-Lilienau and Thomas Uebel (eds.) *The Routledge Handbook of Logical Empiricism*, Abingdon: Routledge, 127–136.

Limbeck-Lilienau, Christoph (2022) “The first Vienna Circle and the Erlangen Conference”, in: Christoph Limbeck-Lilienau and Thomas Uebel (eds.) *The Routledge Handbook of Logical Empiricism*, Abingdon: Routledge, 99–108.

Longino, Helen E. (1990) *Science as Social Knowledge: Values and Objectivity in Scientific Inquiry*. Princeton: Princeton University Press.

Mihaljević, Katarina (2023) “Breaking into British academic life in Second World War Britain: The story of Rose Rand”, *HOPOS: The Journal of the International Society for the History of Philosophy of Science* 13, 297–316.

Reck, Erich H. (2022) “Logic and the foundations of mathematics in early logical empiricism”, in: Christoph Limbeck-Lilienau and Thomas Uebel (eds.) *The Routledge Handbook of Logical Empiricism*, Abingdon: Routledge, 139–147.

Rossiter, Margaret W. (2003) “Der Matthäus-Matilda-Effekt in der Wissenschaft”, in: Theresa Wobbe (Hg.) *Zwischen Vorderbühne und Hinterbühne: Beiträge zum Wandel der Geschlechterbeziehungen in der Wissenschaft vom 17. Jahrhundert bis zur Gegenwart*, Bielefeld: transcript Verlag, 191–210.

Stadler, Friedrich (2015) *Der Wiener Kreis: Ursprung, Entwicklung und Wirkung des Logischen Empirismus im Kontext*, 3. Aufl. Cham: Springer.

Verein Ernst Mach (2013) „Wissenschaftliche Weltauffassung. Der Wiener Kreis“, in: Christian Damböck (ed.) *Der Wiener Kreis: Ausgewählte Texte*, Stuttgart: Reclam, 7–32.

Acknowledgement

Ich bedanke mich bei meinem Dissertationsbetreuer, Mag. Dr. Christian Damböck, Privatdozent, für seine konstruktiven Anmerkungen.

Philosophy is deceitful above all things. On non-dualising speech, the duck-rabbit and JT LeRoy

Thomas Hainscho (Klagenfurt, Austria)

Abstract

The aim of this paper is not to present an argument of its own, but to pose questions. The questions arise from considerations that avoid what Josef Mitterer calls *dualising*: seeking alignment between linguistic statements and the world by referring to objects from a 'beyond of discourse'. The paper is intended to clarify which questions become relevant when this alignment is no longer sought. I address an argument that Mitterer develops in relation to Wittgenstein's discussion of aspect-seeing and apply it to the discourse about the US-American writer JT LeRoy and the hoax associated with him. When it is no longer asked, who JT LeRoy *really* is, then questions asked may pursue change rather than truth.

Introduction

The two books of Josef Mitterer, *Das Jenseits der Philosophie* (1992; *The Beyond of Philosophy*) and *Die Flucht aus der Beliebigkeit* (2001; *The Flight From Contingency*), offer a criticism of philosophy. Mitterer claims that different philosophical traditions share common, unquestioned presuppositions, and he critically reconstructs how philosophical questions have come about, shaped by the way that philosophy has been practiced thus far. His critical reconstruction is threefold, addressing: (1) the dichotomous differentiations at the heart of any philosophical school, such as the dichotomies between statement and object, appearance and reality, etc., (2) the orientation of philosophy towards the truth and (3) the idea of reference, i.e., the orientation of thought or speech to the object of thought or speech (Mitterer 2011a: 21–24, §10). Based on this reconstruction, Mitterer introduces the 'non-dualising speech', a way of dealing with philosophical questions that avoids these three commonalities.

The common aim for diverse philosophical traditions is to align what the philosophers of this tradition say with how the world is, as understood within the respective tradition. This presupposes a distinction between two realms: language on the one hand and the world on the other. For Mitterer, philosophers themselves introduced this distinction. In order to argue, to introduce truth as a regulative principle, the distinction between language and world was created, and thus Mitterer labels his reconstruction 'dualist philosophy'. His alternative avoids distinguishing between language and

world, it is not oriented towards the truth, and its direction of thinking is not referential. Therefore, Mitterer calls this way to philosophise ‘non-dualising speech’.

Accepting Mitterer’s criticism means to change what philosophy strives for. The pursuit of truth is superseded by a pursuit of change (Mitterer 2011: 86, §97). Questions about the relation between object and statement or certain claims about the truth of a proposition may be unmasked as dualist, in a similar manner as logical empiricism argues that the logical analysis of language can reveal certain statements as meaningless pseudo-statements. For Mitterer (2011: 83ff., §§88ff.), the dualist response to a question about the truth of a proposition is made by referring to what he calls the ‘beyond of the discourse’ – the true world, objective facts, etc. –, the extra-linguistic realm: “The ‘language-distinct’ reality is the beyond of discourse” (Mitterer 2011: 86, §99). Philosophers who belong to the same tradition, speak the same language and share the same standards may share common views about the beyond of discourse, but others may not. The incommensurability thesis and the anti-metaphysical stance of logical empiricism echo in Mitterer’s claims about the beyond of discourse.

Whereas logical empiricism presents linguistic philosophy as a research programme with a specific methodology and a promise of eliminating metaphysics, Mitterer does not do so. He emphasises that he does “not want to establish the Non-dualizing Speech [...], the ‘*pursuit of change*,’ as a new paradigm; indeed, one of my concerns is to *prevent the establishment of paradigms*” (Mitterer 2011: 13). An aim is achieved “when the dualist understands my construction as a *reconstruction*” (Mitterer 2011a: 35, §28). Non-dualising speech is also “developed only to the extent necessary to be able to reconstruct the dichotomous distinctions between language and reality, description and object, and their role in philosophical discourse” (Mitterer 2011a: 31, §22). This seems to suggest that the pursuit of change merely aims at reconstructing dualist philosophy.

Pursuing change instead of truth can be understood in at least two ways: First, as a critical endeavour that does not aim to produce positive statements but continuously draws attention to dualisations. This approach is similar to Wittgenstein’s claim in §6.53 of the *Tractatus*, regarding the “correct method in

philosophy”, namely “to say nothing except what can be said, [...] and then, whenever someone else wanted to say something metaphysical, to demonstrate to him that he had failed to give a meaning to certain signs in his propositions” (TLP §6.53), but raising one’s voice only when someone argues in a dualising way. Ernst von Glasersfeld (2008: 123), for example, proposes this view, writing that “Mitterer is the proponent of a [...] conceptual revision”; in the German translation, Glasersfeld suggests that Mitterer seeks to “ausrotten [...], was er ‘dualistische Argumentation’ nennt” (Glasersfeld 2011: 125). In this sense, non-dualising speech raises only one question: Is someone committing dualising acts?

A different understanding of the ‘pursuit of change’ is proposed by Sebastian Kletzel, who, drawing on Richard Rorty, suggests that such a pursuit might appeal to individuals he calls ‘edifying ironists’ – those who hope “that no absolutization of a discourse will ever be successful but that there is an ongoing growth of possibilities” (Kletzel 2017: 70) and embody this hope in their philosophical behaviour. Katharina Neges (2022: 113–114) compares dualising and non-dualising speech by analogy to two different languages: “If we speak in a non-dualizing way, we become unreceptive to ontological questions, they sound strange to us, and their meaning cannot be expressed precisely” (Neges 2022: 114). The change pursued is then a language-switch that renders certain expressions obsolete – i.e., philosophically uninteresting – and sparks interest in new ones. In this sense, non-dualising speech is a practise that does not pose a specific question but aims for change and openness.

However, Mitterer’s effort to make argumentation transparent instead of transcendent (Mitterer 2011a: 35, §28), which is given as a kind of programmatic slogan, and his claim that “the decision for a different philosophy is always also a decision for different problems” (Mitterer 2011: 121), suggest that different, perhaps even new, problems may arise when arguing non-dualistically. The following examination of Wittgenstein and the discourse surrounding JT LeRoy is an attempt to pose such different problems. It is intended to illustrate how the dualising consideration of a problem leads to familiar questions of canonised philosophy, and to offer an alternative.

Wittgenstein describes the duck-rabbit

Mitterer draws on three examples from section XI in part II of the *Philosophical Investigations*. In these examples, Wittgenstein examines cases in which an object is seen as something else: a triangle can be seen as a triangular hole, a mountain, an arrow, etc. (PI 2009: 210f.), children play a game in which “they say of a chest[,] [...] that it is now a house” (PI 2009: 217), and the duck-rabbit (PI 2009: 204ff.). In the context of the *Philosophical Investigations*, Wittgenstein uses these cases to discuss aspect-seeing and investigates a use of ‘to see’ that seems to be both visual experience and thought (PI 2009: 207). Noticing an aspect of something means cognitively acknowledging that something has not changed and yet seeing it differently (PI 2009: 203). This phenomenon can be observed, for example, in puzzle-pictures such as the duck-rabbit figure. Mitterer is interested in how Wittgenstein uses language to introduce the language games, i.e. how Wittgenstein describes the relation between the report of visual experience and the description of the visually experienced object.

For Mitterer (2011: 17, §6), “Wittgenstein defends the view that we ‘see an object according to its interpretation’”. Although this is not a literal quotation from Wittgenstein, it is in line with what can be found in the *Philosophical Investigations*. Examining the use of ‘to see’ in the case of the duck-rabbit, Wittgenstein notes that it would be odd to say, “Now I see it as a picture-rabbit” (PI 2009: 205), referring to both the visual experience (‘picture-rabbit’) and the perceived object (‘it’). He states that we use ‘to see’ just to describe what we see: “I say ‘It’s a rabbit’. Not ‘Now it’s a rabbit’” (PI 2009: 205), and later asks “how is it possible to see an object according to an *interpretation*?” (PI 2009: 211)

In Wittgenstein’s discussion of continuous seeing and dawning of an aspect, Mitterer (2011: 21, §6) points out that a sentence “involves a contradiction”. This concerns the second sentence in this passage: “The picture might have been shown me, without my ever seeing in it anything but a rabbit. [...] I may, then, have seen the duck-rabbit simply as a picture-rabbit from the first” (PI 2009: 204). For Mitterer, the statement about having seen a picture-rabbit ‘from the first’ makes it impossible to also mention the duck-rabbit. Either the

picture-rabbit had been seen or the duck-rabbit. Wittgenstein either describes a “visual (language) experience that did not occur” (Mitterer 2011: 21, §6), or contradicts his claim that an object is seen according to its interpretation.

Wittgenstein’s discussion of the duck-rabbit is given as an example for “the endeavor to arrive at interpretations, perceptions, descriptions, and observations of something, where this something is not itself supposed to be an interpretation, perception, description, or observation” (Mitterer 2011: 23, §6). It illustrates that discourses about interpretation, and even the claim that objects are seen according to their interpretation, presuppose that there are things without interpretation, i.e. things of a language-independent realm, from a beyond of discourse. This is not just a criticism of realism, but of the practise of referring to language-independent object when opinions differ, and of the belief to have the right standards for seeing things ‘how they really are’.

In case of conflicts, dualists argue and justify their claims by referring to the objects from a beyond of discourse. This reference is made possible by a dualist “*apriorization of the object*” (Mitterer 2011: 81, §80). A person’s own opinion, disguised as an objective measure, thus becomes the standard for judgement and comparison. Mitterer’s suggestion for avoiding this is the introduction of the notions ‘so far’ and ‘from now on’: “The description of the object constitutes/forms together with the object it describes a ‘new’ object of further description. Denoting the object of a description means denoting and prescribing a description *so far* that is supposed to be continued in a description *from now on* and to be changed into a new object for further description(s)” (Mitterer 2001: 46, §21). Instead of committing the dualising act of comparing linguistic statements with an alleged language-independent object from a beyond of discourse, non-dualising speech continues the discourse without reference to anything non-linguistic. Describing, comparing, agreeing, dissenting, objecting to statements, all occur within the discourse. Different descriptions repeatedly form and change objects of the discourse.

Modelling a non-dualistic discourse is challenging. Examples of how a description-from-now-on continues a description-so-far are often illustrative, such as continuing /a table/ or /an apple/ (Mitterer uses /execution marks/ to indicate that a description performed is an object of a description, cf. Mitterer 2011: 43, §14). I consider it an open question what problems arise when a

discourse is carried out in a non-dualising way. Are different problems likely to emerge? In the next section, I investigate this question by considering the case of JT LeRoy.

The case of JT LeRoy

JT LeRoy is the author of three books, the novel *Sarah* (2001), the short story collection *The Heart Is Deceitful Above All Things* (2001), the novella *Harold's End* (2005), and several shorter contributions published in literary anthologies and magazines in the late 1990s and early 2000s (Handy 2008). LeRoy writes about growing up with his drug-addicted mother, who worked as a prostitute at truck stops and cross-dressed him as a girl to attract men, selling him for sexual and physical abuse. His writings tell grim tales of trauma, rape and misery; yet they bear witness to survival and the power of literature in overcoming traumatic experience.

LeRoy only communicated by telephone and email, avoiding face-to-face contact with others. He did not make public appearances until 2001, when he gave his first readings in San Francisco, wearing sunglasses and a blonde wig. He then quickly became a cult figure, with several film and music celebrities among his fans. Only three years later, LeRoy promoted Asia Argento's adaptation of his second book at the 2004 Cannes Film Festival. (Beachy 2005; Handy 2008)

In 2006, *The New York Times* published the article *The Unmasking of JT Leroy: In Public, He's a She* (St. John 2006), stating that the literary works of LeRoy were written by Laura Albert, a 1965 born New York author and musician. While this had been claimed earlier (Beachy 2005), it remained unclear, who the person with the wig and sunglasses was, as LeRoy and Albert had been seen together. St. John (2006) revealed that this person was Savannah Knoop, Albert's partner's sister, who dressed up as LeRoy while Albert handled communication via email and phone. St. John's claims were confirmed to be true.

'The Unmasking of JT Leroy' turned out to be both a scoop for the public and a shock to the many (including famous) friends who had trusted LeRoy. Above all, it created confusion. Even after the reveal, director Gus Van Sant, a close friend LeRoy, is quoted as saying, "I still kind of believe that he exists, just not in the flesh" (Handy 2008). In an effort to resolve the confusion, questions

about ‘the real JT LeRoy’ were raised, about ‘the truth’ and even about ‘the real truth’ as, for example, Loontjens (2008) notes: “LeRoy’s hoax shows that the desire to know the real truth suddenly becomes very pressing when what was believed to be ‘real’ reveals itself as pure fiction.”

There are at least three possible answers to question who ‘the real JT LeRoy’ is:

(1) The most common answer is that there is simply no real JT LeRoy. This position was used in court in 2007, when Laura Albert was sued for fraud: “A film production company has sued Ms. Albert, saying that a contract signed with JT Leroy for the rights to make a feature film of ‘Sarah’ should be null and void, for the simple reason that JT Leroy does not exist” (Feuer 2007). The production company won the case.

(2) During the trial, Albert stated that LeRoy was a coping mechanism and that he was real for her: “Albert told jurors [...] that she had been assuming male identities for decades as a coping mechanism for psychological problems brought on by her sexual abuse as a child. To her, she said, LeRoy was real” (Westfeldt 2007). She claimed that, “It was more like two computer programs running in my head. There was him, and there was me” (Feuer 2007a). For Albert, LeRoy fulfilled a psychological function. He allowed her to say things that she would not have said otherwise, similar as Falco for Hans Hölzel or Hannah Montana for Miley Cyrus.

Savannah Knoop, who made the public appearances as LeRoy in disguise, offers a similar account. In their memoir *Girl Boy Girl* (2008), Knoop writes about being LeRoy, noting “I feel like it’s taking over my life” (Knoop 2008: 155), and reflects on their different body perception while being LeRoy (Knoop 2008: 202). After the unmasking, Knoop stated that LeRoy had served as a shortcut for expressing their queer identity (Crummy 2019).

(3) Scholarly works have examined LeRoy in the context of postmodern theories of authorship and gender studies: Caroline Hamilton acknowledges that LeRoy is a hoax but claims that “[a]uthors are made, not born”, and uses LeRoy as an example for her claim that an author is “a collective identity, formed by writers, readers and the publishing industry” (Hamilton 2010: 37). Loontjens (2008) claims that “JT LeRoy still is the functional author”, even after the unmasking. In this sense, Pil and Galia Kollektiv investigate the textual and

physical manifestations of the body of JT LeRoy, not denying his existence, but claiming that he ‘did not exist properly’ (“Despite never having properly existed, JT Leroy had a lung”; Kollektiv and Kollektiv 2018: 254).

It is clear that the three answers are all plausible yet mutually incompatible, and this may now be considered as a result: A philosophical consideration reveals that multiple answers are possible, that none of them is true, and this settles the question about ‘the real JT LeRoy’. This assumption can be juxtaposed with Mitterer’s examination of the duck-rabbit. Wittgenstein’s investigation of aspect-seeing presupposes that the figure is a duck-rabbit. The picture-rabbit and the picture-duck are then introduced as aspects of the figure that had been described as what it is (a duck-rabbit). The three answers to the question of who ‘the real JT LeRoy’ is can be seen as interpretations (i.e., aspects) of a state of affairs (i.e., the literary hoax): Although we know that LeRoy is a hoax, we can still interpret him in various ways. Interpreting him, for example, as a postmodern author presupposes that he is a hoax. Claiming that the philosophical meta-position explains the complicated matter in its totality, or ‘as it is’, uses relativistic measures to explain how lawyers, literary scholars or philosophers perceive the matter, allowing each group to assert that LeRoy does (not) exist in a particular way. This approach is following the same path as assuming the answer (1), (2) or (3) to be true. Assuming this philosophical consideration to be true would be, I am borrowing the title from one of LeRoy’s books, ‘deceitful above all things’.

A case for non-dualism could be made by considering a statement from LeRoy (it is Albert speaking as LeRoy, who does not and cannot speak for himself). He was more concerned with avoiding the establishment of an identity than committing to one, especially with regard to his gender: “When I wrote *Sarah*, I was male-identified, and now I’m not. I don’t know what I am. [...] So many people have claimed me as their own, so I guess the best thing is to confuse them all.” (Press 2001: 34)

Several papers explore the relations between Butler’s ideas about gender and non-dualising speech (Derra 2008; Weiss 2013; Cyzman-Eid 2023). Aleksandra Derra, for example, writes that Butler’s performative conception of the subject is opposed to the idea of the subject “as an entity having a certain foundation” (Derra 2008: 208f.), claiming that feminist philosophy and politics

can proceed without an ontological commitment to what a woman is or a consensus on the nature of female subjectivity. With a performative approach to identity, a description-from-now-on changes the established description-so-far, for example that of /female subjectivity/, or of /JT LeRoy/. LeRoy embodies a refusal to be pinned down, seeking to be interpreted through his performance – as text (Albert), as appearance (Knoop) –, being aware of the confusion.

Knoop's memoir was adapted into a film, entitled *JT LeRoy* (2018). In one scene, LeRoy (played by actress Kristen Stewart, who plays Savannah Knoop, who plays LeRoy) is asked whether his story is real, whether he actually wrote the books and whether he is who he says that he is, and s/he responds with a statement that is paradigmatic of non-dualising speech: "JT's whoever you want him to be". There is no foundation of JT LeRoy that somehow determines who he is. Being whoever someone wants him to be shifts the responsibility for how he is to be described to those who actually describe him. For Mitterer, this is the case with every object. What we say is not static, but in flux. What is called 'the real x', is how people believe x to be described.

Talking about /JT LeRoy/ continuously alters what we are talking about. According to Neges (2022), our convictions, i.e. commonly shared moral beliefs, determine how we talk about (moral) problems, and which (moral) problems we talk about. A change of our moral convictions goes along with changing ourselves: We become another person by choosing differently. Neges argues that certain conflicts should not be carried out over an object, but over the convictions that are used to describe that object (Neges 2022: 192–204). However, I may also retain my convictions while altering the description /LeRoy is performance/ to /LeRoy is not a performance but an author/. Nevertheless, this case addresses convictions, or rather standards, and does raise questions: Is it necessary to adopt different standards in order to understand a different position? Do different standards lead to incommensurable descriptions? Can descriptions change standards? How do standards change? Do new interpretations require new standards? Such questions can only be asked when it is accepted that the things we talk about are not permanent, but in flux.

References

- Beachy, Stephen (2005) "Who Is the Real JT Leroy? A Search for the True Identity of the Literary Hustler", *New York*, 17 October 2005, Online version (7 October 2005): <http://nymag.com/nymetro/news/people/features/14718/index4.html>.
- Crummy, Colin (2019) "The real J.T. LeRoy", *The Face*, 17 April 2019, <https://theface.com/culture/the-real-j-t-leroy>.
- Cyzman-Eid, Marzenna (2023) "Nie-dualizująco o płci. Uwagi wstępne", *Białostockie Studia Literaturoznawcze* (23), 7–16.
- Derra, Aleksandra (2008) "The non-dualizing way of speaking and the female subjectivity problem", *Constructivist Foundations* 3(3), 208–213.
- Feuer, Alan (2007) "At Trial, Writer Recalls an Alter Ego That Took Over", *The New York Times*, 21 June 2007, B1. Online version: <https://www.nytimes.com/2007/06/21/nyregion/21writer.html>.
- Feuer, Alan (2007a) "Her Journey, All True", *The New York Times*, 23 August 2007, E1. Online version: <https://www.nytimes.com/2007/08/23/arts/23albert.html>.
- Glaserfeld, Ernst von (2008) "Can dichotomies be tamed?", *Constructivist Foundations* 3(3), 123–126.
- Glaserfeld, Ernst von (2011) "Sind Dichotomien zähmbar?", in: Alexander Riegler and Stefan Weber (eds.) *Die Dritte Philosophie. Kritische Beiträge zu Josef Mitterers Non-Dualismus*, Weilerswist: Velbrück Wissenschaft, 125–132.
- Hamilton, Caroline (2010) "Passing for an author: The strange case of J. T. LeRoy", *Humanities Research* 16(1), 35–55.
- Handy, Bruce (2006) "The Boy Who Cried Author", *Vanity Fair*. Online version (1 January 2007): <https://www.vanityfair.com/news/2006/05/jtleroy200605>.
- Kletzl, Sebastian (2017) "Who Wants to Be a Non-Dualist and Why?", in: Christian Kanzian, Sebastian Kletzl, Josef Mitterer and Katharina Neges (eds.) *Realism – Relativism – Constructivism* (Publications of the Austrian Ludwig Wittgenstein Society, New Series, Vol. 24), Berlin/Boston: De Gruyter, 59–72.
- Knoop, Savannah (2008) *Girl Boy Girl. How I Became JT LeRoy*, New York: Seven Stories Press.
- Kollektiv, Pil; Kollektiv, Galia (2018) "The Missing Body of JT Leroy", in: Emmy Beber (ed.) *The bodies that remain*, Santa Barbara: Punctum Books, 251–262.
- Loontjens, Jannah (2008) "Resisting the author: JT LeRoy's fictional authorship", *Image & Narrative* 9(22).

Mitterer, Josef (2011) *Das Jenseits der Philosophie. Wider das dualistische Erkenntnisprinzip*, mit einem Nachwort zur Neuausgabe 2011, Weilerswist: Velbrück Wissenschaft. English translation available at: <https://www.3-16am.co.uk/articles/.c/josef-mitterer-s-beyond-philosophy-serialised>.

Mitterer, Josef (2011a) *Die Flucht aus der Beliebigkeit*, mit einem Vorwort zur Neuausgabe 2011, Weilerswist: Velbrück Wissenschaft. English translation available at: <https://www.3-16am.co.uk/articles/.c/josef-mitterer-s-beyond-philosophy-serialised>.

Neges, Katharina (2022) *The Realm of Morality. Lectures on Metaethics* (Dissertation, University of Klagenfurt).

Press, Joy (2001) "The Cult of J.T. LeRoy", *The Village Voice* 46(24), 33–35.

St. John, Warren (2006) "The Unmasking of JT Leroy: In Public, Hes a She", *The New York Times*, 9 January 2006, B1. Online version: <https://www.nytimes.com/2006/01/09/books/the-unmasking-of-jt-leroy-in-public-hes-a-she.html>.

Weiss, Martin G. (2013) "Non-dualistic Sex. Josef Mitterer's Non-dualistic Philosophy in the Light of Judith Butler's (De)Constructivist Feminism", *Constructivist Foundations* 8(2), 183–189.

Westfeldt, Amy (2007) "LeRoy found to be fraud", *Los Angeles Times*, 25 June 2007. Online version: <https://www.latimes.com/archives/la-xpm-2007-jun-25-et-literarylawsuit25-story.html>.

Wittgenstein, Ludwig (2001) *Tractatus Logico-Philosophicus*, translated by D. F. Pears und B. F. McGuinness, with an introduction by Bertrand Russell. London: Routledge.

Wittgenstein, Ludwig (2009) *Philosophische Untersuchungen / Philosophical Investigations*, revised fourth edition by P. M. S. Hacker and Joachim Schulte, translated by G. E. M. Anscombe, P. M. S. Hacker and Joachim Schulte. New York: Wiley-Blackwell.

Acknowledgement

I would like to thank Stephen Beachy for his help in identifying literature sources on JT LeRoy.

"A Separation from the Original Soil": Wittgenstein on Anthropological Explanation in the "Remarks on Frazer's *Golden Bough*"

Kyle Hardy (Toronto, Canada)

Abstract

Wittgenstein ends the first section of his "Remarks on Frazer's *Golden Bough*"—the section comprising loosely-connected reflections written in a manuscript throughout the summer of 1931—with a discussion of the origins of human religion. Throughout the early "Remarks" Wittgenstein offers a commentary on the methods of pioneering social anthropologist James George Frazer. He concludes this discussion with the mystifying remark that "the awakening of the intellect occurs with a separation from the original soil, the original basis of life. (The origin of choice.) (The form of the awakening spirit is veneration.)." In this essay I propose one way of interpreting the closely-related notions of separation and veneration at play in this quote. To do so, I sketch Wittgenstein's view that human nature is distinguished by a ritual instinct. This ritual instinct is the grounds for ritual actions, both those we venerate and those we do not. Veneration then, I argue, is a function of our lack of proximity to a given practice. We do not venerate those things which we take to be an unremarkable part of our lives. With these themes in mind, I conclude by pointing towards an underexplored dimension of the "Remarks": their critical evaluation of which practices we take as requiring explanation, and which we do not.

1. Wittgenstein's Account of Ritualistic Actions

Throughout the first part of the "Remarks on Frazer's *Golden Bough*," Wittgenstein criticizes Frazer for making the magical and religious practices he examines seem like errors (GB 1993: 121) because they do not actually effect any result. Frazer suggests that a rainmaking ritual is an attempt to cause rain to fall, but its practitioners only believe it is successful because "an incantation that is supposed to bring rain certainly seems efficacious sooner or later" (GB 1993: 121). Wittgenstein deems Frazer's view psychologically unrealistic. The fact that they did not "do [these practices] in the dry periods of the year in which the land is 'a parched and arid desert'" (GB 1993: 137) evidences that they did not believe in a causal connection between their practices and the coming of the rain. Furthermore, Wittgenstein wonders why they would have begun the rainmaking practice at all if they believed it caused rain, since before they instituted the practice they had already experienced that the rains would come at a certain time of year (GB 1993: 137). The question of the origin of a ritualistic practice is one to which I will return.

Wittgenstein suggests we instead ought to see magical practices as "ceremonial" or "ritualistic" actions. They are not means to an end which they might or might not accomplish. He states this claim first in tentative terms, writing that "one could almost say that man is a ceremonial animal. That is, no doubt, partly wrong and partly nonsensical, but there is also something right about it" (GB 1993: 129). He goes on to suggest that: "one could begin a book on anthropology by saying: When one examines the life and behaviour of mankind throughout the world, one sees that, except for what might be called animal activities ... men also perform actions which bear a characteristic peculiar to themselves, and these could be called ritualistic actions" (GB 1993: 129). Wittgenstein encourages us to think that the human form of life is distinguished by its inclusion of ritualistic actions. By implication, animals do not have ritualistic actions. By the same token, the various forms of human life across the world have the fact that they perform ritualistic action in common.

Wittgenstein motivates this claim by suggesting that ritualistic actions are also held in common by the various forms human life has taken across time. They play a role in his own life, and by analogy other modern lives, just as they do in the lives of ancient humans. As he writes: "when I am furious about something, I sometimes beat the ground or a tree with my walking stick. But I certainly do not believe that the ground is to blame or that my beating can help anything. 'I am venting my anger'. And all rites are of this kind. Such actions may be called Instinct-actions" (GB 1993: 137). The significance of the example is that Wittgenstein, a person with modern knowledge, obviously does not strike an object with his walking stick because he believes some restitution or positive effect will come by punishing that object. But were one to interpret this behavior in the way Frazer interprets the ancient practices he examines, they would assume Wittgenstein has some erroneous causal belief. Wittgenstein therefore encourages us to think that practices, both ancient and modern, are responses to a human ritualistic instinct. These actions don't aim at some effect: "we just behave this way and then we feel satisfied" (GB 1993: 123).

2. Ritualistic Instinct and Human Nature

Ritualistic actions are, Wittgenstein tells us, necessarily associated with an instinct distinctive to—and common among—human beings. This is an account of one element of what might be called human nature, or the human form of life.

The emphasis I place on Wittgenstein's account of ritualistic actions as related to a view of human nature is inspired by a thought of Frank Cioffi's, in his 1981 paper "Wittgenstein and the Fire-festivals." Cioffi focuses on the later-written "Remarks," where Wittgenstein examines Frazer's comments on the Beltane Fire Festival. Cioffi disambiguates Wittgenstein's criticism of Frazer on this practice into two different theses (Cioffi 1981: 213). The first thesis is that the Beltane Fire Festival is "intelligible as [it] stand[s]," (Cioffi 1981: 213) and therefore an account of its development from some earlier practice is not required. The second thesis revolves around Frazer's failure to see that the feeling that these practices are "deep" and "sinister," and thus needing to be explained, is because of an inner human impulse. If we moderns want to understand why such a festival strikes us as deep and sinister, Cioffi suggests, we need an account of "the experience in ourselves from which we impute" this "deep and sinister character" (Cioffi 1981: 213).

While there is much correct about Peter K. Westergaard's suggestion that the later "Remarks" have "significant differences ... in terms of style, approach, and thematic content" (Westergaard 2016: 292) from the earlier, a version of the division Cioffi identifies is already clearly present in the early "Remarks." I propose to align many of the statements in the early "Remarks" along the lines of two loose arguments: 1. Wittgenstein's criticism of Frazer's tendency to posit developmental relationships between the practices he studies, and 2. Wittgenstein's claim that ritual actions are associated with a human ritualistic instinct.

The difference between these two arguments is exemplified in the difference between Wittgenstein's remark that "And so the chorus points to a secret law' one feels like saying to Frazer's collection of facts. I can represent this law, this idea, by means of an evolutionary hypothesis ... but also by means of the arrangement of its factual content alone, in a 'perspicuous' representation" (GB 1993: 133) and his claim that "one could almost say that

man is a ceremonial animal. That is, no doubt, partly wrong and partly nonsensical, but there is also something right about it." (GB 1993: 129). The former remark is a methodological criticism: Wittgenstein suggests that Frazer's evolutionary hypothesis is only one way of representing the facts collected in *The Golden Bough*, and need not be the correct way of representing them. By contrast, the second is a positive claim about human nature.

3. Ritualistic Instinct and Perspicuous Representation

One role which this dimension of human nature, the ritualistic instinct, plays in the "Remarks" is as part of Wittgenstein's critique of Frazer's methodology. P.M.S. Hacker's 1992 paper "Developmental Hypotheses and Perspicuous Representations: Wittgenstein on Frazer's '*Golden Bough*'" is instructive on this point.

Frazer posits that the magical practices he examines develop into one another, on the basis of similarities between their features. As Frazer writes in the preface to the 1922 edition, the edition which Wittgenstein read during the composition of the early "Remarks," "the rule of succession to the priesthood of Diana at Aricia"—a ritual foremost in importance among those which Frazer proposes to explain in *The Golden Bough*—"clearly exemplifies a widespread institution, of which the most numerous and the most similar cases have thus far been found in Africa. How far the facts point to an early influence of Africa on Italy or even to the existence of an African population in Southern Europe, I do not presume to say" (Frazer 1922: vi-vii). He draws this historical inference on the basis of formal similarities between the rituals involved.

Wittgenstein's association of ritualistic actions with a common human ritualistic instinct provides a way of accounting for the formal similarities Frazer examines, without resorting to a specious historical inference. This is how Wittgenstein's notion of perspicuous representation comes into play. A perspicuous representation is here a way of looking at the data which does not mislead one into drawing illicit connections. As Hacker writes, "In the case of Frazer's anthropological data ... [one can represent] the pattern of formal similarities and analogies ... by means of a developmental hypothesis (as Frazer did) or ... one can also represent the 'law' by means of an arrangement

of the data alone in a perspicuous representation" (Hacker 1992: 294). Wittgenstein's analysis of the cases of ritualistic actions discussed by Frazer is such a perspicuous representation. It is a way of looking at the data which does not mislead one into drawing illicit connections. Frazer draws illicit connections by positing a historical development between disparate practices.

Hacker therefore explains the significance of Wittgenstein's account of human nature to the "Remarks" as a "point of reference for rendering intelligible the meaning of magical rites of primitive societies" (Hacker 1992: 298). Humans share the ritualistic instinct, which is the foundation for ritualistic actions. Without this point of reference, hermeneutic investigation into ritual actions has a tendency to be confused. Frazer, in his claim that magical practices are mistaken attempts at understanding and influencing the world, runs into this very problem.

4. Ritualistic Instinct and Anthropological Explanation

While Hacker is correct in thinking that awareness of the ritualistic instinct Wittgenstein identifies enables a perspicuous representation of Frazer's data, this does not exhaust the role the ritualistic instinct plays in the "Remarks." Wittgenstein's account of the ritualistic instinct points beyond that argument, towards a critical evaluation of which practices we take to require anthropological explanation, and which we do not.

Towards the end of the early "Remarks," Wittgenstein writes that "it was not a trivial reason, for really there can be no reason, that prompted certain races of mankind to venerate the oak tree, but only the fact that they and the oak were united in a community of life" (GB 1993: 139). This moment marks his return to the question of the origin of ritual practices. There is no reason for the role of oak trees in certain religious practices, because nothing caused people to start venerating oak trees. The relation between tree and human has existed from time immemorial. But this relation isn't reducible to proximity between humans and oak trees. Puzzlingly, Wittgenstein says that "it was not their union (the oak and man) that has given rise to these rites, but in a certain sense their separation. For the awakening of the intellect occurs with a separation from the original soil, the original basis of life" (GB 1993: 139). In this comment Wittgenstein offers something like a history of ritual actions, albeit one which does not seek a causal explanation for the actions themselves.

At the earliest point of human ritual practice—as I take Wittgenstein to mean by “the awakening of the intellect”—primitive humanity was separated from the conditions in which they had been formed.

Glossing his comment about the separation between oak trees and humans, Wittgenstein suggests that the “form of the awakening spirit is veneration” (GB 1993: 139). Wittgenstein is here referring to some point in deep antiquity where we first became recognizably human. The origin of ritual practices is associated with the origin of human beings as such. At some point in history, humans ceased being just another animal, and began being something distinctively human. This moment was occasioned by their beginning to engage in ritual practices, which (as we have already seen) Wittgenstein takes to be distinctively human. But why does the beginning of ritual practice, and the feeling of veneration, involve a separation from the original basis of life?

To make the answer to this question explicit, it is helpful to draw on a few of the later “Remarks.” There, Wittgenstein tells us that “if [ancient people] were to write it down, their knowledge of nature would not differ fundamentally from ours. Only their magic is different” (GB 1993: 141). By drawing a distinction between “their” magic and “ours,” Wittgenstein implies that there is still some distinction between modern ritual practices (as in his hitting a tree with his walking stick) and ancient ritual practices like those Frazer examines. This is intuitively true: we do not regard the modern practices Wittgenstein has in mind as “dreadful, magnificent, horrible, tragic, etc.” (GB 1993: 121) the way we do accounts of ancient ritual practices. Nor do we regard those modern practices as standing in need of an anthropological explanation. They are simply familiar parts of our ordinary lives. We do not feel the need to ask why Wittgenstein would strike a tree with his walking stick when frustrated, we immediately understand his action by considering our own experience of frustration.

However, there are practices which continue into the modern world which we do regard as dreadful, magnificent, horrible, tragic, etc.: those which we believe to be the last vestiges of genuinely ancient practices. Wittgenstein contends that when a ritual practice strikes us as deep and sinister, we often imagine this is because it originates in the distant past. We are inclined to feel that “through this aspect, it suddenly gains depth” (GB 1993: 145). Our

tendency to impute ancient provenance to these practices (even where that tendency is not consistent with history) is something which Wittgenstein wants to valorize. As he writes: "Where do we get the certainty that such a practice must be ancient ... might we not be mistaken and convicted of our mistake by history? Certainly, but then there still remains something of which we are certain. We would then say: 'Good, the origin may be different in this case, but in general it is surely prehistoric'" (GB 1993: 147). Our intuition that a practice is prehistoric, and thus that it has depth, cannot be dispelled by the discovery that the practice is not actually ancient. Furthermore, Wittgenstein suggests that our feeling that it must be ancient, regardless of its real origins, is not at all misplaced.

Our perception that a practice must be ancient speaks to what really makes a practice strike us as deep and serious: our ritualistic instinct. With this insight, we can return to the early "Remarks," and to the question of the awakening human spirit. The impression that a practice is dreadful, magnificent, horrible, tragic, etc.—that it has depth—maps well onto the notion of veneration, and for Wittgenstein they seem to be synonymous. Ancient humanity's feelings of veneration must not be related to the real history of a practice, for a people with no history (as their human spirit has only just been awakened) must equally feel the depth of the practices in which they take part. The dawning of wonder surrounding these familiar things signals "a separation from the original soil, the original basis of life" (GB 1993: 139) because we see those natural things which sustain our lives as worthy of veneration. That which could not be more familiar suddenly seems dreadful, magnificent, horrible, tragic, etc.

Wittgenstein takes it that we seek to explain those practices which we find dreadful, magnificent, horrible, tragic, etc. As he writes, "Frazer begins by telling us the story of the King of the Wood of Nemi, he does this in a tone which shows that he feels, and wants us to feel, that something strange and dreadful is happening" (GB 1993: 121). It is at this point, Wittgenstein suggests, that we are inspired to ask "why does this [practice] happen?" (GB 1993: 121). But there are many ritual actions with which we are intimately familiar. We do not find them strange and dreadful, nor do we feel the need to ask why they happen. If they are all ritual actions, why shouldn't we feel the need to explain them all?

The "Remarks" includes a view which goes beyond arguing that hermeneutic investigation into ritual practices will be misled if they are not attentive to the ritualistic instinct in humans. Wittgenstein tells us that we do not take a practice to require explanation because it is a ritual practice: there are many ritual practices in our own lives which do not seem to stand in need of anthropological explanation. We might think that some practices require explanation because of how strange they are. To the contrary, Wittgenstein writes: "no phenomenon is in itself particularly mysterious, but any of them can become so to us" (GB 1993: 129). What makes a practice mysterious to us does not inhere in that practice. It is rather a function of our lack of proximity. Our feeling that some practice requires explanation is related to our feeling that it is deep and sinister. Our feeling that a practice is deep and sinister is related to its distance from the practices of our ordinary life.

References

Cioffi, Frank (1981) "Wittgenstein and the Fire-festivals", in: Irving Block (ed.) *Perspectives on the Philosophy of Wittgenstein*, Cambridge: The MIT Press, 212-237.

Frazer, James George (1922) *The Golden Bough: A Study in Magic and Religion*, Abbreviated edition, London: MacMillan Publishing Company.

Hacker, P. M. S. (1992) "Developmental Hypotheses and Perspicuous Representations: Wittgenstein on Frazer's 'Golden Bough'", *Iyyun: The Jerusalem Philosophical Quarterly* 41, 277-299.

Westergaard, Peter K. (2016) "Thirteen Loose Sheets of Varying Size: On Part II of *Bemerkungen über Frazer's 'The Golden Bough'*", in: Lars Albinus, Josef G. F. Rothhaupt and Aidan Seery (eds.) *Wittgenstein's Remarks on Frazer: The Text and the Matter*, Berlin and Boston: Walter de Gruyter, 291-310.

Wittgenstein, Ludwig (1993) "Remarks on Frazer's *Golden Bough*", in: James C. Klagge and Alfred Nordmann (eds.) *Philosophical Occasions: 1912-1951*, Indianapolis and Cambridge: Hackett Publishing Company, 115-155.

Acknowledgement

This essay draws on research supported by the Social Sciences and Humanities Research Council of Canada.

The Window and the Soul: Wittgenstein, Woolf, and the Ethics of Recognition

Iain Haslam (London, United Kingdom)

Abstract

This paper examines Wittgenstein's notion of "an attitude towards a soul" as an alternative to standard accounts of interpersonal recognition in philosophy. Recent humanist approaches to social injustice emphasise dehumanisation as a key concept, yet face a paradox: perpetrators of inhumanity often seem to recognise their victims' humanity in classificatory terms. Drawing on Aleksy Tarasenko-Struc's work, I argue that Wittgenstein's approach allows us to reconceive interpersonal recognition as a form of engaged attention rather than mere classification. This reconception provides resources for understanding both the breakdown of recognition in dehumanisation and its successful realisation in interpersonal relations and the understanding of another individual. To illustrate this approach, I turn to Martha Nussbaum's analysis of Virginia Woolf's *To the Lighthouse*. In this novel, we find a rich literary exploration of both the failures and successes of interpersonal recognition through the contrast between Lily Briscoe's failed attempt to know Mrs. Ramsay and the Ramsays' mutual, loving understanding. This approach has significant implications for feminist humanism, suggesting that dehumanisation is not just a matter of false beliefs about others' capacities but a failure of engaged attention—a failure in one's attitude towards their soul.

Telling someone something he does not understand is pointless, even if you add that he will not be able to understand it. (That so often happens with someone you love.)

(CV, 1998: 10)

Introduction

Recent work in feminist philosophy has witnessed an increased focus on conceptualising questions of social and political injustice in terms of *dehumanisation*. Mari Mikkola, to use one example, argues that we should “stop taking the gender concept *woman* as the organizing notion of feminist philosophy and reframe feminist analysis of injustice in humanist terms” (2016: 4).

So what does it mean to see – or fail to see – someone's humanity? The standard answer treats such interpersonal recognition as essentially a cognitive achievement, a matter of correctly classifying others as possessing the psychological capacities characteristic of persons. But this picture faces significant challenges, most notably that perpetrators of inhumanity often seem to recognise their victims as human in some important sense. If this is so,

it suggests that dehumanisation cannot play the explanatory role often attributed to it.

In this paper, I propose that Wittgenstein's notion of "an attitude towards a soul" (*eine Einstellung zur Seele*) offers a promising alternative to this standard picture. Building on the work of Aleksy Tarasenko-Struc (2020, 2024), I will argue that Wittgenstein's approach allows us to reconceive interpersonal recognition as a form of engaged attention rather than mere classification. This reconception, I suggest, provides resources for understanding both the breakdown of recognition in cases of dehumanisation and its successful realisation in interpersonal relations, particularly in loving relationships.

To illustrate the explanatory power of this Wittgensteinian approach, I will turn to Martha Nussbaum's analysis of Virginia Woolf's *To the Lighthouse*, first presented at this very Symposium some thirty years ago. Woolf's novel, as read by Nussbaum, offers a rich literary exploration of both the failures and successes of interpersonal recognition, thereby providing an illuminating insight into what is at stake in this discussion.

The Paradoxes of Dehumanisation

The dominant way of thinking about interpersonal recognition treats it as a species of belief or classification. On this model, recognising someone as a human being consists in classifying them as possessing the psychological capacities characteristic of persons, such as rationality, autonomy, and emotion. Dehumanisation, correspondingly, occurs when we fail to classify others correctly, when we do not attribute to them these capacities.

This picture, however, faces significant challenges. As Manne (e.g. 2016, 2018) and others have argued, perpetrators of inhumanity often *do* seem to recognise their victims as human in the classificatory sense. They may attribute to their victims distinctively human mental states and may even be motivated by attitudes – like contempt, resentment, or indignation – that presuppose the humanity of their targets. As Manne puts it, perpetrators of misogynistic violence often do "not deny women's power, independence, or the reality of their minds. Rather, [they] hate and seek to punish them for evincing these capacities in ways that frustrated [them]" (2018: 146).

This has led some critics to deny that dehumanisation, in the sense of failing to see others as human, plays a significant role in explaining inhumane treatment. If perpetrators of inhumanity recognise their victims' humanity, then it cannot be their failure to see this humanity that explains their inhumane treatment.

Tarasenko-Struc (2020, 2024) has offered an insightful diagnosis of this paradox. The problem, he suggests, lies in the overly intellectualised conception of interpersonal recognition that underlies the classification model. This model treats seeing someone as human as a purely cognitive achievement, a matter of forming the right kind of belief about them. But this fails to capture the distinctively engaged and affective character of interpersonal recognition.

The upshot, according to Tarasenko-Struc, is that we need a richer conception of interpersonal recognition – one that can account for both the classification of others and the engaged, affective dimension of our relations with them. It is here that Wittgenstein's notion of an "attitude towards a soul" becomes relevant.

Wittgenstein's "Attitude Towards a Soul"

In a famous passage from the *Investigations*, Wittgenstein writes:

"19. "I believe that he is suffering." — Do I also believe that he isn't an automaton?

Only reluctantly could I use the word in both contexts.

(Or is it like *this*: I believe that he is suffering, but am certain that he is not an automaton? Nonsense!)

20. Suppose I say of a friend: "He isn't an automaton." — What information is conveyed by this, and to whom would it be information? To a human being who meets him in ordinary circumstances? What information could it give him? (At the very most, that this man always behaves like a human being, and not occasionally like a machine.)

21. "I believe that he is not an automaton", just like that, so far makes no sense.

22. My attitude towards him is an attitude towards a soul. I am not of the *opinion* that he has a soul. (PI II iv).

In these sections, Wittgenstein draws our attention to an important distinction. While we can speak of believing or having an opinion about certain empirical matters concerning others – for instance, that they are suffering – there is something odd about saying we merely “believe” or “have the opinion” that they are not automata. Our basic stance toward others as minded beings is not well characterised as a belief or opinion. Rather, it is an “attitude towards a soul” – a distinctive stance that cannot be reduced to any set of beliefs about them.

This stance reflects a different way of seeing others – not as objects to be categorised but as subjects with whom we stand in relation. It involves seeing their behaviour not as mere movements but as expressions of a mental life. As Wittgenstein puts it (characteristically pithily) elsewhere, “If I see someone writhing in pain with evident cause, I do not think: all the same, his feelings are hidden from me” (PI II xi §324). The point is not that I infer their pain from their behaviour, but that I see their behaviour as itself an expression of pain.

What is involved in this understanding of others is a sort of ‘pattern recognition’. Wittgenstein depicts life as a “weave” or “tapestry” containing patterns that “vary in a multiplicity of ways” and are “interwoven” with each other (see Z §§56-89). When we say that someone is irritated, or that they feel grief or joy, we are identifying patterns in the “tapestry of their life” (*Lebensteppich*) or the “weave of their life” (*Band des Lebens*).

“Grief describes a pattern which recurs, with different variations, in the tapestry of life.” (PI II i §2)

This pattern recognition is not a matter of inference from behaviour to hidden mental states. It is a matter of seeing the behaviour itself as expressing mental states. We might think of it as a form of direct perception rather than inference.

Building on these Wittgensteinian insights, Tarasenko-Struc (2020) develops what he calls the “engagement model” of interpersonal recognition. On this model, seeing someone as a person is fundamentally a matter of being

emotionally and practically engaged with them – of being susceptible to such sentiments as sympathy and the Strawsonian reactive attitudes (love, gratitude, resentment) toward them. It is the “overall, ground-floor orientation of the emotions and will vis-à-vis the recognized party’s perspective” (2020: 16).

The key difference between the classification model and the engagement model, then, lies in their respective conceptions of the role of affect in interpersonal recognition. On the classification model, affect plays at most an auxiliary role – it may accompany our recognition of others but is not constitutive of it. On the engagement model, by contrast, affect is constitutive of interpersonal recognition. To recognise someone as a person is, in part, to be emotionally engaged with them in certain ways.

Resolving the Paradoxes

With this Wittgensteinian model of interpersonal recognition in hand, we can now turn to the question of why dehumanisation remains a significant phenomenon despite the objections raised by Manne and others. The engagement model suggests that perpetrators of inhumanity may classify their victims as human in the cognitive sense while still failing to engage with them in the ways characteristic of interpersonal recognition. They may not be emotionally and practically responsive to them as subjects.

This is not to deny that perpetrators of inhumanity sometimes do attribute distinctively human mental states to their victims or that they may be motivated by attitudes that presuppose the humanity of their targets. But these attributions and attitudes do not necessarily involve the kind of engaged recognition that the Wittgensteinian approach highlights. One can attribute mental states to another while still failing to adopt an attitude towards a soul with respect to them – failing to be emotionally and practically responsive to them as subjects.

Moreover, the engagement model also helps us understand the positive counterpart to dehumanisation: successful interpersonal recognition. If dehumanisation involves a failure to adopt an attitude towards a soul with respect to others, then successful recognition involves precisely the adoption of such an attitude. It involves a form of engaged attention and concern that makes us responsive to others as subjects rather than mere objects of knowledge.

This gives us a richer way of thinking about what it means to recognise another's humanity or personhood. It is not just a matter of believing certain things about them – that they have such-and-such capacities or qualities. It is a matter of attending to them in certain ways, of being responsive to them as subjects rather than objects. And this form of recognition is exemplified in various forms of human relationship, from friendship to love, where we are particularly attuned and responsive to others as subjects.

It is in literature, perhaps, that we find the most vivid and compelling portrayal of both the failures and successes of interpersonal recognition. So to draw out this comparison, I will now turn to Virginia Woolf's novel *To the Lighthouse*, as analysed by Martha Nussbaum.

Literary Expression: To the Lighthouse

In her essay "The Window" (1995), originally presented at this symposium in 1994, Martha Nussbaum offers a reading of Woolf's novel as an exploration of the problem of other minds. The novel, she suggests, dramatises both the failures and successes of interpersonal recognition in a way that illuminates the Wittgensteinian approach I have been developing.

The novel centres on the Ramsay family and their guests at a summer house on the Isle of Skye. Among the guests is Lily Briscoe, an artist who is fascinated by Mrs. Ramsay, the mother of the family. Lily wonders, "How, then, did one know one thing or another thing about people, sealed as they were?" She imagines knowledge of Mrs. Ramsay as a kind of access to a hidden inner realm, "like the treasures in the tombs of kings, tablets bearing sacred inscriptions, which if one could spell them out, would teach one everything, but they would never be offered openly, never made public" (Woolf, quoted in Nussbaum 1995: 741-742).

But Lily's attempt to understand Mrs. Ramsay fails. As Nussbaum notes, "The hives are sealed" (1995: 731). Lily cannot gain access to the inner chambers of Mrs. Ramsay's mind and heart. And this failure is not just a contingent one – it is not just that Mrs. Ramsay happens to be particularly reticent or that Lily happens to be particularly unperceptive. It is a failure that stems from the very conception of interpersonal knowledge that Lily is working with. She conceives of understanding others as a matter of penetrating their hidden inner realms, and this conception inevitably leads to frustration.

In contrast to Lily's failed attempt at understanding, Woolf presents the relationship between Mr. and Mrs. Ramsay. Their understanding of each other is not based on prying open each other's inner chambers. Rather, it is based on a kind of attentiveness and responsiveness that respects the other's separateness. As Nussbaum puts it, Mr. Ramsay "attends to [his wife] more fully as a person separate from himself existing in her own right, rather than as an instrument of consolation for himself" (1995: 744).

We see this, for instance, in Mr. Ramsay's sensitivity to his wife's need for solitude:

"He turned and saw her. Ah! She was lovely, lovelier now than ever he thought. But he could not speak to her. He could not interrupt her. He wanted urgently to speak to her now that James was gone and she was alone at last. But he resolved, no; he would not interrupt her. She was aloof from him now in her beauty, in her sadness. He would let her be, and he passed her without a word, though it hurt him that she should look so distant, and he could not reach her, he could do nothing to help her."

What we see here is not an attempt to penetrate Mrs. Ramsay's inner chambers but a willingness to respect her separateness, to "let her be." Mr. Ramsay's understanding of his wife is expressed precisely in his recognition of the limits of understanding her – that she exists as a separate person with her own inner life. And far from being a limitation, this recognition is a form of respect and concern for her as a subject.

This form of understanding is beautifully captured in the climactic scene at the end of Part I, where Mrs. Ramsay silently communicates her love to her husband:

"Then, knowing that he was watching her, instead of saying anything she turned, holding her stocking, and looked at him. And as she looked at him she began to smile, for though she had not said a word, he knew, of course he knew, that she loved him... She had not said it: yet he knew."

Mr. Ramsay knows that his wife loves him, not through probing or questioning, but simply by looking. Nussbaum characterises this kind of understanding as a form of "reading":

“Having lived together for a long time, they have gathered a lot of information about patterns of speech, action, reaction. Among other things, they have learned a lot—partly by making mistakes with one another—about the limits of analogizing, about relevant similarities and differences. They have gathered this information, furthermore, not in the manner of a detached scientist, but in the course of interactions to which both ascribe enormous importance... Each learns the idiosyncratic text of the other in the way that one might learn a foreign language—never having a once-for-all guaranteed translation manual, but holistically piecing it all together, trying to make the best sense, over time, of all the words and phrases.” (Nussbaum 1995: 745)

What Nussbaum describes here chimes nicely with Wittgenstein’s account of how we come to understand others through pattern recognition in the tapestry of life. The Ramsay’s have learned to read the patterns in each other’s lives, to see how their behaviour expresses their mental states. This reading is not a matter of inference from behaviour to hidden mental states. It is a matter of seeing the behaviour itself as expressive *of* mental states.

Crucially, this reading is not a matter of dispassionate observation. It is a matter of engaged attention and response. The Ramsay’s do not merely observe each other’s behaviour; they are emotionally and practically responsive to it. They care about each other and are concerned with each other’s well-being. Their understanding is bound up with their love for each other.

Thus, I propose that the contrast between Lily’s approach and the Ramsay’s mutual understanding illustrates the difference between the classification model and the engagement model of interpersonal recognition. Lily treats Mrs. Ramsay as an object to be known, as containing hidden inner contents that she wants to access. The Ramsays, by contrast, treat each other as subjects with whom they stand in a relationship, as persons whose separateness they respect and whose expressions they are attuned to. They do not merely classify each other; they are attuned and responsive to each other as subjects. They have, in sum, correctly attuned attitudes towards each other’s souls.

Conclusion

I have argued that Wittgenstein's notion of an attitude towards a soul offers a fruitful alternative to standard accounts of interpersonal recognition. On this view, recognising someone as a person is fundamentally a matter of affective and engaged attention, not dispassionate classification. This approach, I have suggested, helps us understand both the breakdown of recognition in cases of dehumanisation and its successful realisation in interpersonal relations, particularly in love.

This reframing offers resources for feminist critique that are not available on the classification model. It allows us to identify forms of dehumanisation that involve not the denial of human capacities but the refusal to engage with others as subjects. It helps explain how it is possible for someone to classify another as human while still failing to treat them with basic moral consideration.

At the same time, this approach offers a positive vision of interpersonal recognition that goes beyond mere classification. It suggests that to recognise someone as a person is to adopt a specific stance toward them – one characterised by engaged attention and concern. This stance is exemplified in the successful mutual recognition of the Ramsays, and more generally in love.

The recognition of others' humanity is not just a theoretical matter but a deeply practical one. It is a matter of how we attend to and engage with others in our everyday interactions. In the end, what Wittgenstein's notion of an attitude towards a soul offers is a way of thinking about interpersonal recognition that is both more realistic psychologically and more ethically rich than the classification model. It acknowledges the complexity of our relations with others and the irreducible role of affect and engagement in those relations. It reminds us that to recognise another's humanity is not just to classify them correctly but to be attuned and responsive to them as subjects – to have what Wittgenstein called an attitude towards a soul.

References

- Manne, K. (2016) 'Humanism: A Critique', *Social Theory and Practice*, 42(2), pp. 389–415.
- Manne, K. (2018) *Down girl: the logic of misogyny*. New York, NY: Oxford University Press.

Mikkola, M. (2016) *The Wrong of Injustice: Dehumanization and its Role in Feminist Philosophy*. Oxford University Press.

Nussbaum, M.C. (1995) 'The Window: Knowledge of Other Minds in Virginia Woolf's *To The Lighthouse*', *New Literary History*, 26(4), pp. 731–753.

Tarasenko-Struc, A. (2024) 'Humanism: A Reconsideration', *Journal of the American Philosophical Association*, 10(3), pp. 542–561.

Tarasenko-Struc, A. (2020) 'Interpersonal Invisibility and the Recognition of Other Persons', in D. Kaspar (ed.) *Explorations in Ethics*. Cham: Springer International Publishing, pp. 219–242.

Tarasenko-Struc, A. (forthcoming) 'An Attitude Towards a Soul—and its Corruptions: A Wittgensteinian View of Racial Alienation', in J. Beale and R. Cosker-Rowland (eds) *Wittgenstein and Contemporary Moral Philosophy*.

Wittgenstein, L. (2009) *Philosophical investigations*. Rev. 4th ed. Translated by G.E.M. Anscombe, P.M.S. Hacker, and J. Schulte. Chichester, West Sussex, U.K.; Malden, MA: Wiley-Blackwell.

Woolf, Virginia (1955). *To the Lighthouse*. New York: Harcourt, Brace & Company.

Nonbinary Identity and Social Meaning: A Reply to Kadji Amin

Matthew Hoffman (Bloomington, Indiana, United States)

Abstract

Kadji Amin argues we ought to reject ‘nonbinary’ as a category, because it arises from a neoliberal, internalist understanding of the self. Amin rightly criticizes gender internalism, but his conclusion depends on failing to recognize gender as a normative system. His view of social categories misses the evaluative elements that characterize gender’s normative sociality. Contra Amin, we should understand someone who identifies as nonbinary to be rejecting, not “binary gender” as an identity category, but the gender binary as a normative structure. The “announcement” of one’s gender that Amin criticizes is always social, since it involves rejecting social positions one is supposed to inhabit. The problem with a “private gender” internalist view of gender identity is not that it is bad politically, as Amin suggests, but that it is a false picture of the relation between our concepts and our mental states. Since no one could adopt an identity internally without relating to external criteria, that cannot be what someone is doing in identifying as nonbinary.

Kadji Amin argues that ‘nonbinary’ is an empty category, depending on historical accident and an “autological” understanding of gender (Amin 2022a). While I agree with much of his historical overview and his criticisms of how gender identity is understood under neoliberalism, I find Amin’s conclusions less agreeable. His most provocative conclusion – that in rejecting neoliberal idealization and an autological notion of gender we ought to reject ‘nonbinary’ as an identity category – depends on failing to come to terms with how gender is a normative system. Amin clearly understands gender to be a social category, but his view of social categories is missing the evaluative and regulative elements that characterize gender’s normative sociality.

My starting place is, following Naomi Scheman, to recognize that nonbinary lives “are lived, hence livable,” and so an analysis of nonbinary identity must begin by seeking “the conditions of that possibility” (Scheman 2011: 120). In understanding gender as normative, we can understand identifying as nonbinary to be a social activity, not a meaningless psychic act. Contra Amin, identifying as nonbinary is rejecting not “binary gender” as an identity category, but the gender binary as a normative structure.

I. Voluntarism About Gender

Amin argues that *divergence*, *binarism*, *idealization*, and *autology* are the logics driving new gender categories, categories that depend only on inner identification rather than outward expression. His analysis draws these logics

together as follows. Identity terms are idealized, in that one's feelings about falling under a category are taken to require a lack of ambivalence. Amin takes terms like 'cisgender' and 'heterosexual', for example, to be idealized in that they are assumed to be "uncontaminated by any gender trouble whatsoever." This idealization leads to the construction of binary opposites for any category. That is, the logic of neoliberalism is that everyone belongs to an identity category; and that to be a member of that category, one must feel unambivalent about one's group membership. Thus, 'heterosexual' and 'cisgender' become identities had by those who feel "perfectly comfortable in their sexed bodies and gendered social roles" (Amin 2022a: 109), as the opposite identity categories from homosexual and transgender. Likewise, Amin argues, nonbinary discourse has created 'binary gender' as its "fictive opposite" (Amin 2022a: 114).

Self-identification is what puts one into one of these identity categories, rather than gender performance or anything external. Amin sees this as the "apotheosis of the liberal Western fantasy of self-determining 'autological' selfhood" (Amin 2022a: 116). In particular, Amin is critical of the way he understands nonbinary identity to have no determining grounds outside of identification, which leaves nonbinary identity "being claimed by people who look and behave in a manner indistinguishable from ordinary lesbians and gays, or even ordinary heterosexuals" (Amin 2022a: 113). Since one might look any number of ways and still identify as nonbinary (Amin 2022a: 114), it is empty as a gender category, a fiction of neoliberal identity politics.

I agree with Amin regarding what (following Naomi Scheman), I will refer to as voluntarism, the view that we are whatever we say we are, whatever we take gender terms to mean (Scheman 2011: 129-130). I agree that this is a bad way to understand gender, and that it looks reasonable only because of our current social/political moment. The idealization of gender categories in neoliberalism functions to encourage misunderstanding of the relation between the self and social categories. However, Amin draws the wrong conclusions from this. He criticizes people identifying as nonbinary because he sees nothing to gender identity beyond announcement of one's self-determined gender.

But voluntarism is not simply a politically questionable way to understand gender, as Amin suggests by his criticism. Rather, it is a false picture of the relation between our mental states and concepts. Hilde Lindemann Nelson draws a connection between Butler's project and Wittgenstein's discussion of the possibility of a private language, one whose "words refer to what can be known only to the person speaking, so another person cannot understand" (Nelson 2002: 217). Just as with other concepts, there is content to gender concepts only because they exist within a set of social practices (Scheman 2011: 129). Gender performance is never "fully self-styled" because it requires "criteria for application whose correctness can be determined independently of the person wielding them" (Nelson 2002: 219).

Amin appears to recognize this at certain points, as his criticism of 'gender identity' as a "phantasmic and labile relation to [gender] categories" (Amin 2022a: 115) takes the form of distinguishing the identification he criticizes from social identity. But there can be no identification in his sense, something wholly individual and private. Since no one could adopt an identity purely psychically without relating to external criteria, there will always be some relation between gender identity and our social world. Voluntaristic announcement *could not* be what someone is doing in identifying as nonbinary. So identification is "neither capricious nor unconnected to cultural meaning" in the first place (Scheman 2011: 130). In "emptying out the fictive core of gender" (Amin 2022a: 106), we need to look elsewhere to understand the relation between the self and gender categories.

II. Gender as a Normative System

Gender categories are constructions, but as social categories they do not merely group people together, they regulate behavior, setting standards for how someone is supposed to behave and punishing those who behave differently. In what follows, I use Elizabeth Barnes's social position account of gender to provide the theoretical backing for rejecting Amin's claims about nonbinary identity. But I am not endorsing Barnes's account, necessarily. I think whatever analysis of gender one prefers, there will be similar claims to be made about how identification as nonbinary has social meaning.

On Barnes's analysis, gender has many aspects (social role, gender identity, expression etc.), but these are "ultimately explained by a hierarchical social

structure” (Barnes 2020: 716). Bodies are coded as male or female, assigned roles as a result, and those roles coded as feminine are systematically devalued (Barnes 2020: 715). Being masculinized or feminized, in Barnes’s terms, is being understood to occupy some social position on the basis of presumed biological features, and being privileged or subordinated on this basis (Barnes 2020: 714-715).

Barnes also thinks there are other social positions, which she labels as “gender outliers” and “gender confounders” in which one’s perceived physical characteristics and social role come apart, where this explains why one is subordinated. One’s social position is dependent on context, so someone who is a gender outlier in one context (for example, trans people who don’t pass as cis, as well as gender nonconforming cis people) might in other contexts be feminized or masculinized. Likewise, someone who in certain contexts is masculinized or feminized might be a gender confounder in other contexts where they cannot be identified as having characteristics associated with either male or female role, regardless of their gender identity (Barnes 2020: 717).

While being masculinized or feminized is central to her account of gender roles, gender identity and gender expression are responses to this. Thus, while the social division is binary, gender identity is not. Everyone stands in some relation to gender ideology, and one element of that ideology is that gender is binary. So an account of gender needs to track the social significance of presuming biological sex, while having a broader view of how we respond to being masculinized or feminized (Barnes 2020: 715). Our use of gender terms, therefore, needs to be flexible, in order to combat this binary logic and gendered oppression based on social role (Barnes 2020: 723).

Amin wants to go on without ‘gender identity.’ He urges us to create social categories with positive content (unlike, as he sees it, ‘nonbinary’), “socially legible gender categories – based on presentation and behavior, not self-identification alone” (Amin 2022a: 116). While Amin’s gender categories involve social position, this does not mean that Amin has a social position theory of gender, in which social expectations and differing treatment of the basis of social categories explains gender. Amin’s categories seemingly depend for their coherence on having distinct gender performances. Yet, Amin’s

examples, such as “fairies and queens, butches and ‘he-shes,’ hermaphrodites and sexual intermediaries” (Amin 2022a: 111) and “feminine male-assigned people” (Amin 2022a: 117), involve individuals who fall under those categories on the basis of assumed sexual characteristics. In other words, being (or not being) antecedently masculinized or feminized determines which categories are open to one. Amin’s social categories thus cannot fail to have content, since that content is partially determined by binary gender norms.

III. Felt Ambivalence and Dominant Social Categories

Amin contends that that ‘heterosexuality’, ‘cisgender’, and ‘binary gender identity’ are fictive identity categories created by neoliberal categorization. No one feels unambivalent about their assigned gender role, he claims, so these categories do not have meaning any more than nonbinary does (Amin 2022a: 114). He concludes that we should reject all these identity categories and instead create more legible social categories with positive content. However, in rejecting these terms as identities created by a process of idealizing binary categories, we lose the ability to criticize gender as a normative structure which constricts our lives.

While Amin says that nonbinary “could seemingly apply to almost anyone” (Amin 2022a: 113), not everyone (not even every queer person whose gender performance reflects ambivalence or discomfort towards their assigned gender and gender role) does identify as nonbinary, and this is not because they have not heard the term. Amin recognizes this, writing “Feeling comfortable and ‘at home’ in straight culture is more powerfully predictive of heterosexual identification than is an exclusive desire for the ‘opposite’ sex” (Amin 2022a: 111). We should not ignore this when considering how to think about the gender binary and about nonbinary as a social category.

We need the terms ‘heterosexual,’ ‘cisgender,’ and ‘binary gender’ not as identity terms, but as terms by which we can talk about those people for whom dominant social practices are home. We need these terms not because some people have no homosexual desire, or because they identify with binary gender as such, but because most people identify, however inchoately, with a system of oppressive gender norms. Most people, however much ambivalence they feel about gender roles and expectations, are at home in the gender binary as a social system. That felt ambivalence, which Amin points to as

evidence that 'binary gender' is an idealization, is real. Yet that ambivalence does not destabilize most people's understanding of gender. It does not lead them to think about gender as a social system rather than as a natural part of the self, far less to think of gender as a system with which they could contest. It is true that there is no feeling of being 'binary'. Amin concludes from this that there is no nonbinary identity, either. But chasing a feeling by which we could alone determine gender identity is a mistake. In losing sight of gender's social normativity, Amin fails to recognize that, while people do not identify as 'binary', they do identify as men and women, and do so in ways that take for granted dominant gender ideology.

IV. Identifying to Others is Social

Amin sees no way of understanding gender identity talk other than as voluntaristic. However, on any view on which gender is a system of social positions that systematically privilege or subordinate people on the basis of various assumed biological factors and social roles, self-ascriptions can make sense non-voluntaristically. Identifying as nonbinary, rather than being a meaningless act of self-declaration, must be understood in relation to the social positions on which we are (whether we like it or not) assessed. Rather than being an interior, mental act, identifying as nonbinary is itself a social activity, relating the self to social expectations. The "announcement" that Amin criticizes (Amin 2022: 115) is always relational, since it involves rejecting social positions that one is supposed to inhabit (rejecting, in Barnes's terms, being masculinized or feminized). Rather than an announcement of an identity voluntaristically taken on, identifying as nonbinary can be understood as the refusal of letting others view you (as they might have) as inhabiting one of two intelligible social roles. 'Nonbinary', then, is what Robin Dembroff terms a *critical gender kind*, a category whose members "destabilize one or more core elements of the dominant gender ideology in that society" (Dembroff 2020: 2). Dembroff includes more internalist content to their proposal than either Amin or I would support. Nevertheless, Dembroff builds externalist features into their view as well, and that is enough to use it to criticize Amin's position.

In targeting voluntarism, Amin fails to read the claim "There is no one way to be nonbinary" (Talusán 2017) in the way I suspect it is intended. Amin takes there to be no content to being nonbinary because he thinks discourse around gender identity is making the claim that there is no external content necessary

to being nonbinary (and since he has rejected a private, internalist picture of gender). But he concludes this from there being no single gender performance for the social identity ‘nonbinary’. This conclusion is mistaken. External content is necessary, although there is no *necessary* external content (that is, no *specific* external content is necessary, and so there is no one way to be nonbinary, no one gender performance that will count one as nonbinary). Thus we find, on Dembroff’s view, “Many forms of resisting binary categorization are possible” (Dembroff 2020: 17). For example, “using gender neutral pronouns (and other terms, like the title ‘Mx.’), cultivating gender non-conforming aesthetics, asserting nonbinary categorization (e.g., ‘I am agender’), queering personal relationships, defying sexual binaries, and... moving between male and female gendered spaces” (Dembroff 2020: 17). We see this, I think, even in the examples Amin uses to argue that nonbinary is an empty category. Rather than there being nothing external to nonbinary identity, Amin’s own citations show at least one common feature: identifying to other people as nonbinary.

Amin differentiates identification – the purely psychic, voluntaristic act – from transitioning, a socially relevant shift in social gender categories (Amin 2022a: 115). Yet, even as he cites popular articles as support for the claim that identifying as nonbinary needs no external expression, those same articles involve discussion of how interviewees navigate social expectation and “external expression in choice of dress, hairstyle, [and] pronouns” (Amin 2022a: 114). Similarly, in discussing his students who use they/them pronouns, Amin says he does not see them as expressing transmasculinity, transfemininity, or genderqueerness (Amin 2022b). His students identify as nonbinary, but they are not performing these distinct social roles. For Amin, this is more evidence that nonbinary has no positive social content. This only follows, however, if the social content of nonbinary identity *must be* legible to Amin himself.

Amin’s criticism is that modern discourse has created an identity that finds no outward expression. Yet, even as these popular articles suggest that there is no specific way to be nonbinary, all the people interviewed have socially transitioned, in the sense of changing social position. They are identifying to others as nonbinary. That is, they are all out, recognized as being what Barnes would name as gender outliers or confounders (and thus potentially subject to

subordination). For this reason, the examples Amin discusses are not cases of voluntarism at all. Amin is not, therefore, discussing the fiction of a private gender, but talking about people who identify to others as nonbinary. Identifying to others is not internal, but external, and so it is not the voluntaristic sense of “identification” that Amin would target.

Amin’s criticism is also that there is no *shared* gender expression among people who identify as nonbinary. But that is because there is no social position ‘nonbinary’ in the way that there are on Barnes’s view a feminized and a masculinized social position. That is because gender is a normative system, in which the only social positions so defined are ‘man’ and ‘woman.’ Because of this, there is no particular gender expression by which one can be perceived as occupying the nonbinary social role. So, it is true that there is nothing in particular by which all and only people who self-identify as nonbinary will be regularly identified by others as nonbinary. Nevertheless, changes in gender expression and social position *are* social changes, not purely mental acts of identification.

The examples Amin discusses are examples of people who identify to others as nonbinary. It is false, therefore, that their gender identity is “autonomous from the social, to the extent that it may entirely contradict one’s actual gender performances” and that their gender identity “references a core selfhood that requires no expression” (Amin 2022a: 116). If there were no outward expression, no one could judge that a person’s gender identity contradicted that (nonexistent) performance. On the other hand, if one’s expression (here, at least telling others things like “I’m nonbinary” or “I use they/them pronouns”) is sufficient in one’s community to be recognized as inhabiting the social category ‘nonbinary’, one’s identity is not autonomous from the social. In the cases under discussion, there is an outward expression of identity. And on a variety of accounts, there will be no “contradiction” here between gender performance and gender identity, because there is a legible social category, *nonbinary*. Amin might not like self-identification being sufficient for that social role categorization, but that is not up to him to decide. People who identify as nonbinary are transitioning from one social category to another, even if that new social category is not understood or liked by Amin himself.

Thus, despite Amin's claims, 'nonbinary' has just as much potential to be a real social category as Amin's preferred categories. 'Nonbinary' has meaning whether or not Amin himself understands certain social identities better than others. But, to deliberately mis-quote Wittgenstein (PI 2009: 224), if Amin finds certain things about gender puzzling, it is because he does not find the whole business of gender queer enough.

References

Amin, Kadji (2022a) "We Are All Nonbinary: A History of Accidents", *Representations* 158, 106-119.

Amin, Kadji (2022b) "We Are All Nonbinary: A History of Accidents." *Center for the Study of Gender and Sexuality at the University of Chicago (YouTube Channel)*, <https://www.youtube.com/watch?v=BLigPSnvqSc>.

Barnes, Elizabeth (2020) "Gender and Gender Terms", *Noûs* 54, 704-730.

Dembroff, Robin (2020) "Beyond Binary: Genderqueer As Critical Gender Kind", *Philosophers' Imprint* 20, 1-23.

Nelson, Hilde Lindemann (2002) "Wittgenstein Meets 'Woman' in the Language-Game of Theorizing Feminism", in Naomi Scheman and Peg O'Connor (eds.) *Feminist Interpretations of Ludwig Wittgenstein*, University Park: Pennsylvania State University Press, 213-234.

Scheman, Naomi (2011) "Queering the Center by Centering the Queer", in *Shifting Ground: Knowledge and Reality, Transgression and Trustworthiness*, Oxford: Oxford University Press, 111-144.

Talusan, Meredith (2017) "This Is What Gender-Nonbinary People Look Like", *them*, <https://www.them.us/story/this-is-what-gender-nonbinary-people-look-like>.

Wittgenstein, Ludwig (2009): *Philosophical Investigations / Philosophische Untersuchungen*. Hg. von P. M. S. Hacker und Joachim Schulte, übersetzt von G. E. M. Anscombe, P. M. S. Hacker und Joachim Schulte. New York: Wiley.

The Limits of Certainty: Distinguishing Moral Certainties from Empirical Certainties

Minoo Hojjat (Tehran, Iran)

Abstract

In his later works, Wittgenstein does not explicitly discuss ethics. However, we may still attempt to apply his method to the moral domain. Some philosophers, drawing on *On Certainty*, argue that moral certainties function similarly to the empirical certainties Wittgenstein examines. I aim to show that, while moral propositions may serve a comparable role, they differ in several key ways from the certainties Wittgenstein describes. These differences can be summarized as follows: (a) unlike moral certainties, empirical certainties are embedded in our instinctive actions, (b) empirical certainties form the foundation of a *general* language-game, whereas moral certainties ground *specific* language-games, (c) ethical language-games are plural, built upon diverse foundational beliefs, (d) empirical certainties are typically expressed as singular propositions, whereas moral certainties take the form of general propositions—leading to further distinctions: (e) moral rules are often ambiguous, subject to exceptions, and open to interpretation, which creates room for disagreement and dispute, (f) singular moral judgments (i.e., general rules when applied to specific cases) are intertwined with judgments of other kinds (e.g., metaphysical, empirical, psychological, and sociological), making them more complex. Furthermore, (g) unlike epistemic skepticism, moral skepticism allows for consistent action—that is, moral skeptics can violate moral rules.

Due to these fundamental differences between moral and empirical certainty, Wittgenstein's insights in *On Certainty* cannot be directly applied to ethics.

Introduction

Although Wittgenstein did not speak much about ethics in his later philosophy, given the importance of his special philosophical method, we might hope to draw on some of his insights into ethics. In *On Certainty*, Wittgenstein discusses propositions that, in his view, are neither knowable nor open to doubt; they merely express our way of acting and, in this sense, are neither reasonable nor unreasonable, neither questionable nor justifiable. At the same time, these beliefs form the foundation of all inquiry and human research.

Some have claimed that there are similar beliefs in ethics. Among them is Russell Goodman, who believes in objective moral certainty and sees all of Wittgenstein's claims about certainty in *On Certainty* as equally applicable to moral certainties (Goodman 1982: 144-145). Similarly, Nigel Pleasants argues for basic moral certainty and attempts to show that moral certainty plays the same fundamental role in our ethical practices and judgments as empirical certainty does in our epistemic practices and judgments (Pleasants 2009: 669).

I do not intend to examine such claims here; rather, I aim to highlight differences between moral and empirical certainties that prevent the former from being classified among the certainties Wittgenstein discusses in *On Certainty*. I will point out key differences between these two types of certainty to show that, while basic certainties in ethics play a similar role, they differ in important ways from the certainties Wittgenstein discusses.

1. Empirical Certainties as Instinctive Actions

One of the key points Wittgenstein emphasizes in *On Certainty* is that "justifying the evidence comes to an end, but the end ... is not a kind of seeing on our part; it is our acting which lies at the bottom of the language-game" (OC: §204); "it is an ungrounded way of acting." (OC: §110). From the very beginning of life, children start acting—or more precisely, reacting (OC: §538)—and in the process of learning how to act (e.g., fetching books, sitting in armchairs, etc.), they swallow certain beliefs (OC: §143). It is as if certainty in these beliefs is embedded in all human instinctive actions. For example, when I get up from a chair and walk, I do not think about having two legs, but this belief is implicit in my action. Our instinctive actions reveal propositions that take the form of empirical statements, but they do not seem to reveal moral norms. However, Julia Hermann argues that moral certainties are, to some extent, instinctive:

the acting underlying our moral language-games can be understood in at least three different ways: as involving primitive reactions to pain or signs of distress, which are non-verbal and instinctive and can also be found among non-human animals; as the natural reactions of children and caregivers within the process of moral teaching and learning (for example a parent's expression of approval and affection in response to good behavior, the cries of another child which has been hurt, the child's reaction to those reactions, and so on); as the immediate responses of morally competent agents. On the third interpretation, it includes spontaneous emotional reactions as well as the exercise of linguistic competence, both of which are animal in the sense that they do not require conscious thought. (Hermann 2015: 112)

But none of these interpretations seems to make this similarity convincing. First, one should not assume that characterizing moral certainty as animalistic helps establish its instinctive nature. Unlike belief in the external world—

which animals, in a sense, share with us—moral belief is not animalistic. No one considers animals to possess a form of ethics. Thus, even reactions like feeling compassion for another being are not considered *moral* reactions. Feeling compassion is different from having a moral sense or feeling obligated to perform an action. Therefore, we cannot say that moral certainties, like empirical certainties, are observable in a child's behavior from the very beginning of life. This creates a significant distinction between certainty in these two realms.

At the same time, the encouragement of an action by educators is not necessarily a moral reaction either. An action may simply be pleasing to an educator, but finding someone's action pleasing is different from a moral reaction. It may be possible to establish a relationship between human inclinations and morality, but these inclinations themselves do not constitute morality. For instance, we all prefer to be told the truth, but this preference, in and of itself, is not a moral reaction.

Another example Hermann provides for basic moral behavior is the immediate reaction of a morally competent agent. However, it should be noted that certain teachings that are deeply internalized are also sometimes exhibited in an individual's behavior spontaneously; so mere spontaneity does not indicate intrinsicness. Furthermore, Hermann considers the reactions of specific agents, whereas for a behavior to be instinctive, it must be a universal reaction among all agents. Despite moral certainties, empirical certainties are embedded in all our actions, from those of a newborn just entering the world to those of all rational individuals. This key distinction connects to another major difference between moral certainty and the certainty discussed in *On Certainty*.

2. General and Specific Language-Games

According to Wittgenstein's remarks in *On Certainty*, human beings participate in various language-games throughout their lives, each involving its own foundational and certain beliefs. However, it seems that some language-games are more general than others. For example, the language-game of science is more general than that of applied techniques, in the sense that anyone engaging in technical practices may first need to have participated in the game of science, since applied techniques involve using scientific results in various

fields. Given this assumption, it appears that some language-games are universal—those in which almost all human beings participate. These are the games that make other games possible. Learning and using words in a language is of this kind. Many examples Wittgenstein mentions pertain to the human language-game, or more precisely, to the language-game that belongs to those whom we consider "rational" beings (see OC: 220, 223, 252, 254, 324). His discussion in *On Certainty* concerns certainties that are "beyond any reasonable doubt" (OC: §§416, 607).

The language-game of ethics becomes possible *within* this broader game, rather than existing alongside it. Empirical certainties are far more foundational than moral certainties because, without empirical certainties, ethical propositions would be neither comprehensible nor applicable. Long before anyone contemplates killing a person and the norms related to it, they must first have learned the meaning of "human" and "killing." Even making a judgment about such matters depends on how these words are understood. For a normative claim about killing to be made, numerous factual beliefs must already be established: that "humans exist," that "a certain action leads to death," that "a certain action can be performed," and many other beliefs about reality.

This distinction appears to align with the difference Stephan Rummens draws between *basic* certainties and *absolutely basic* certainties. He identifies seven characteristics of hinge propositions of both types, but the key distinction between these two types of certainty is that, in the case of *basic* certainties, these characteristics are confined to a particular language-game (Rummens 2013, 134, 141–143). For example, while absolutely basic certainties serve as the precondition for *any* doubt, inquiry, or reflection, basic certainties are only the precondition for doubt, inquiry, or reflection *within* a particular language-game.

So, when a language-game is the most general one in which all other games take place, there is no discourse outside or beyond it where its certainties can be questioned. However, when it comes to moral certainties, even if they may not be questioned *within* the language-game of ethics itself, they *can* be challenged in other discourses.

At the same time, the ethical language-game is not a singular, unified game; rather, in ethics, we encounter multiple language-games.

3. The Plurality of Ethical Language-Games

Wittgenstein frequently speaks about the possibility of different world-pictures, but what is discussed in *On Certainty* is a *human* world-picture. As mentioned in the previous section, this world-picture belongs to all rational human beings. He states that, "the reasonable man does not have certain doubts" (OC: §220) and that "every reasonable person" believes in this way (OC: §252). However, moral certainties are not equally certain for everyone. Michael Kober argues that even within a single community, rather than one moral language-game, there can be "a family of moral practices overlapping and criss-crossing one another" (Kober 1997: 374). Naturally, what differentiate these games are the basic beliefs they involve. The impact of these differences can be observed in disagreements over moral judgments.

Of course, the plurality of moral language-games itself arises from various other characteristics of these games, which I will address in the following sections.

4. The General, Not singular, Nature of Moral Certainties

If we examine the examples Wittgenstein provides in *On Certainty*, we see that almost all of them are singular propositions—statements like "This is my hand," "The earth has existed for many years," or "My name is L.W." In contrast, moral certainties, e.g., "Murder is wrong," are usually expressed in the form of general propositions. Wittgenstein states that, "the squirrel does not infer by induction that it is going to need stores next winter as well. And no more do we need a law of induction to justify our actions or our predictions" (OC: §287). Elsewhere, he remarks, "The 'law of induction' can no more be grounded than certain particular propositions concerning the material of experience." (OC: §499). He also says, "why should we have learnt one universal law first, and not the special one straight away?" (OC: §133)

Unlike these certain propositions, the statements presented as examples of moral certainty are all general propositions. Kober, highlighting this distinction, explains that even ordinary moral discourse quickly enters the realm of philosophy. Even young children may ask why they must perform a particular action, and if parents wish to give a reasonable answer, they refer

the child to one of the general certainties that justify the action—certainties that themselves are not justified (Kober 1997, 378). The generality of these propositions leaves room for different interpretations, which carries certain consequences. Unlike examples of empirical propositions in *On Certainty*, these general moral norms tend to be ambiguous or subject to exceptions, intertwined with various other types of propositions, and open to questioning. In the following, I will elaborate on these three characteristics.

5. Moral Rules: Ambiguity and Exceptions

When we search for certain propositions that serve as definitive bases for moral evaluations and are universally accepted, it may seem that statements such as "injustice is wrong," "lying without reason is wrong," or "killing an innocent person without reason is wrong" are among them. Of course, if what is meant is a proposition, however vague, lacking clear instances, and subject to exceptions or interpretations, then such propositions can be considered examples of moral certainties. But the problem is that when a proposition is vague, resolving the ambiguity within it becomes the starting point for many disputes. For example, the concept of "justice" is vague to the extent that it does not provide any clarity in determining its instances. It is common to see people considering completely contradictory cases as instances of justice.

Similarly, when a proposition is subject to exceptions, disagreements frequently arise over where and under what conditions exceptions should apply. We may disagree about when and in which situations lying or killing is not wrong.

Furthermore, when we speak of "killing without reason," the question arises as to what qualifies as a "reason," which can lead to various disputes. Even in the phrase "innocent person," both words—"innocent" and "person"—are open to multiple interpretations, leading to differing positions. The mere fact that different interpretations or instances exist for a single moral norm means that it cannot serve as an unequivocal and decisive foundation.

6. The Complexity of Singular Moral Judgments

When general moral norms are applied to specific cases, various matters of other kinds are also involved. That is, the specific case has many characteristics that influence the judgment. As a result, the judgment regarding specific cases becomes much more debatable than the general

judgment. For example, the moral judgment "killing a person is wrong" potentially encompasses a wide range of instances, from abortion to the execution of a criminal; and abortion may be judged in different ways depending on the health condition of the fetus, the economic situation of the mother, and other factors. Therefore, many factors interfere with the judgment.

This is not the case with empirical propositions. For example, when I see a chair in a particular place and say, "There is a chair over there," this proposition does not involve as many factors as the general proposition "Human senses are a way of discovering reality." Wittgenstein states, "... and I also trust my memory the whole time, and trust it without any reservation. The certainty here is the same as that of my never having been on the moon." (OC: §337). He explicitly considers certainty about a general principle and certainty about one of its instances to be the same. However, this is not the case with moral principles. When we have a specific moral issue to decide, we are not deciding just about a universal moral rule, but about a much more complex issue, arising from the combination of different types of matters. Basic moral norms, when combined with other types of judgments—such as metaphysical, empirical, psychological, sociological, economic, etc.—result in more complex moral judgments. These outcomes are just as dependent on basic moral beliefs as they are on beliefs from other domains. Therefore, before resolving moral disagreements, consensus must be reached in other areas. For instance, when discussing the fairness of a particular tax policy, we must evaluate its impact on both the taxpayer and society as a whole and reach a consensus. Similarly, when arguing about the responsibilities of parents, we must already have formed opinions about its psychological aspects. In fact, our understanding of issues related to a moral judgment continually refines our normative stance on that issue.

Furthermore, these non-moral beliefs are sometimes themselves basic beliefs of different kinds, often of a metaphysical nature. In such cases, resolving disagreements becomes significantly more difficult. For example, the debate over the morality of abortion stems from several fundamental disagreements that are extremely difficult, if not impossible, to resolve. One such disagreement concerns whether a fetus is a human being and, if so, at what point it can be considered as such. Another concerns the priority of life over

reducing pain and suffering, or vice versa. The reason why this issue remains unresolved among moral philosophers lies in these kinds of fundamental disagreements. Otherwise, they all agree that one should not unjustly harm or kill an innocent human being. Thus, religious beliefs also influence moral perspectives, whereas they do not intervene in scientific beliefs. In fact, since ethics—unlike factual matters—is not neutral regarding our interests and inclinations, its potential for such interventions is much greater.

7. Despite Epistemic Skepticism, Moral Skepticism Allows for Consistent Action

If the analogy between moral certainty and empirical certainty is intended as a response to moral skepticism, it will not be a successful one. This is because the examples Wittgenstein provides for empirical certainty are beliefs with which we begin and continue our lives—beliefs that cannot be removed from the lives of each of us. Our lives show that we know or are certain about these beliefs (OC: §7). Some facts stand before us and do not allow us to deny them. Wittgenstein asks, "Doesn't it seem obvious that the possibility of a language-game is conditioned by certain facts?" (OC: §617).

However, while we do live with moral beliefs that hold an important place in how we coexist, they do not necessarily manifest in everyone's lives, and an individual can live without considering them. If we do not believe in the toxicity of poison and fail to act accordingly, we will die. But we can reject—or even accept—the wrongness of killing and still choose not to act upon it. As Williams points out, *for the skeptic about reality, there is no alternative, but "there does seem to be an alternative to accepting ethical considerations. It lies in a life that is not an ethical life."* (Williams 2006: 24). In this case, the question "*Why should I be moral?*" demands an answer, and no matter which proposition we consider certain, there will always be someone who does not see themselves as bound to accept it.

Overall, it can be said that while morality is indeed an important aspect of our human lives and in fact, "it is there—like our life" (cf. OC: §559), it is there with its own characteristics. It is true that in morality, there are certain beliefs shared among us, and these shared beliefs enable our moral interactions with one another. However, there are crucial differences between these moral certainties and empirical certainties, particularly in terms of generality,

ambiguity, interpretability, and their being subject to exceptions. Therefore, Wittgenstein's insights in *On Certainty* cannot be directly replicated in ethics.

References

- Goodman, R. B. (1982) "Wittgenstein and ethics", *Metaphilosophy* 13(2), 138-148.
- Hermann, J. (2015) *Moral Certainty*, Palgrave Macmillan UK.
- Kober, M. (1997) "On epistemic and moral certainty: A Wittgensteinian approach", *International Journal of Philosophical Studies*, 5(3), 365-381.
- Pleasants, N. (2009), "Wittgenstein and basic moral certainty", *Philosophia* 37(4), 669-679.
- Rummens, S. (2013), "On the Possibility of a Wittgensteinian Account of Moral Certainty", *Philosophical forum* (Vol. 44, No. 2), pp. 125-47.
- Williams, B. (2006), *Ethics and the Limits of Philosophy*, Routledge.
- Wittgenstein, Ludwig (1969) *On Certainty / Über Gewißheit*. Hg. von G. E. M. Anscombe und G. H. von Wright, übersetzt von Denis Paul und G. E. M. Anscombe. Oxford: Basil Blackwell. [Deutsch mit englischer Übersetzung]

Discovering Ethics in Wittgenstein's Picture Theory

Herbert Hrachovec (Vienna, Austria)

Abstract

Michael Kremer, in writing about the purpose of Tractarian nonsense, takes Wittgenstein's concluding words of his *Tractatus logico-philosophicus* (TLP) quite literally. Its statements "... do not begin by making sense, only to be gradually reduced to nonsense. They are nonsense all along". Readers only realise the deception too late. No theories presented are even provisionally justifiable. "This may seem too revisionary. Where is the famous 'picture theory'? Where is the isomorphism between propositions and states-of-affairs? The answer is that the *Tractatus* has shown all that talk to be nonsense."

Kremer's throwaway attitude will be rejected here. Wittgenstein's theatrical finale distracts from the intellectual labour he invested in his book. One has to *climb* the ladder before doing away with it. The purpose of the objections presented is, however, not to defend substantial theses of the TLP; their aim is more modest and provocative at the same time. Details of the picture theory are not examined in order to defend Wittgenstein's theses as correct. The theory is shown to be questionable in several respects. Yet, the effort is by no means pointless. And neither does it necessarily lead to silence. A plausible ethical perspective can be gained from close reading.

Michael Kremer, in writing about the purpose of Tractarian nonsense, takes Wittgenstein's concluding words of his *Tractatus logico-philosophicus* (TLP) quite literally. Its statements "... do not begin by making sense, only to be gradually reduced to nonsense. They are nonsense all along" (Kremer 2001:42). Readers only realise the deception too late. No theories presented are even provisionally justifiable. "This may seem too revisionary. Where is the famous 'picture theory'? Where is the isomorphism between propositions and states-of-affairs? The answer is that the *Tractatus* has shown all that talk to be nonsense ..." (loc.cit.). Kremer's throwaway attitude will be rejected here. Wittgenstein's theatrical finale distracts from the intellectual labour he invested in his book. One has to *climb* the ladder before doing away with it. The purpose of the objections presented is, however, not to defend substantial theses of the TLP; their aim is more modest and provocative at the same time. Details of the picture theory are not examined in order to defend Wittgenstein's theses as correct. The theory is shown to be questionable in several respects. Yet, the effort is by no means pointless. And neither does it necessarily lead to silence. A plausible ethical perspective can be gained from close reading.

1 A crack within 'the famous picture theory'

The key points of the theory are well known. "We use the perceptible sign of a proposition (spoken or written, etc.) as a projection of a possible situation." (TLP 3.11) 'A sign does not determine a logical form unless it is taken together with its logico-syntactical employment.' (TLP 3.327) The image alone cannot determine whether what is depicted is true or false (cf. TLP 2.224). True elementary propositions are pictures that reflect the existence of facts (cf. 4.25). In general "A picture presents a situation in logical space, the existence and no-existence of states of affairs." (TLP 2.11) These definitions are inspired by everyday experiences like the pictorial re-enactment of a car crash (TB 29.9.1914) or the mutual fit of "a gramophone record, the musical idea, the written notes and the sound waves" (TLP 4.014). A more formal version of the idea of correspondence is discussed with regard to Newtonian mechanics (TLP 6.342 ff), concerning an image consisting of spots on a canvas and a coordinate system tracing these spots. They are registered according to their size and distribution. The resulting 1:1 mapping between an appropriately selected grid and the spots may be regarded as an isomorphism. The example seems to match the chain of correspondences observed between a record and its correlating sound waves. This juxtaposition of logic and mechanics (cf. TLP 6.342) passes over a crucial point, though, namely the disparity between isomorphism and truth. Two different notions of "mapping" are involved. Figuratively speaking: acoustic waves emanating from a record are expected to conform to a system of harmonies and not consist of deafening screeching. Such noise is still an isomorphism, though.

A subtle distinction is at work here. Wittgenstein calls "the arrangement of irregular black spots" a *picture* ("whatever kind of picture they make") and -- in the same sentence -- a *surface* "which I can always approximate as closely as I wish to the description of it ...". (TLP 6.341) There is a small but crucial difference between those two points of view. The presumed surface is a picture in the *extended sense* of being a visual form, but lacking a delimiting frame. Its description, on the other hand, *actually frames* its rendering of the patches, deploying optionally selectable grids to capture the underlying data. Here is the catch: Mechanics, given these prerequisites, "... is an attempt to construct according to a single plan all the true propositions that we need for the description of the world." (TLP 6.343) But unframed visual specks do not by

themselves suggest a system of coordinates. They are registered by data points, but this barely deserves to be called “description of the world”. ‘Truth’, in a loose sense, might be regarded as the systematically controlled replication of a structure according to a pre-established grid. But this falls short of a proper world view based on sentences attributing properties to things. Logic’s procedure includes content-related judgements over and above duplication. (A clone does not represent its original cell.) A “surface as a picture” may be reproduced by mechanical means, or else represented by conceptual specifications of a target area. Wittgenstein is aware of the difference, but gets into unexpected difficulties in the course of his treatise.

2 Isomorphy and Truth drift apart

The problem arises from an idiosyncrasy of TLP-logic which will be shown to have farreaching consequences, not least in ethics. The book sets up a striking difference between propositional logic and the realm of conventional predicate logic. The former offers a self-contained system for the truth-functional compilation of sentences referring to (possibly existing) facts. Its interaction between logical operators and assertions combines calculation and content, algebra with sentence tokens. Since this is done without regard to internal propositional structure, picture theory is useless here. One needs (at least) first order predicate logic to capture the attribution of concepts to things, i.e. properties. But Wittgenstein, in a baffling move, does away with quantors in one succinct sentence: “I dissociate the concept *all* from truth functions.” (TLP 5.521) The effects of this dictum are often overlooked. Generality, according to this move, is no operational aggregation of individual entities under the name of a property. Wittgenstein recurs to the image of a spotted surface to distinguish his notion of properties from quantifier logic. Tracing patches does not require an understanding of “black” and “white”. Representations, on the other hand, involve interpretation of signifiers. ‘I must first have determined in what circumstances I call “p” true, and in so doing I determine the sense of the proposition.’ (TLP 4.063) The meaning of sentences does not lie in verbal tokens, but in their ability to be assessed in the light of statements made. “Black” gets its meaning by its role in specifying facts.

One must keep in mind how predicate logic in Wittgenstein’s time (and in current model theory) deals with conceptual generality. A vocabulary provides terms and relational expressions the former to be assigned to elements of an

universe of discourse, the latter ones to sets of its elements. A statement is regarded as true if the bearer(s) of its names are elements of the set that its predicative expressions designates. An early remark from the "Notes on Logic" indicates Wittgenstein's reason to reject this approach. "... It comes to seem as if logic deals with things which have been deprived of all properties except thing-hood, and with propositions deprived of all properties except complexity." (Ts 201, a2,b23). To put it in terms of ideological criticism: Conventional quantification tears things and properties apart in order to then put them back together again (at one's own discretion). The truth of sentences is thus determined by freely selectable interpretation functions. Wittgenstein's move turns against this set-theoretic mechanisms. He regards names/things and sentence structure/facts as inseparable. For him, truth is not an isolated mark of assent, but ultimately the structural underpinning of the entire net of sentences correctly describing the world. His properties are formal traits shared by all sentences matching a certain feature of reality. Nothing is to be taken out of context in syntax; respect is shown towards the embedding of objects in their native environment (semantics).

The details of Wittgenstein's proposal are somewhat sketchy and rarely commented upon in the scientific literature. A brief outline must suffice. Expressions from which sentences are formed contribute to their truth values in different ways. "Caesar", for example, indicates the person to whom all true or false statements concerning his existence are attributed. Wittgenstein calls the set of all sentences equi-functional to a meaningful sentence-part a sentence variable (cf. TLP 3.313). According to this definition, predicates are sentence variables representing all sentences fulfilling a given assertive function. To be more precise: all sentences that are to be taken into account when determining content-assignments within the sentences in question. In "Caesar crossed the Rubicon," "crossed" represents all sentences between whose subject(s) and object(s) the relationship of crossing may exist. This construction is based on the vaguely Hegelian concept of holistic truth: "The force of a proposition reaches through the whole of logical space." (TLP 3.42) But there is trouble ahead. Wittgenstein's concern with meaning by verification clashes with his logical absolutism. Logic taken as a priori discipline is supposed to exert exclusive control over its variables. Applied logic, on the other hand, is concerned with actual state of affairs. It cannot freely invent its

variables. In rejecting the model-theoretical approach Wittgenstein opts for including globally contextual to-be-confirmed content. But this does not square with truth-functional completeness. Once again in the jargon of critical theory: logical formalism is pitted against logic's applicability. Power of domination against power of judgement.

3 Syntax teams up with ontology

To recapitulate the problem: "What values a propositional variable may take is something that must be stipulated." (TLP 3.316) Meanings are explicitly excluded which equals the logical fiat of model theory. But propositional variables incorporate worldly content. They contain all sentences used with a certain purpose, namely contributing to establish the truth of a given, assessable expression. Correct statements refer to existing facts external to the a priori realm of logic. As a mere sentence variable, "crossing" is a formal term. We may determine its meaning at the drawing board. This will, obviously, be arbitrary if the expression is not familiar from sentences applying it to a factually determined relationship. But, on the other hand, the "fixing" of an empirically legitimate variable will have to be a matter of experience. How does this relate to picture theory? The formal features of isomorphisms are unavailable as long as one is unaware of the complete set of instances comprised in sentence variables.

Wittgenstein, at this point, introduces a second eye-raising feature: "What is peculiar to the generality sign is first, that it indicates a logical prototype and secondly, that it gives prominence to constants." (TLP 5.522) The constants, representative of the sentences concerned, embody structural relationships that permeate the facts of this world. But what purpose do "logical prototypes" serve? In the TLP empirical generality is established by adding up judgements. The sentences thus collected have, however, to be *selected* according to a pattern. The meaning of an expression must be established prior to judgement (TLP 4.063). Wittgenstein's diaries contain a succinct description of this move: "... the proposition would consist of proto-pictures, which were projected on to the world." (TB 12.11.14, 30e) and: "Even if the sentences which we ordinarily use all contain generalizations, still there must surely occur in them the proto-pictures of the component parts of their special cases." (TB 7.5.15, 46e) Patterns are to be imposed upon visual experiences in order to be recognized as pictures; as paradigms, if you wish. Warrants of generality required by picture

theory. They are the determining force behind assertive practices (underexposed in the TLP).

Prototypes are metaphysical constructs. We have, to get a clearer picture, to look at the ontology outlined in the TLP's first two main propositions. The most prominent ontological feature of a thing is its embedding in facts: "It is essential to things that they should be possible constituents of state of affairs." (TLP 2.011) The possibility of those state of affairs themselves is determined by the entirety of logical forms available/confirmable in judgements. Since nothing in logic is accidental (cf. TLP 2.012) those forms must serve a purpose, i.e. already be inherent in a thing's internal composition. (Superfluous forms fall victim to Occam's razor, cf. TLP 3.328). The possibility of the state of affairs is the ontological counterpart to the device of sentence variables, among them those amounting to names (cf. TLP 3.314). These are containing the complete structural information of any entity's obligatory embedding in state of affairs, in other words its essential features. Think of a visual speck: logic cannot assign it any particular color, but the speck *has to have* one or the other. Knowing an object amounts to knowing "all its possible occurrences in state of affairs" (TLP 2.0123), its indispensable features which Wittgenstein refers to as "formal properties of objects and state of affairs" (TLP 4.122). Instructive examples are his (albeit cursory) remarks on space, color and tones (cf. TLP 2.0131). Logic has to be up to the *type* of phenomena, not to the specification of its *tokens*.

Sentence variables register the nodes at which propositions overlap in terms of their contribution to the truth of propositions. They are Wittgenstein's version of general terms, recording a common functionality of the sentence parts concerned. Names (as propositional variables) are consequently inseparable from the formal properties of their bearers and as such cover all of the world's possibilities, which TLP-logic, in turn, regards as its own facts (cf. TLP 2.0121). No surprises are therefore to be expected (cf. TLP 6.1251). But here we come up against a pressing conceptual challenge. TLP-Logic does not recognise development: "A new possibility cannot be discovered later" (TLP 2.0123), which means that it is to account for the syntactic pattern of precisely every required sentence variable in advance. Yet, as our analysis of the picture

theory has shown, in “making sense” logic depends on confirmation from outside its realm. Where does that leave us? Substantial changes to the world, and in particular ethical intervention, seem to be ruled out.

4 Showing by doing

Two questions remain to be addressed. How is the systematic stalemate between TLP-logic and the world to be assessed and what consequences does this have for ethical considerations. Wittgenstein does not shy away from addressing the first question. Admitting (1) that his logic does not have *in toto* access to atomic sentences he asks whether it could get into the position of needing a 27-digit relation. “If I cannot say *a priori* what elementary propositions there are, then the attempt to do so must lead to obvious nonsense.” (TLP 5.5571) Therefore (2) the following supposition is to be rejected: “It is supposed to be possible to answer *a priori* the question whether I can get into a position in which I need the sign for a 27-termed relation in order to signify something.” (TLP 5.5541)” Taken together those statements amount to a methodological impasse. Logic cannot be dependent from its application and, at the same time, be unalterable. Wittgenstein neither subscribes to model-theoretical formalism nor to logical empiricism. What remains is an epistemological equilibrium in which both sides are irrevocably opposed to each other.

The present result has been arrived at without recourse to the saying/showing distinction which generations of interpreters have tried to work out. The TLP's finale might be regarded as a consequence of the standoff described. General concepts, indicated by logical proto-types (cf. TLP 5.522) make up the formal structure of the world. By insisting on logic being strictly *a priori* and applicable to state of affairs Wittgenstein is blocking further arguments, leading over to silence. But must discussion stop here? The current reading has reconstructed a methodological *tugging-game* between conflicting intuitions, triggered by Wittgenstein's key move to abandon quantificational logic. Now, consider it as a *game*, i.e. as an ingenuous competition with the participants taking turn in advocating their position. The point of the enterprise might not to pick a winner. The TLP, contrary to prevailing opinion, is not simply a monolithic *tour de force*, but just as much a compilation of contradictory motives.

Wittgenstein, in a famous letter to Ludwig Ficker, claims that his book limits the ethical from within. This limitation is in itself an ethical endeavour. Has anyone pointed out that, therefore, the picture theory's elaborate technical apparatus is (to a certain extent) an act of resistance against attempts of order and entropy to overwhelm each other? (This is where Wittgenstein's "mysticism" seems to reside.) The *admirabile commercium* (a detour into scholastic theology may be forgiven here) between TLP-logic and its world is a solution balancing the interests between the urgency of rationality and its susceptibility to disruption. A durable ceasefire (another detour, as these lines are written). We have dealt with a monologue, but let's think of it as a dialogue, e.g. in the format of the *Philosophical Investigations*. Its "iron must" of logic vis à vis the vagaries of rule-following are already in play in the TLP, see for example the book's "famous picture theory". Note to Michael Kremer: it shows if one looks for it.

References

Kremer, M. (2001). "The Purpose of Tractarian Nonsense", *Noûs*, 35(1), 39–73. <https://doi.org/10.1111/0029-4624.00287> (23.3.2025).

Wittgenstein, L. 2015 Bergen Nachlass Edition. Edited by the Wittgenstein Archives at the University of Bergen under the direction of Alois Pichler. Electronic Document: <http://www.wittgensteinsource.org> (23.3.2025).

Wittgenstein, L. 1969. *Notebooks 1914-1916*, Ed. Georg H. von Wright and Gertrude E. M. Anscombe. Trans. Gertrude E. M. Anscombe. New York: Harper & Row. [NB].

Wittgenstein, L. 1972. *Tractatus Logico-Philosophicus*. Trans. D. F. Pears und B. F. McGuinness. London: Routledge & Kegan Paul. [TLP].

Anscombe and Diamond on Wittgenstein's Notion of Sense in the *Tractatus*

Andrej Jandrić (Belgrade, Serbia)

Abstract

Elizabeth Anscombe pointed out that correctly identifying Frege's influence on Wittgenstein is necessary for understanding the *Tractatus*. In her view, Wittgenstein adopted Frege's notion of *Sinn*, or sense, together with his context principle. According to Anscombe, Wittgenstein used the same notion of sense as Frege, but claimed, in contrast to Frege, that only sentences could have sense. Cora Diamond criticised Anscombe and argued that Wittgenstein's notion of sense was different from Frege's. She emphasised that Wittgenstein operated with a directed notion of sense. It is essential for Wittgenstein's understanding of *Sinn* that from the sense of a sentence one can immediately read the sense of its negation: the sense of the negation is the reversal of the sense of the original sentence. I will argue that Frege understood the *Sinn* of an expression as a binary classification associated with the expression, and that such a notion of sense is already directed. Differences between Frege's and Wittgenstein's claims about which expressions can have sense are to be attributed to the non-Fregean aspects of the *Tractatus*, which form the background against which Wittgenstein employed the Fregean notion of *Sinn*.

1

In her 2019 book *Reading Wittgenstein with Anscombe, Going Back to Ethics*, Cora Diamond pays homage to Elizabeth Anscombe's *An Introduction to Wittgenstein's Tractatus*, the work which has exerted a profound influence on her thinking about Wittgenstein. Diamond honours Anscombe in the only appropriate way: by engaging in a critical discussion with some of her views.

Anscombe (1959: 12) was perhaps the first to point out that many misinterpretations of the *Tractatus* result from the failure to properly address Frege's influence on Wittgenstein. Evidence of this influence is, in her opinion (Anscombe 1959: 75), visible in Wittgenstein's adoption of Frege's notions of sense (*Sinn*) and reference (*Bedeutung*) (Frege 1984b) and his context principle from *Die Grundlagen der Arithmetik* (Frege 1960: § 60). However, Wittgenstein formulates his version of the context principle in § 3.3:

Only propositions (*Satz*) have sense (*Sinn*); only in the nexus of a proposition (*Satz*) does a name (*Name*) have meaning (*Bedeutung*).

Diamond (2019a: 17) correctly observes that, while the second part of the sentence bears witness to the impact of Frege's philosophy of language on Wittgenstein, the first part of the sentence is doubtlessly meant to be criticism

of Frege's views. It is not just sentences which have sense in Frege, but all kinds of linguistic expressions: names, predicates, etc. Anscombe believes that Wittgenstein here uses Frege's notion of sense, but makes different claims about that very notion than Frege does. According to Diamond (2019a: 8), on the other hand, this difference is a clear indication that Frege and Wittgenstein have different notions of sense behind the common term *Sinn*. In Diamond's view (2019a: 18-23), Wittgenstein's notion of sense is, unlike Frege's, directed. It is essential for Wittgenstein's notion of sense that from the sense of a sentence one can immediately read the sense of its negation: the sense of the negation is the reversal of the sense of the original sentence (TLP § 5.2341); doubling the reversal would thus cancel out and yield the initial sense. Wittgenstein assimilates sense to an arrow or a vector (TLP § 3.144); negating a sentence is, accordingly, like rotating the vector for a straight angle. A sentence can be a logical picture of a fact only if it can be used to say how things stand, but also how things do not stand (TLP § 4.0621): the latter is just the reverse orientation of the former vector. Directedness of sense also explains why sentences with *Sinn* can be combined into more complex sentences which would also have *Sinn* (Diamond 2019b: 105).

In what follows I will defend Anscombe's view from Diamond's criticism and argue that Frege already had a directed notion of sense.

2

How Frege understood *Sinn* is, perhaps, most clearly visible in his considerations of the sense of a predicate. For Frege, predicates are just a special case of functional expressions: they are unsaturated expressions in that they contain a gap, or an empty place, for their argument. Since unsaturated expressions refer to unsaturated entities, references of predicates cannot be objects: instead they are concepts (*Begriffe*). A predicate needs to be supplemented with a name to make up a meaningful sentence (*Satz*); at the ontological level, the concept, which the predicate refers to, needs to be supplemented with an object, which is the name's *Bedeutung*, in order to give a truth value of the corresponding *Satz*. Concepts are thus functions of one argument which assign truth values to objects: the value True to objects which fall under the concept (i.e. which belong to the concept's extension), and the value False to all other objects.

A predicate can pick out a specific concept only via its *Sinn*; as with other expressions in Frege, there is no direct reference. Consequently, the *Sinn* of a predicate has to enable sharp discrimination between two groups of objects (Frege 1984a: 148): it is none other than a *binary classification* of objects. If the conditions contained in a predicate's sense are such as to leave open if a certain object (of any kind whatsoever) falls under the concept (the predicate aims to refer to) or not, then, according to Frege, no sense has been given to the predicate and, consequently, it is bound for reference failure. Hence, if the predicate "red", for instance, is to refer to the concept of redness, it has to be given a determinate sense that is sharp enough to divide objects into two mutually exclusive and collectively exhaustive sets of objects: one of them is the set of red objects, the other the set of all objects that are not red, among which there will be objects which are not coloured at all, such as numbers, England or the Earth's axis; the former set is *designated* as the extension of the concept *red*.

This understanding of *Sinn* of an expression as a binary classification (of entities of some kind) is that which provides for the directedness of sense, as described by Diamond. To every predicate there corresponds the opposite, or the reversed predicate; the opposite or "red" is "non-red". If "red" is given a sense, then it immediately follows that "non-red" has also been given a sense and that its sense is determined by the sense of "red": to opposite predicates there corresponds the same binary classification, namely that into red and non-red objects, it is only that a different set (of the two) is being designated as the extension. By reversing the sense of "red", we keep the binary divide intact, and switch from one of the elements of the divide to the other; this is how we arrive at the sense of "non-red". Reversing it once again would lead us back to the sense of "red" from which we started. It also becomes quite obvious that predicates with *Sinn* combine into a new predicate which will also have *Sinn*: if "red" is being associated with one binary classification and "round" with another, then to "red and round" would also correspond a binary classification: namely that into the objects which possess both properties and all the other objects. The disjunctive predicate "red or round" would also be guaranteed a sense, as another binary classification generated by those for "red" and "round" would correspond to it.

The *Sinn* of a sentence is also a binary classification, only not of objects but of possible worlds. According to Frege, the *Sinn* of a sentence is the thought (*Gedanke*) it expresses, which is understood as the sentence's truth-conditions. Although Frege does not explicitly use the terminology of possible worlds, it is obvious that a sentence's truth-conditions divide possible worlds into those in which these conditions are satisfied and those in which they are not. The directedness of thusly conceived sense is evident: to the negation of a sentence, there again corresponds the same binary classification, only with the different element of the pair being designated as the worlds in which the sentence is true.

The sense of a name should enable singling out a unique object: hence, it divides objects into a designated object (if there is such in the world) and all the other objects. The sense of a name, say, "Venus" is closely connected to the sense of predicate "is identical with Venus": to both there corresponds the same binary classification of objects, only with the name it is object, i.e. Venus, that is being designated and with the predicate a singleton set, i.e. {Venus}. To say of a thing that it is not Venus (i.e. not a reference of "Venus") is to ascribe it the concept of not being identical with Venus. Hence, for all types of expression Frege understands *Sinn* as a binary classification (with a designated element) associated with the expression; such understanding of sense makes it directed.

3

In the *Tractatus*, Wittgenstein also understands *Sinn* as a binary classification. This is apparent in the paragraph 4.063, in which he compares the meaningful sentence (which is a logical picture) with a black spot on a white paper: its sense can be given by specifying for every point in the paper whether it is black or white, i.e. by specifying a binary classification of points. The points stand for "the truth-possibilities of elementary propositions" (TLP 4.063), and the black or white colour for "agreement or disagreement" with those truth-possibilities. In § 4.431 he adds that "the expression of agreement and disagreement with the truth-possibilities of elementary propositions expresses the truth-conditions of a proposition (*Satz*)" and that "a proposition (*Satz*) is the expression of its truth-conditions": they are its *Sinn*. The directedness of sense is an immediate consequence of this Fregean conception of sense.

But if Wittgenstein understands sense in the same way as Frege, as Anscombe claims he does, then why does he, in contrast to Frege, denies sense to all expressions but sentences (TLP § 3.3)? The answer is to be found in some non-Fregean aspects of the *Tractatus*, which form the background against which Wittgenstein uses the Fregean notion of sense as a binary classification.

Predicates cannot be attributed sense in the *Tractatus*, simply because there are no predicates in the Tractarian language (TLP §§ 4.22, 6-6.001). The only expressions there are beside sentences are names. So, the question is rather if a sense can be given to names, i.e. if names can be associated with a binary classification of entities of some kind.

Could names, perhaps, have a sense which would divide possible worlds into two sets, exactly like the sense of a meaningful sentence does? Just as the *Sinn* of a sentence groups together the possible worlds in which the sentence is true, as opposed to the possible worlds in which it is false, so could the *Sinn* of a name distinguish between the possible worlds which contain the object the names refers to and those worlds the domain of which does not contain the object in question. This is, however, precluded by Wittgenstein's understanding of objects in the *Tractatus*. In paragraph 2.0271, Wittgenstein says that "objects are what is unalterable and subsistent". In the same group of paragraphs, he also states that "objects make up the substance of the world" (TLP § 2.021) and that "substance is what subsists independently of what is the case" (TLP § 2.024). This suggests that the objects exist of necessity, i.e. in every possible world, since they exist "independently of what is the case". In contemporary metaphysics, Wittgenstein is often regarded as the twentieth century predecessor of the view now labelled as *necessitism* (cf. Williamson 2013: 1-4), according to which everything that exists, necessarily exists (i.e. domains of all possible worlds coincide). Evidence for ascribing him such a view is typically found in paragraphs 2.022 and 2.023, in which he explicitly addresses the difference between the actual and merely possible worlds:

It is obvious that an imagined world, however different it may be from the real one, must have *something*—a form—in common with it. Objects are just what constitute this unalterable form.

Thus, possible worlds are indistinguishable with regard to objects they contain: the only difference between them lies in the way their common objects are organised into states of affairs (*Sachverhalte*). Therefore, if a name had a Fregean *Sinn*, it could not consist in a binary classification of possible worlds with regard to their containing or not containing the name's reference.

But could a name be given sense by associating it with a binary classification of states of affairs (*Sachverhalte*) or situations (*Sachlagen*) in the world? The answer is again negative. This time it results from Wittgenstein's rejection of Leibniz's law of identity of indiscernibles: according to this controversial metaphysical principle of Leibniz, objects which are indistinguishable with regard to their (intrinsic) properties have to coincide. Objects in the *Tractatus* are distinguished with regard to their formal and material properties (TLP § 2.0231). An object cannot be combined with just any other object in a state of affairs. The logical form of an object, which Wittgenstein also labels as the object's nature (TLP § 2.0123), is determined by its combinatorial possibilities: formal properties of an object are the states of affairs it can appear in. States of affairs are possible configurations of objects and not all of them exist in the world, those that do are situations; material properties of an object are situations it appears in. Thus, it may seem natural to ascribe to a name a binary division of *Sachverhalte* into those in which the object the name refers to may appear and those in which it may not, and hence to ascribe a name a Fregean sense. However, the sense of a name has to be such as to aim to pick out a unique object; names with different references thus cannot have the same sense. But that would not be precluded if the sense of a name is understood as a binary classification of *Sachverhalte*, since Wittgenstein allows for numerically distinct objects which are indistinguishable with regard to their formal properties: every possible configuration of objects which contains one of these objects also contains the other. If two objects are indistinguishable with regard to their formal properties, they are automatically indistinguishable with regard to their material properties: since *Sachlagen* are nothing but *Sachverhalte* which obtain in the world, and as every *Sachverhalt* which contains one of them contains the other as well, then every *Sachlage* which contains one of these objects also contains the other. Wittgenstein explicitly rejects Leibniz's law of the identity of indistinguishable objects in § 2.02331:

Either a thing has properties that nothing else has, in which case we can immediately use a description to distinguish it from the others and refer to it; or, on the other hand, there are several things that have the whole set of their properties in common, in which case it is quite impossible to indicate one of them.

Hence, neither *Sachverhalte* nor *Sachlagen* can serve as a ground for attributing sense to names.

But could not a name be given a sense in the manner of Frege: by associating it with a binary classification of objects, such that one side in the classification contains a single object the name refers to and the other side all the other objects. Anscombe provides compelling reasons why this cannot be done. If a name is to pick out an object among all the objects by means of its *Sinn*, then this *Sinn* has to be a definite description. In that case the name would be merely an abbreviation for the description, and it could be replaced with the description in every sentence in which it appears. Names would thus become semantically complex expressions; there would be no semantically simple ones in the language. This would, in turn, make the *Sinn* of every sentence indeterminate; consequently, the language would lose its ability to represent the world. In paragraph 3.23, Wittgenstein says:

The requirement that simple signs be possible is the requirement that sense be determinate.

And in § 3.24, he adds:

When a propositional element signifies a complex, this can be seen from an indeterminateness in the propositions in which it occurs.

Anscombe (1959: 41-50) claims that in these passages we should recognize Russell's influence on Wittgenstein: in her view, they are the result of Wittgenstein's adoption of Russell's analysis of definite descriptions. According to Russell, a definite description may have a primary or a secondary occurrence in a sentence: depending on its occurrence, the sentence will express different propositions. Hence, if the description "the king of France" occurs primarily in "The king of France is not bald", the sentence claims that

(KF1) There is a unique king of France and he is not bald,

whereas if it occurs secondarily, the sentence says that

(KF2) It is not the case that there is a unique king of France who is bald;

consequently, the sense of the sentence becomes indeterminate between the two readings (Russell 1905: 490; 2010: 88-89). The sentence in question contains negation; the occurrence of the description determines the scope of the negation. But how can it follow from such considerations that every sentence is vague? There are sentences with no such operators as negation, as, for instance, "The king of France is bald". Here there seems to be no place for the Russellian difference in the description's occurrence. However, to think thus would exactly mean to disregard the directedness of sense: due to its directedness, the sense of the latter sentence must be the same as the sense of "It is not the case that it is not the case that the king of France is bald", i.e. doubly reversing the sense with negation yields the initial sense. By applying Russell's analysis to the last sentence, it becomes clear that its sense oscillates between expressing the negation of (KF1) and the negation of (KF2). Hence, an attempt to ascribe names sense in this way would have disastrous consequences: it would make the whole language indeterminate and deprived of meaning. Thus no binary classification, and hence no *Sinn*, can be associated with names.

4

If the above considerations are correct, then Wittgenstein inherited from Frege not only the term *Sinn*, but also the notion behind it, just as Anscombe claimed. They both understood the sense of an expression as a binary classification associated with that expression. Such a notion exhibits directedness that Diamond has noted. In the *Tractatus* it is only sentences that can have sense, but not because Wittgenstein's notion of sense is directed and Frege's is not; it is rather because Wittgenstein applied Frege's notion of sense in a radically non-Fregean setting: elimination of predicates, necessitism, rejection of Leibniz's law of identity of indiscernibles and adoption of Russell's analysis of definite descriptions are the aspects of the *Tractatus* which are not to be found in Frege's philosophy. It is these aspects that are responsible for Wittgenstein's making divergent claims about the very same notion of sense that Frege was using.

References

Anscombe, G. E. M. (1959) *An Introduction to Wittgenstein's Tractatus*, New York: Harper and Row, Publishers.

Diamond, Cora (2019a), "Wittgenstein, Anscombe and the Activity of Philosophy", in her: *Reading Wittgenstein with Anscombe, Going On to Ethics*, Cambridge, Massachusetts and London: Harvard University Press, 7-44.

Diamond, Cora (2019b) "Reading the *Tractatus* with G. E. M: Anscombe", in her *Reading Wittgenstein with Anscombe, Going On to Ethics*, Cambridge, Massachusetts and London: Harvard University Press, 97-153.

Frege, Gottlob (1960) *The Foundations of Arithmetic*, New York: Harper and Brothers.

Frege, Gottlob (1984a) "Function and Concept", in: *Collected Papers on Mathematics, Logic, and Philosophy*, ed. B. McGuinness, Oxford: Basil Blackwell, 137-156.

Frege, Gottlob (1984b) "On Sense and Meaning", in: *Collected Papers on Mathematics, Logic, and Philosophy*, ed. B. McGuinness, Oxford: Basil Blackwell, 157-177.

Russell, Bertrand (1905) "On Denoting", *Mind* 14, 479-493.

Russell, Bertrand (2010) *The Philosophy of Logical Atomism*, London and New York: Routledge.

Williamson, Timothy (2013) *Modal Logic as Metaphysics*, Oxford: Oxford University Press.

Wittgenstein, Ludwig (1972): *Tractatus Logico-Philosophicus*. Übersetzt von D. F. Pears and B. F. McGuinness. Reprinted with corrections in the light of the publication of Wittgenstein's correspondence with C. K. Ogden about the first translation. London: Routledge and Kegan Paul.

Acknowledgement

My research has been supported by the Ministry of Science, Technological Development and Innovation of the Republic of Serbia (451-03-137/2025-03/200163), the Croatian Science Foundation (IP-2022-10-2550) and the University of Rijeka, Croatia (uniri-human-23-109).

Embodied Hermeneutic Injustice

Miriam Jerade, Celeste Vecino (Santiago, Chile)

Abstract

This paper examines how hermeneutic injustice manifests through the lived body, building on Iris Marion Young's phenomenological analysis of women's embodiment as a site of systemic oppression. While Miranda Fricker's theory of hermeneutic injustice highlights how conceptual gaps prevent understanding and communication, it treats embodiment as peripheral to epistemic concerns. The paper explores the corporeal dimension of hermeneutic injustice by examining specific aspects of female embodiment, presenting how social meanings actively constitute -rather than merely interpret - bodily experience. This reconceptualization advances our understanding of hermeneutic injustice by showing how interpretive frameworks both shape and are shaped by embodied experience.

Introduction: Bridging Phenomenology and Epistemic Injustice

Consider how a teenage girl experiences her first period. The physical sensation is inseparable from internalized social meanings of shame, disgust, and required vigilance. These social meanings, constituting collective frameworks for interpretation and coordinate action, thought, and affect (Haslanger 2013), don't simply overlay a neutral bodily experience - they fundamentally shape how the experience is lived and understood. Is this a case of hermeneutic injustice? While Fricker's valuable framework (2007) shows that the lack of hermeneutical resources prevents marginalized groups from making sense of or communicating their experiences, it treats embodiment as secondary to epistemic concerns. This paper argues that social meanings actively shape and constrain embodied experiences, potentially constituting a form of hermeneutic injustice that operates at the level of lived corporeality.

Iris Marion Young's phenomenological analyses of women's embodied experiences provide crucial insights for expanding our understanding of hermeneutic injustice. She positions the Female Body Experience as both a vital source of meaning-making and a site where societal structures constrain women's embodied experiences. Through works like "Throwing Like a Girl" and "Menstrual Meditations," Young demonstrates how social meanings in a patriarchal society restrict women's full bodily engagement with the world, shaping their comportment and motility. According to Young, women often experience a sort of unintelligibility in their ability to place themselves entirely in their bodies and their bodies entirely in the world- a phenomenon

that extends beyond individual psychology to reflect broader social and historical constructions.

Building on Young's insights, I argue that hermeneutic resources are not merely absent or distorted but actively shape how we experience and perceive reality. While José Medina & Tempest Henning (2021) have addressed the body as a source of testimony in their expressive behavior, considering the embodied positionality of the knowers, Young phenomenology positions the embodied subject not merely as expressive but as a source of meaning-making - "significant as a ground of existence" (Salamon 2017: 47). As Young defines it: "The lived body is a unified idea of a physical body acting and experiencing in a specific sociocultural context; it is body-in-situation" (Young 2005: 16).

Twenty years after her original work, Young reflected that she used existential phenomenology to "theorize the effects of feminine socialization and sexual objectification on a woman's world-making movements, describing us as unable to be free in movement." (Young 1998: 286) While acknowledging limitations in her approach, particularly her reliance on a masculine style and overemphasis on oppression in women's comportment, Young maintained the enduring value of her insights into women's embodied experiences. Although phenomenology and epistemology remain two distinct fields, Young's critique of phenomenology gender-blind universalism and feminist theory's occasional neglect of embodiment reveals the need for their intersection. While scholars like Ahmed and Alcoff have made significant strides in bridging this theoretical divide (Ahmed 2006; Alcoff 2006; 2018), developing a comprehensive embodied epistemology remains an ongoing philosophical challenge.

Though Fricker frames her chapter on Hermeneutic Injustice using Marxist and feminist standpoint theory, she primarily adopts an epistemological rather than material reading of structural disadvantage. While acknowledging that "certain material advantages will generate the envisaged epistemological advantage" (147), Fricker's interpretation focuses primarily on distributive injustice through access to resources like education and technology. Young's framework, however, provides deeper insight into how structures - simultaneously ontological, material, and epistemic - create conditions for hermeneutic injustice. The phenomenological approach reveals the intricate

relationship between experience as embodiment and social and historical constituted interpretations, demonstrating how systemic inequities fundamentally impair women's ability to articulate their lived experiences.

This paper develops its argument in two stages: first, by examining how female body experiences manifest as cases of hermeneutic injustice, and second, by demonstrating why Fricker's conception of collective hermeneutical resources as a shared pool of meanings (Fricker 2016) fails to capture the dynamic relationship between meanings and embodied experiences. Rather than viewing collective hermeneutical resources as a neutral repository of meanings, I argue that these resources actively shape and constrain bodily experience. Cultural narratives, categories, scripts, and discourses about femininity don't merely provide interpretations but actively shape and constrain women's bodily experiences, potentially constituting a form of hermeneutic injustice.

1. Cases of Embodied Hermeneutic Injustice

In "Menstrual Meditations," an unpublished text on Young's life, she explains how in certain cultures the meaning of menstruation "covered over by patriarchal biases and interest" (Young 2005: 97), determines the experience of menstruation for adolescent girls and adult women with feelings of shame, disgust, and humiliation. In some modern 20th-century societies, the social meaning of menstruation is paradoxical: menstruation is supposedly acceptable, with women expected to participate in the same way as men in social and physical activities, yet simultaneously they must vigilantly conceal their bleeding. Even advertising of companies like Kotex, notes Young, present menstruation as a healthy process that is nevertheless dirty, "presenting a hygiene problem that needs managing with their products" (Young 2005: 103). There is a cultural meaning of menstruation that women internalize as repugnant and arouses certain moods and affections.

This might not initially appear as a case of hermeneutic injustice if we claim that there is no gap or a lacuna in our collective hermeneutical resources. However, as Arianna Falbo pointed out, Fricker's "lacuna-centered analysis" (Falbo 2022) precludes other explanations of hermeneutic injustice related to the productivity of concepts. Or as Sally Haslanger puts it: "Although hermeneutical lacunae are a significant source of harm, Fricker's focus on

lacunae side-steps the ways in which symbolic power is, as Fraser put it, productive. Culture provides tools for constituting ourselves as subjects, and the problem is not just what tools are left out, but what sort of subjects we become given the tools we have” (Haslanger 2019: 15). Menstruation carries symbolic power that influences every woman's life at every age. It also has a normative dimension that requires menstrual etiquette. The meaning of menstruation and its symbolic and normative power constitute women's subjectivity, counterposing the abjection of menstruation to the clean and proper masculine. This has influenced not only their behavior but also their place in society; as Young writes: “If a woman wishes to walk among men as she bleeds, if she wishes to lay claim to the rights and privileges of a solid self who stand forth and achieves, then she had better keep her private fluidity secret.” (Young 2005: 111)

Young is precisely working on how social meanings constitute us as embodied subjects while recognizing that menstruation has different meanings in different contexts; it can be annoying, it can also be a sign of pregnancy – desired or avoided, or a sign of menopause. The meaning of menstruation requires an intersectional analysis: “My guess is that women often experience greater problems around menstruation in settings of greater poverty and sexual inequality” (Young 2005: 106). Her insight points toward the necessity of an analysis that considers how class, race, disability, and gender identity shape menstrual experiences. This intersectional approach reveals how multiple, overlapping structures of power and meaning-making mediate bodily experiences.

Young moves beyond phenomenological description to advocate for transforming the social meaning of menstruation, aiming to reframe women's bodies from “abject or monstrous to the merely different” (Young 2005: 117). This transformation is crucial because menstruation's physiological effects, including mood changes, are often weaponized against women through “affective injustice” (Whitney 2018) – a systematic invalidation of women's emotional testimonies that results in an improper uptake. Young argues that rather than suppressing menstruation-related moods, they should be recognized as sources of reflective insight.

This reframing reveals a double dimension of hermeneutic injustice: first, in how patriarchal interpretations that are embedded within broader structures of oppression shape women's lived experience of menstruation, and second, in how these interpretations can be transformed to enable new modes of embodied experience. This parallels Carmita Wood's experience of sexual harassment, where an absence of a concept with normative properties led to interpreting workplace harassment as mere flirting. An interpretation that shaped not only their understanding but also their embodied responses to these encounters. According to Brownmiller's chronicles, Wood's experience of hermeneutic injustice manifested initially as physical symptoms: a corporeal response to an unnamed injustice. The embodied experience of sexual harassment has degrees of intelligibility, where bodily reactions and emotions can either illuminate the presence of injustice or remain obscured by prevailing gender ideologies. Wood's case demonstrates how hermeneutic resources affect cognitive interpretation and bodily experience, suggesting that addressing hermeneutic injustice requires attention to both dimensions.

2. Social Meanings and Embodiment

While menstruation exemplifies how social meanings shape intimate bodily experiences, Young's analysis of feminine motility demonstrates how hermeneutic injustice extends to the broader domain of bodily capabilities and spatial engagement. In her groundbreaking essay "Throwing Like a Girl" (1980), Young employs phenomenological analysis to examine a seemingly simple observation: from an early age, girls typically don't engage their entire bodies in throwing motions as boys do. Challenging Straus's biological determinism, Young locates this difference in women's socialization rather than biological or anatomical distinctions.

Young examines how the body orients itself toward its environment and possibilities for action by drawing on Merleau-Ponty's phenomenology. In his *Phenomenology of perception*, Merleau-Ponty characterizes embodiment as "the vehicle of being in the world" (Merleau-Ponty 2012: 84). The body enacts our disclosure of the world, which does not exist as an independent reality prior to this disclosure. Body and world form a unified correlation that expresses our existence in passive and pre-objective ways. Our habitual bodily movements and worldly engagements (our distinctive *style* of comportment) are meaningful, though this meaning isn't represented explicitly in

consciousness, but rather lived-through. Challenging a standard view of constitution as meaning-giving (*Sinngebung*), Merleau-Ponty stresses that meanings do not originate in the subject but are part of a natural and social history that the subject incorporates in her own experience.

Young centers on how women's socialization fundamentally shapes their bodily comportment, movement patterns, and spatial relationships. Particularly significant is what she terms "inhibited intentionality." In phenomenological terms, intentionality describes how the embodied subject encounters the world as a field of possibilities correlative to an "I can." Still, it also has a terrain of limitations expressed through an "I cannot." Young's crucial observation is that women experience this intentionality distinctively: they project the "I can" as belonging to others while internalizing the "I cannot" as their own. This "I cannot" becomes constitutive of feminine embodiment, while the "I can" remains the province of an imagined universal masculine subject. This 'I cannot' represents not just a physical limitation but a hermeneutical constraint that shapes how women interpret their bodily capabilities and possibilities for action.

Young writes: "Feminine bodily existence is also self-referred in that the woman takes herself to be the *object* of the motion rather than its originator. Feminine bodily existence is also self-referred to the extent that a woman is uncertain of her body's capacities and does not feel that its motions are entirely under her control" (Young 2005: 38-39). The phenomenon runs deeper than conscious restriction - it manifests in how girls learn to move through the world "like a girl," internalizing limitations that become inscribed in their bodily schemas.

This analysis gains additional force when placed in dialogue, as Alia Al-Saji did, with Fanon's work on racialized embodiment: "Having internalized or 'epidermalized' the 'I cannot' of objectifying vision into one's body schema, this 'I cannot' leads not only to a split sense of bodily self (between a subject who can and an object who cannot), but also to a form of hesitation that can be paralyzing" (Al-Saji 2014: 153). Fanon builds upon Merleau-Ponty's concept of the dynamic organization of bodily experience in relation to the tasks at hand (Merleau-Ponty 2012, 102-103), to demonstrate how a racialized man struggles to elaborate his body schema within a colonized context. A racial epidermal

schema adheres to his embodied experience, rooted in racial discourses and ideology that not only shape and restrain his motility and engagement with the world but also affect his internal bodily sensations and affective responses. (Fanon 1967) Fanon's "epidermalized" schema disrupts the development of embodied self-understanding that Merleau-Ponty describes.

Both Young and Fanon reveal how collective hermeneutical resources do not simply exist as a "shared pool of meanings" that anyone can access, as Fricker suggests (2016, 163). Fricker's conception assumes that these resources function primarily as tools for interpretation that subjects use to make sense of experiences that would exist regardless. In contrast, Young and Fanon demonstrate that these resources actively constitute our embodied subjectivity by structuring what is possible within embodied experience itself. The meanings embedded in these resources become inscribed in our bodily schemas, determining in advance how we experience our bodies and their capabilities. These resources remain historically contingent and subject to transformation - as Young noted when observing her daughter's generation's more confident movement patterns two decades later. Instead, the dynamic interplay between social interpretation and embodied experience reveals how hermeneutic injustice operates not just at the level of cognitive understanding but through the very structure of bodily existence. Young's analysis of restricted movement patterns reflects and perpetuates a collective interpretation of "feminine" that constrains women's physical engagement with their environment, creating what we might call an embodied hermeneutic disadvantage.

Beyond Western feminist thought, South American indigenous feminist concept of "body-territory" (*cuero-territorio*), warrants further exploration. This framework, developed by theorists and territorial rights activists like Lorena Cabnal (2010) and Julieta Paredes (2015), conceptualizes bodies not as individual entities but as fundamentally connected to ancestral lands. They demonstrate how colonial and extractivist violence simultaneously targets Indigenous women's bodies and communal territories. This perspective resonates with Husserl's phenomenological insight that the Earth as ground shares some key features with our body – serving as the zero point of orientation– (Husserl 2001). These analyses reveal how embodied hermeneutic injustice operates through severed relationships between bodies and

territories. Through extractive practices, women experience not only material harm but also a disruption of embodied meanings tied to their relationship with the land. This expanded framework suggests that addressing hermeneutic injustice requires attention to both conceptual resources and the spatial-material dimensions through which social meanings become incarnated in specific contexts.

Conclusion

This paper analyzes how hermeneutic injustice manifests not only through gaps in collective understanding but also through the embodied experience of social meanings. By examining female bodily experiences, we see how hermeneutic injustice operates at the level of lived corporeality, fundamentally shaping how women experience their bodies as inhibited intentionality. Young's phenomenological insights reveal that collective hermeneutical resources are not merely tools for interpretation but active forces that constitute embodied subjectivity.

This reconceptualization advances our understanding in several ways. First, it shows how hermeneutic injustice can operate through embodied experience even when relevant concepts exist - the problem is not just conceptual absence but how existing concepts structure experience itself. Second, it reveals how social meanings actively constitute rather than merely interpret bodily experience. Third, it demonstrates the need for an intersectional approach, considering how multiple meaning-making structures mediate embodied experience.

This reconceptualization suggests that addressing hermeneutic injustice requires attention to both conceptual frameworks and the body constitution through which social meanings become incarnated. The dynamic between interpretive resources and embodied experience shown here opens new possibilities for understanding how marginalization operates through the body while indicating pathways for resistance and transformation.

References

- Ahmed, Sara (2006) *Queer Phenomenology. Orientations, Objects, Others*. Durham: Duke University Press.
- Alcoff, Linda M. (2006) *Visible Identities: Race, Gender, and the Self*. Oxford: Oxford University Press.

- Alcoff, Linda M. (2018) *Rape and Resistance*. Cambridge: Polity.
- Al-Saji, Alia (2014) "A Phenomenology of Hesitation", in: Emily S. Lee (ed.) *Living Alterities. Phenomenology, Embodiment, and Race*, NY: Suny Press, 133-172.
- Cabnal, Lorena (2010) "Acercamiento a la construcción de la propuesta de pensamiento epistémico de las mujeres indígenas feministas comunitarias de Abya Yala", in: *Feminismos diversos: el feminismo comunitario*, Acsur-Las Segovias, 11-25.
- Falbo, Arianna (2022) "Hermeneutical Injustice: Distortion and Conceptual Aptness", *Hypatia*.
- Fanon, Frantz (1967) *Black Skin, White Masks*, trans. Charles Lam Markmann, New York: Grove Press.
- Fricker, Miranda (2007) *Epistemic Injustice: Power and Ethics of Knowing*. NY: Oxford University Press.
- Fricker, Miranda (2016) "Epistemic Injustice and the Preservation of Ignorance", in: Rick Peels and Martijn Blaauw (eds.) *The Epistemic Dimensions of Ignorance*, Cambridge: Cambridge University Press, 160-177.
- Haslanger, Sally (2013) "Social Meaning and Philosophical Method", American Philosophical Association 110th Eastern Division Annual Meeting.
- Haslanger, Sally (2019) "Cognition as a Social Skill", *Australasian Philosophical Review* 3(1), 5-25.
- Husserl, Edmund (2001) "Foundational investigations on the phenomenological origin of the spatiality of Nature: the originary ark, the Earth, does not move", in: Leonard Lawlor and Bettina Bergo (eds.) *Husserl at the limits of phenomenology*, Evanston: Northwestern University Press.
- Medina, José; Henning, Tempest (2021) "My Body as a Witness", in: Jennifer Lackey (ed.) *Applied Epistemology*, NY: Oxford University Press, 171-190.
- Merleau-Ponty, Maurice (2012) *Phenomenology of Perception*, trans. Donald A. Langdes, London: Routledge.
- Paredes, Julieta (2015) "Despatriarcalización. Una respuesta categórica del feminismo comunitario (descolonizando la vida)", *Revista de Estudios Bolivianos*.
- Salamon, Gayle (2017) "Phenomenologies of Relation: Re-Worlding Gender with Iris Marion Young", *Symposium* 21(2), 44-62.
- Whitney, Shiloh (2018) "Affective Intentionality and Affective Injustice: Merleau-Ponty and Fanon on the Body Schema as a Theory of Affect", *The Southern Journal of Philosophy* 56(4), 488-515.

Young, Iris Marion (1998) "Throwing Like A Girl: Twenty years Later", in: Welton (ed.) *Body and Flesh: A Philosophical Reader*, Cambridge: Blackwell, 286-290.

Young, Iris Marion (2005) *On Female Body Experience. "Throwing Like a Girl" and Other Essays*, Oxford: Oxford University Press.

Gender in Psychoanalysis. A brief note referring to Wittgenstein

Ulrike Kadi (Vienna, Austria)

Abstract

In her critique of the concept of gender, Alenka Zupančič argues for the special epistemological significance of sexuality. According to Zupančič, the position of sexuality between ontology and epistemology forces us to consider difference. In her view, the recognition of gender identities is not very helpful. Her concept of identity is inspired by Jacques Lacan's text on the mirror stage. In this famous essay, Lacan not only addresses the role of visual impressions in the formation of identity, but also emphasizes deceptive and denying aspects of identity in a generalizing way.

My contribution is directed towards a queer-theoretical extension of psychoanalytical approaches to gender. Ludwig Wittgenstein criticizes a common form of generalization, referring to Galton's composite photographs. He thus draws attention to processes in the formation of concepts such as identity that Zupančič overlooks. The example of a composite image of Wittgenstein and his three sisters illustrates these neglected processes in the attribution of gender identities.

During the last century, many approaches to gender have been developed using psychoanalytic concepts (Hagemann-White 1979). For example, the term “gender” has a psychoanalytically influenced history in the wake of Robert Stoller's (1985) work. Difference feminist approaches and their many forms, often interwoven with literary studies, philosophy or aesthetics (e.g. Kristeva 1989, Irigaray 1996, Bersani 1986), are based in one way or another on psychoanalytic knowledge. In particular, Judith Butler's (1990, 1993) post-structuralist project of softening gender norms was embedded in a psychoanalytic framework. On the other hand, debates about (female) gender and its meanings have accompanied the theoretical development of psychoanalysis from the very beginning (Mitchell 1974/2000, Chasseguet-Smirgel 1976). Starting from the well-known androcentrism of Freud's theorems, psychoanalysts have continuously argued about the meaning of femininity, the mother, the father, siblings, male and female homosexuality, and the questionable universality of the Oedipus or castration complex.

Given the outwardly effective theoretical fruitfulness of its assumptions, the virtual disappearance of psychoanalysis from current discussions of gender theory may come as a surprise. Superficially, this development is also related to the concept of “gender”. In addition to efforts to connect psychoanalysis with current gender discourses (e.g. Gylter 2010, Hutfless, Zach 2017), a critical

distance from the concept of gender is not uncommon in psychoanalysis (see Reiche 1997, Copjec 2014a). A bouquet of psychoanalytic responses to positions on gender can thus be compiled today, which limits the ability of psychoanalysis to form alliances with other gender-political endeavours.

Of course, looking at the discussions within psychoanalysis, it is clear that there is a struggle about gender, gender norms, their genesis and validity that does not only concern psychoanalysts. It echoes many of the controversies that characterise the political discourse surrounding the prohibition of gendered language practices (Butler 2024). Inside and outside psychoanalysis, proponents of a normative and/or biologically based binary gender are pitted against proponents of gender diversity that is more closely tied to an individual's self-image.

My following reflections on the question of gender identity are situated on the borderline between psychoanalysis and gender theory. In a small borderline dialogue, I make a proposal based on Ludwig Wittgenstein's remarks on composite photography. My reflections follow a controversy that contributes to current psychoanalytic discussions on the understanding of identity.

1 The primacy of difference

With Sally Haslanger (2000: 32f.), three questions can be asked if we are philosophically interested in gender as a concept: What is our concept (1)? What does it refer to (2) and why is it necessary (3)? The first question, the clarification of the concept, aims, among other things, to ensure that a new concept within an existing theory does not lead to unresolvable contradictions. In this sense, the contradictions inherent in psychoanalysis will be discussed below. The limitation to this aspect would be misunderstood if it were read as an indication of a lack of psychoanalytic contributions to the other two questions. With regard to the extension of the concept of gender, for example, the great influence of fantasy on perception must be emphasised from a psychoanalytic perspective. Psychologically, this leads to a prioritisation of the fantasmatic anatomy over a biologically based anatomy (see Laplanche 2017). With regard to the third question of strategies in the field of the socio-cultural establishment and maintenance of the power of gender norms, intensive research is underway within psychoanalysis (see Allouch 2007, Laufer 2020).

Like many adherents of structural psychoanalysis, philosopher Alenka Zupančič is critical of the concept of gender. This criticism primarily concerns gender as a shorthand for a multiplicity of genders. Zupančič belongs to the so-called Ljubljana School of psychoanalysis, which emerged largely outside of clinical practice in the wake of Slavoj Žižek or Mladen Dolar and their reception of Jacques Lacan. Zupančič refers to the Freudian beginnings of psychoanalysis, when sexuality and the multiple manifestations of its symptomatic defence were at the centre of clinic and theory. She emphasises that sexuality does not only constitute the uniqueness of psychoanalysis in the scientific field (cf. Zupančič 2022: 302). She also sees sexuality as a short-circuit between ontology and epistemology (cf. *ibid.*: 307). This is why she claims that Freud's discovery does not concern empirical sexuality as such, “but its problem, its negative core and the role this core plays in the escalation of the sexual” (*ibid.*, my translation). This negative core makes sexuality, and with it sex, an unrecognisable element that forms the fulcrum of the subject (cf. *ibid.*: 317). Consequently, it is sex and not gender that matters. Because “sex forces us to think difference” (*ibid.*: 318, my translation). According to Zupančič stressing the recognition of gender identities does not lead very far. This can already be seen in the fact that the concept of gender has brought women little in the way of fundamental improvements in their social position (cf. *ibid.*: 319).

Although these descriptions with Wittgenstein could be seen as a form of “metaphysical puzzlement” (TS 309: 92), I do not deny the subtlety of Zupančič's theorem, which is fed by Kantian, Heideggerian and Lacanian motifs. The primacy of difference corresponds to an accentuation of alterity that is indispensable for the relational content of psychoanalysis. In structural psychoanalytic terms, this is the primacy of language or – as the associated vocabulary puts it – of the symbolic order. The primacy of negativity can also be found in other psychoanalytic approaches. It emphasises the importance of castration, which metaphorically denotes a set of boundaries and limits that are necessary for every human being to grow up.

Remarkably, the conceptual web woven with these components helped to break the rigid link between biological anatomy and gender in the mid-twentieth century. However, the arguments about (non-)being, identity, difference, sexuality and sexuation have since been transformed into an isolationist vote in favour of a binary gender order, not only by Zupančič, but

also in some other post-Lacanian approaches (see Copjec 2014b, Soiland et al 2022). This gives the impression that a bivalent logic of being and non-being should be reflected in an order of two genders. Women and all subjects who position themselves as women in the symbolic order should recognise themselves as governors of negativity, of nothingness (cf. Zupančič 2022: 319). The negative is highly valued. This is reminiscent of the rituals of courtly love, in which gender roles were clearly divided. At the same time, such speculation also puts a stop to thinking about a queer diversity of genders.

2 Nodes in the network of family resemblances

A special feature of conceptual psychoanalytic contributions (not only to the concept of gender) is their dual practical relevance. In addition to the strategic task of defining a politically applicable concept of gender, the psychoanalytical treatment clinic opens up a space in which all concepts – and thus “gender” and “identity” – have to demonstrate their worth. Such a description of the task, however, primarily conceals the difficulties involved. For it remains unclear where each term derives its meaning and what it refers to.

Zupančič (cf. 2022: 300) emphasises that the psychoanalytic clinic should not be regarded as a 'laboratory' for the formation of concepts. It is appropriate to have some reservations here, because the clinic is where the relational core of psychoanalysis appears in vivo, enabling a conceptual co-construction between the psychoanalyst and analysand to describe experiences. Conversely, the importation of general philosophical concepts such as difference, identity or negativity into psychoanalytic practice presents several difficulties. Firstly, practising psychoanalysts are confronted with a multitude of translation problems between everyday language and philosophical expressions, which can sometimes be overwhelming. Secondly, under these circumstances, jargon easily arises that lacks precise knowledge of the terms (see Kadi 2025). This can lead to misunderstandings and erroneous conclusions, such as the prescription of a psychoanalytic norm of bisexuality, as described above.

Such problems are familiar to those engaged in a philosophical practice that refers to Wittgenstein. Here, the difficulty is not limited to the use of philosophical terminology. Ultimately, the problems concern all language. Where do the individual words that we string together as sentences come from? What do they aim at? Wittgenstein identifies a significant challenge in

creating larger units from smaller ones, which he attributes to our “craving for generality” (TS 309:27, and elsewhere). He associates this desire with a fixation on the scientific approach developed in, among other things, the natural sciences. Incidentally, Wittgenstein draws attention to Galton's composite images, which make moments of generalisation identifiable.

In the 19th century, Francis Galton developed a technique for representing multiple photographs in a single image through sequential multiple exposures of a photographic plate (see also Gruber 2021, Lee-Morison 2019: 86-99). Galton first applied his new method to landscapes, which could be depicted as they changed over time. Later, he used his technique to merge photographs of multiple faces. This led to several applications that are still important today: Composite photographs have been used to find averages of statistical normality in the study of pathological or criminal deviance. In addition, Galton's work in the search for prototypical faces continues to influence techniques for automatic face recognition.

There is a composite picture of Wittgenstein and his three sisters. The circumstances surrounding its production with the help of photographer Moritz Nähr remain unclear to this day (see Gruber 2021). The four faces can be seen both individually and as a single face, representing a “puzzle for perception” (Lee-Morison 2019, p. 105). While the common features from all four images are condensed and clearly defined, dim transition zones emerge where different pictorial elements are mixed, sometimes revealing more and sometimes revealing less. Prototypical elements can be revealed by removing the veiled parts.

Wittgenstein questions such a procedure for our language. Our questionable search for the general leads us to look for the common parts, the one and only overall picture in the composite photograph. Like Galton, we collect the same features and use the same linguistic expression to describe things that correspond in terms of that feature (TS 309: 27). Wittgenstein's examples are the terms “game” and “leaf”. We use them as if there were only one kind of leaf or game.

In his critique, Wittgenstein makes a clear shift in focus: it is not enough to pay attention to what we have in common. It is essential not to cut away the parts

where there is nothing in common. This poses a significant challenge, given the imprecision of sensory impressions. It is not possible for us to perceive all parts with absolute precision. We have to deal with a “Verschwommenheit, Unbestimmtheit unserer Sinneseindrücke” (TS 211: 746). The objective of our “craving for generality” (TS 309: 27) is to eliminate this indeterminacy. Wittgenstein, however, takes indeterminacy as the starting point of an oscillating movement. The quick connection to an overall picture that is common to all obscures the view of the individual pictures. It categorically ignores the zone of variation and possibility that arises in Galton's photography where there is no correspondence between all the images. Wittgenstein calls Galton's photography the „Bild einer Wahrscheinlichkeit“ (TS 211: 747). It is made up of a „komplizierte[n] Netz von Ähnlichkeiten, die einander übergreifen und kreuzen“ (PI 1971: §66), of family resemblances.

Generalisation is not the result of a process of reduction; it is the result of a movement that owes itself to indeterminacy. The tendency to arrive at something general through reduction is reflected in the surviving composite photograph of the Wittgenstein siblings. As Andreas Gruber (2021) rightly points out, the surviving version can hardly contain Ludwig's image with the same intensity as his sisters'. Perhaps a less exposed version of his face was inserted. Alternatively, the indeterminate areas around the nose and head of the composite figure may have been removed through retouching.

In our context, it is crucial to address the gender of the mixed figure. Looking at the individual images, should we say that she is three-quarters female and one-third male? Or should we refer to an overall visual impression and postulate a female gender in accordance with the binary norm that prevailed in Freud's time? Wittgenstein's reading of the photographs suggests a third approach. It sharpens the eye for the differences, as opposed to the uniform centre that is all too quickly sought. The gender of the face in the mixed image becomes less clearly legible. References, similarities and differences between the individual images are given greater weight. The relationship of the face to the pearl necklaces, and the conventions for pearl necklace wearers in Wittgenstein's time, indicate a female gender. However, in relation to the philosopher brother, who is also included in the picture, it is likely to be a gender other than female. The sister Hermine's features, perhaps post-

menopausal and therefore more masculine, might inspire us to look for new terms for types that do not fit easily into given labelling systems.

If we think of gender as a network-like structure of similarities, we need to consider more parameters. The question of gender identity thus becomes more complex than Zupančič suggests in her rejection of the concept of identity. Gender would become gender in relation to x. It is different to be considered a woman in relation to prevailing notions of social normality than in relation to the visual impression someone makes. Another point of reference would be, for example, self-attributions, which can develop and change throughout life.

It also gives more weight to the differences in each case. This fits in with a theory of gender that is oriented towards both structural psychoanalysis and queer theory. Apart from a biological anatomy and the social assignment of a sex, gender is a complex interplay of individual experiences and fantasies. Psychologically, gender involves more nodes in the network of family resemblance than a simplistic two-valued logic would suggest. Experiences, identifications and disidentifications with people from one's early social circle are just as important as different parts of a phantasmatic anatomy in which infantile fantasies can be found as well as unprocessed fragments of traumatic experiences.

3 Consideration of marginal zones in the mirror

How does Zupančič arrive at her rigid concept of identity? Identity has a canonical place in structural psychoanalysis: Lacan's text on *The Mirror-phase as formative of the Function of the I* (Lacan 1949/2016). In the context of family resemblances, Wittgenstein recommends not thinking but looking (PI 1971: §66). Lacan takes an example from optics. He describes a young child's first confrontations with its own reflection and its visual perception of early caregivers as sites of identification, i.e. the process that corresponds to the emergence of an ego identity. At this early stage, one's own body is only experienced in fragments. The mirror and the perceived figures of others give rise to a fantasy that serves as a defence against the irritating early experience of helplessness and lack of motor coordination. According to the theory, the images of one's own body in the mirror (or in the eyes of others) form the basis of fantasies of perfection, which, like an externally applied orthopaedic armour, help to deny the experience of limitation due to immaturity of motor

function. The ego identity that refers to the experience of the mirror is generally considered to be deceptive, not real, denying.

The desire for generality stressed by Wittgenstein is present throughout the Lacanian text. And the reference to some laws of nature described in the 1930s (Köhler's monkey experiments, the fetalisation hypothesis) can also be found in Lacan. With regard to the question of gender, the lack of reference to gender in the mirror throughout the text is surprising, since the confrontation with the image of one's own body is psychologically part of gender development (Kadi 2013).

Despite this omission, the theory of the mirror stage addresses key issues of identity formation. These include the observation that the formation of an ego identity begins early in life or that it arises in a field of tension. Every identity formation is an experience that goes beyond conscious attitudes and has its own unconscious parts. Not only language and speech are important for the development of the mental apparatus. Visual experiences often have an almost transformative power in the process of acquiring an identity (cf. Lacan 1949/2016, 94).

According to Lacan's account, these and other moments mentioned in the text belong to zones at the centre of a general concept of identity. In this way, identity becomes associated with deception, a lack of reference to reality, and a denial of one's own need for help. However, such moments override the other zones of the composite image. If Zupančič and other post-Lacans equate identity with fantasies of perfection, this would deny difference. The veiled part of the image thematised by Wittgenstein with Galton's photographs, and subsequently with his reflections on family resemblance, would be overlooked. Standing in front of a mirror, one perceives more than can be found in the generalisable zone of the Lacanian mirror. Alongside singular visual experiences that do not fit into the generalised picture, it is also important to mention blurred sensory impressions and incomprehensible, enigmatic aspects, as well as the fantasmatic attempts to unravel them. It is these many small differences that determine the image of each individual ego identity.

4 Postscript

The psychoanalytic critique of the concept of gender is intertwined with reservations about the concept of identity. The zones in Galton's composite photographs that are not common to all images can reveal these reservations as the result of a questionable approach to the general. In Wittgenstein's later reflections, the veiled areas are transformed into short fibres: „[W]ir dehnen unseren Begriff aus wie wir beim Spinnen Faser an Faser drehen. Und die Stärke des Fadens liegt *nicht darin daß eine Faser durch seine ganze Länge läuft, sondern darin daß sich viele Fasern übergreifen*“ (Ms 152: 74). Why not think of gender dynamically as a series of short threads that are spun in and out to form a concept that a person develops of their gender?

References

- Allouch, Jean (2007) *La psychanalyse est-elle un exercice spirituel? Réponse à Michel Foucault*, Paris: Epel.
- Bersani, Leo (1986) *The Freudian Body: Psychoanalysis and Art*, New York: Columbia University Press.
- Butler, Judith (1990) *Gender Trouble. Feminism and the Subversion of Identity*, London: Routledge.
- Butler, Judith (1993) *Bodies that matter. On the discursive limits of sex*, London: Routledge.
- Butler, Judith (2024) *Who's afraid of Gender?* London: Penguin.
- Chasseguet-Smirgel, Janine (1976) *Psychoanalyse der weiblichen Sexualität*, Frankfurt/M.: Suhrkamp.
- Copjec, Joan (2014a) "The Inheritance of Potentiality. An Interview with Joan Copjec", *Revue électronique d'études sur le monde anglophone*, <https://journals.openedition.org/erea/4102> (25.3.2025)
- Copjec, Joan (2014b) "The Return to Two", *Political Concepts. A Critical Lexicon*. <https://www.politicalconcepts.org/sexual-difference-joan-copjec/8/> (24.3.2025).
- Gyler, Louise (2010) *The Gendered Unconscious: Can Gender Discourses Subvert Psychoanalysis?* London: Routledge.

Gruber, Andreas (2021) "Das Wittgenstein-Kompositportrait Reconstructed", in: Uwe Schlögl; Sandra Tretter; Peter Weinhäupl (eds.) *Moriz Nähr (1859-1945). Fotograf für Habsburg, Klimt und Wittgenstein*. Wien: Catalogue Raisonné.

Hageman-White, Carol (1979) *Frauenbewegung und Psychoanalyse*, Frankfurt/M.: Stroemfeld / Roter Stern.

Haslanger, Sally (2000) "Gender and Race: (What) Are They? (What) Do We Want Them To Be?" *Nous* 34/1, 31-55.

Hutfless, Esther; Zach Barbara (eds.) (2017) *Queering Psychoanalysis. Psychoanalyse und Queer Theory – Transdisziplinäre Verschränkungen*, Wien: Zaglossus.

Irigaray, Luce (1996) *Speculum: Spiegel des anderen Geschlechts*, Frankfurt/M.: Suhrkamp.

Kadi, Ulrike (2013) "Das Geschlecht des Bildaktes", *Zeitschrift für psychoanalytische Theorie und Praxis* 28/1, 61-74.

Kadi, Ulrike (2025) "Von der Codevermittlung zur Übersetzung: Lacan lehren – nicht ohne Laplanche", *RISS* 101, <https://zeitschrift.riss-psychoanalyse.de> (1.5.2025).

Kristeva, Julia (1989) *Geschichten von der Liebe*, Frankfurt/M.: Suhrkamp.

Lacan, Jacques (1949/2016) "Das Spiegelstadium als Gestalter der Funktion des Ichs, so wie sie uns in der psychoanalytischen Erfahrung offenbart wird", in: Jacques Lacan (2016) *Schriften I*, Wien: turia+kant, 109-117.

Laplanche, Jean (2017) "Gender, Geschlecht und Sexual" [2003], in: Jean Laplanche *Sexual*, Gießen: Psychosozial-Verlag, 137-171.

Laufer, Laurie (2020) "Michel Foucault: The Queer Gender for Psychoanalysis?" *Psychoanalytic Inquiry* 40, 579-590.

Lee-Morison, Lila (2019) *Portraits of Automated Facial Recognition. On Machinic Ways of Seeing the Face*, Bielefeld: transcript.

Mitchell, Juliet (1974/2000) *Psychoanalysis and Feminism*, London: Penguin.

Reiche, Reimut (1997) "Gender ohne Sex. Geschichte, Funktion und Funktionswandel des Begriffs 'Gender'", in: *Psyche* 51(9-10), 926-957.

Soiland, Tove; Frühauf, Marie; Hartmann, Anna (eds.) (2022) *Sexuelle Differenz in der post-ödipalen Gesellschaft*, Wien: turia+kant.

Stoller, Robert (1985) *Presentations of Gender*, New Haven: Yale University Press.

The Psychodynamic Formulation Collective (2022) "The Effects of Culture and Society", in: The Psychodynamic Formulation Collective *Psychodynamic Formulation. An expanded approach*, New York: Wiley-Blackwell, 193-204.

Wittgenstein, Ludwig *Bergen Electronic Edition*. Ed. Wittgenstein Archives at the University of Bergen. References to the Nachlass are in the form "Manuscript number: page", according to the Bergen Electronic Edition. <http://www.wittgensteinsource.org/> (5.3.2025).

Wittgenstein, Ludwig (1971) *Philosophische Untersuchungen* [PI], Frankfurt/M.: Suhrkamp.

Zupančič, Alenka (2022) "Philosophie oder Psychoanalyse? Ja, bitte!" In: Tove Soiland; Marie Frühauf; Anna Hartmann (eds.) *Sexuelle Differenz in der post-ödipalen Gesellschaft*, Wien: turia+kant, 299-327.

Anscombe's Social Epistemology

André Kfouri (Lisbon, Portugal)

Abstract

Anscombe is not traditionally read as an epistemologist, but her writings contain several insights that could be valuable to current debates about belief and knowledge. I will present some of these points, with a particular focus on her works that suggest a social emphasis for knowledge in a way that is relevant to the landscape of testimony and social epistemology, as well as being deeply Wittgensteinian. In the first section, I will present her lesser-known writings on testimony, detailing her classification of testifiers in terms of authority and originality, as well as the role of trust. In the second and third sections, I will suggest that her account of brute facts and institutions carry significant implications as to how knowledge is shared within a broad community and assumed in descriptive language. I will conclude by presenting a reading of the brute facts account as a structure of questioning and justification, helpful in determining whether particular instances of such epistemic games are relevant to wider practices. Throughout, I will highlight how Anscombe's thoughts on knowledge are deeply connected to her broader effort to expose the issues with description that lie behind many philosophical problems.

Believing Testimony

The most explicitly epistemological contribution in Anscombe's lesser known works is her exploration of testimony in a pair of papers in the 1970s. I will now go over some of their key points, arguing that Anscombe's work foreshadows some key developments and positions in the current epistemology of testimony, particularly as developed by Greco (2021). Namely, her account of degrees of originality and authority will be presented as analogous to his central argument for the irreducibility of knowledge transmission to generation; and her account of the relation between the speaker's authority and the listener's trust can be compared to his description of the transmission of knowledge as joint agency, which plays a key part in the lively debates on the role of trust in testimony.

A first noteworthy point is that Anscombe recognizes the importance of testimony for epistemology at all, and that she presents it as distinct from and irreducible to other sources, such as causality: “We must acknowledge testimony as giving us our larger world in no smaller degree, or even in a greater degree, than the relation of cause and effect; and believing it is quite dissimilar in structure from belief in causes and effects” (Anscombe 1979b: 26). Beyond this general recognition, Anscombe's main concern in these papers is

what she calls 'belief with a personal object' — believing someone, as opposed to believing in someone's existence or believing a statement — expressible through the form "believing x that p" (1979b: 25). Ultimately, she analyzes a case of believing someone as "trusting him for the truth — in the particular case" (Anscombe 1976b: 32), given the presupposition of a certain set of conditions for attribution and credibility, which include the proper perception and interpretation of the transmitted message along with its attribution to the correct person and understanding it as addressed to the perceiver.

The presuppositions are taken to occupy the same place as the traditional notion of *preambulae fidei*, or preambles of faith, in Thomistic philosophy (see ST I.2.2 ad. 1). With this equivalence, she denies significant structural distinctions between the testimonial 'believing someone' and the religious 'believing God': the former is what she calls 'human faith', and the latter 'divine faith', but the properties she describes purportedly hold for both. This proximity makes clear the extent of the philosophical implications of Anscombe's roots in the Catholic tradition, but also sheds light on its usefulness for current explorations of the topic: Greco (2021: 161-3) argues that the 'externalist turn' in religious epistemology must now be followed by a 'social turn', and makes the point that such a move is only natural given the central role that testimony plays in various religions, especially in the transmission of revelation in the Abrahamic faiths. Anscombe's framing of *preambulae fidei* as preconditions for belief in testimony in general goes further in suggesting that the problems of epistemology of testimony already have a rich history even in the western philosophical tradition, and that present day epistemologists could benefit from the study of these features of scholastic accounts.

Given that background, I can now present the main features of her account that can be compared to ones in current debates. These revolve around the classification of participants in testimony in terms of authority and originality. A subject who simply transmits received information may be an authority on it, but not an original one. They're 'an original authority' if the information includes some contribution on their part, such as witnessing an event or some personal inference, and they're 'the total original authority' if what they transmit doesn't rely on any received information (Anscombe, 1979b: 28-9). To use her examples, someone who is told that Leonardo made drawings of a

flying machine and goes on to tell a friend is not an original authority at all, but is just passing along received information. If they saw the drawings, they're a partially original authority, because they contribute their personal experience but still rely on a scholarly tradition that identified the drawings and their author. If they talk instead about eating an apple this morning they're a totally original authority, given that they're simply drawing on their own experiences and basic linguistic competence. With this, she gives an early version of a distinction that is still key to current debates on testimony, namely that between generators and transmitters of knowledge, to use Greco's (2021) terminology. We can chart the position of the speaker in a particular case of testimony as a point along the axes of authority and testimony, and go over each in turn.

The axis of authority is a broad heading that includes two senses of the word, as superiority both in terms of power relations and of expertise. Either way, an authoritative subject has what Anscombe (1962a: 44) elsewhere calls a “right to demand belief”, which doesn't, of course, necessarily imply unquestioning belief, and, given the fallibility of teachers or any transmitters of information, isn't a “right to be believed” either. The upper limit of authority falls short of totality, given that it must always be met by the receiver's trust at the other end of the transmission relation. Without it, any piece of testimony may not be believed, even coming from a high authority in both senses. This emphasis on trust as the final stop in believing someone, even if it's not described in much detail by Anscombe, places her alongside Greco's joint agency account in arguing, contra Lackey (2008: 238-40) that the matter of trust is intrinsic to the discussion of testimony (see Greco 2021: 47-67).

On the other hand, the axis of originality highlights the differences between imparting information about one's personal experience and partial or total reliance on “traditions of information”. In contrast to the authority axis, one may in fact be totally original, as happens often with ordinary tellings of our own experiences. These depend only on linguistic competence — being “taught the concept” in learning to use language in everyday life” (Anscombe 1979b: 29) — which is taken for granted by both sides of testimony and thus doesn't impair the speaker's originality. Here Anscombe is evidently drawing from Wittgenstein's remarks on conceptual mastery (PI: §§20, 31-33, 145-150, 199, 338, 508; OC: §§526-7, 534), taking insights that Greco (2021: 103-25) would

also go on to incorporate into his account by presenting hinge commitments as common knowledge. The contrast with traditions of information seems to be her own contribution, and it may be read as a way of working out the puzzlement in cases such as a child's learning natural history expressed in OC §534: we are comfortable with speaking of knowing in cases that rely on such traditions, not so much when they rely on everyday language. As Anscombe suggests elsewhere, knowledge derived from widespread traditions should hardly count as testimony, but rather “as if I had been taught to join in doing something [...] for here knowledge is no other than certainly correct belief in pursuit of a practice” (Anscombe 2015: 189). It is in this shared acting that we find the second part of her social epistemology.

The Relation of Bruteness

The next piece of this reconstruction of Anscombe's social epistemology goes beyond testimony and into what is widely known or assumed in social groups, using her concepts of brute facts, institutions and normal circumstances. I will introduce her account of brute facts by contrasting it to the more common use of the term as promoted by Searle, arguing that her version is not vulnerable to some criticisms that have been posed against that variant. Then, I will present a reading of Anscombe's account as providing a solid ground for many epistemic practices by describing a structure of justification, where the fact that our statements about the world rely on unjustified assumptions is not only unproblematic but is in fact a necessary part of the workings of descriptive language. These background features can still be contested and justified, but only when relevant to established practices and moral concerns.

I will start by presenting what Anscombe's brute facts are not. Searle (1965, 1995) sets out a binary account of ‘brute facts’ — “matters of brute physics and biology” — as opposed to ‘institutional facts’ that are “matters of culture and society” (Searle 1995: 27). He concedes that any fact can only be stated within the institutions of language, but counters that the brute facts themselves exist independently. In his account, any institutional fact depends on a logically prior brute fact, and, therefore, holding all facts as institutional would cause a problem of infinite regress or circularity (Searle 1995: 34-5; 55-6; 191).

Searle's key mistake, seen from Anscombe's Wittgensteinian approach, is presenting brute facts as a physicalist foundation isolated from the

institutional character of language itself. Our talk about nature does, of course, often mention things that are independent of human constructions, but any description we can give of them still depends on language and other practices, and those can parse out phenomena in significantly different ways. Searle's (1995: 27) example of the ninety-three million mile distance between the Sun and the Earth as a paradigmatic brute fact illustrates this well: not only does it rely on the 'institution of measuring', as Wittgenstein would say it (RFM III: 36; see PI: §242) and a number of procedures, but, crucially, even the very general fact that there is a distance between the two bodies is still a construction — the average distance of an elliptic orbit — which could have been different according to the needs of our astronomy. The same point holds for more immediate examples: as Anscombe (1976: 177) says, horses would still exist if humans weren't around, but the fact that a horse is a horse and a donkey is not is up to our classifications, which themselves are integrated into our linguistic and extralinguistic activities.

The difficulty in these cases is what Anscombe (1976: 177) refers to as the "narrow channel" between idealism and realism, or as Wittgenstein's own enterprise of "realism without empiricism" (id.: 133; RFM IV: 23). Searle seems to fall into realist empiricism by these metrics, but, on the other hand, is also vulnerable to the criticism of not being enough of a realist insofar as he defends the equal ontological footing of brute and institutional facts, with both serving as components of our "one world" (Searle 1995: ix). On that basis, Dörge and Holweger (2021) argue that institutional facts, as Searle defines them, don't meet some significant realist criteria of existence, such as representation-independence, causal integration and spatiotemporality. Anscombe's account, I hope to show, resists such criticisms by integrating brute facts and institutions so deeply into a broader account of descriptive language that they can't be intelligibly separated.

On to Anscombe's account. The key contrast with the brute-institutional fact dichotomy is that she treats bruteness as a relative property, where, for every description, there may be one of the same events that is more brute or more complex. To use her famous example, that Anscombe's grocer supplied her with potatoes is already quite a complex description, and the facts that he carried a bag of potatoes to her house and left a bill are brute in relation to it. In turn, the description of supplying is brute in relation to the ones that she

owes the grocer money, that she is insolvent, and so on. The difference in bruteness and complexity is one of degrees, and is marked by the background consideration of additional institutions: the more complex description of insolvency requires our institutions of buying, selling and owing, as well as some negative evaluation of debtors; but even the relatively brute fact that the grocer carried a bag of potatoes assumes a number of institutions relating to the activities of a grocer, the act of carrying, the function of a bag, and vegetable identification. Searle's concern for an infinite regress is irrelevant to this account, even if explanations don't bottom out at absolutely brute facts, simply because explanations come to an end, and procedures of sensible questioning and justification are constrained by their relevance to the rest of our lives: "The rock-bottom explanation [...] is the procedures themselves" (Anscombe 1976: 116).

In short, for a given description A there is a set of facts xyz that are "brute relative to the fact described by A" (Anscombe 1958: 24). In addition to the brute facts xyz, the complex description A also has additional institutions behind it which aren't assumed by xyz and come with their own assumptions of normality. It's in these last two features that we will find the most points of epistemological interest in Anscombe's account.

Institutions and Normality

The key to linking Anscombe's account to its Wittgensteinian roots is the concept of an institution. The example of money, found both in Anscombe (1958: 22) and Wittgenstein (PI: 120, 142; RFM I: 143-150), as well as in Searle (1995: 2, 27-9), may give the impression that what these authors mean as 'institutions' are simply the social organizations and traditions to which we would ordinarily apply the term. That is partially the case for Searle — one of the points critiqued by Dörge and Holweger (2021: 4954) is that his definition of an institution is too narrow for the broad class of institutional facts — but is misleading for the others. In Wittgenstein's writings, the term was generally associated with rules and rule-following, and he variously referred to language (MS 124: 188-9) and to rules (MS 124: 63r-v; MS 129: 183; MS 180a: 36r; RFM: §334) as institutions. These instances, as argued by Baker and Hacker (2009: 122), don't suggest the identification of Wittgenstein's use of 'institutions' with 'social institutions', but rather with 'well-established practices' or with "established regularities apprehended as uniformities, that guide

conduct” (2009: 143ff2). Of course, this doesn't exclude social institutions in the ordinary sense from the concept, but instead suggests a much broader category. The question of whether an institution can exist outside of a community, as they suggest, or is necessarily a communitarian concept, as defended by Bloor (1997: 27-42), is not entirely relevant to the present purposes: suffice it to say that Anscombe (1958: 22-3) seems to only refer to them under the social aspect, speaking of ‘our institutions’ and ascribing them to a society. We can still safely say that they cover all shared normative linguistic and extralinguistic practices of the kind that is needed in order to use any sort of descriptive or epistemic language intelligibly and usefully. Anscombe particularly emphasizes the range beyond language: “‘Linguistic practice’ here does not mean merely the production of words properly arranged into sentences on occasions we vaguely call ‘suitable’. It is important to say that it includes activities other than the production of language, into which the use of language is interwoven” (Anscombe 1976: 117).

A valid description, however, can't only require that the relevant institutions exist and are shared by the people involved; it also assumes some kind of normality, without which the same facts and institutions could hold and yet the description could not, such as in staged scenarios or cases of deception — to use Anscombe's example, an actor playing a grocer isn't really selling anything. This normality is not a general norm of societal functioning, but rather a context of ‘normal procedure’ for each of the institutions at play (Anscombe 1958: 23). This makes it impossible to demarcate the exact bounds of normal procedure, but that's not a considerable obstacle to the brute facts account: as she argues, “it is not theoretically possible to make provision in advance for the exception of extraordinary cases; for one can theoretically always suppose a further special context for each special context, which puts it in a new light” (Anscombe 1958: 23). Special cases must therefore be dealt with as they happen, but their possibility doesn't impair the description of normal procedure and the philosophical endeavor of preparing for them in advance would seem to be unproductive — a point that seems ever more insightful in the light of the efforts in limiting the reach of defeaters in post-Gettier epistemology.

At the Wittgensteinian core of the account is the point that the relations between descriptions and their underlying factors are grammatical rather

than inferential, and therefore complex descriptions are made by competent speakers with no mediating consideration of the underlying facts and institutions. To take a remark by Anscombe on practical syllogisms, if the “[...] account were supposed to describe actual mental processes, it would in general be quite absurd” (Anscombe 1957: §42). The layers of facts and institutions are only revealed when a situation is analyzed in depth, which happens in ordinary cases of disagreement.

I propose that this central aspect of Anscombe's account is best captured by a reading that takes the four key concepts in “On Brute Facts” — complex descriptions, brute facts, normal circumstances and institutions — as a structure of instances of appeal in the game of questioning and justification of a normal description. In most cases, an adequate description is understood by every competent speaker of that language and doesn't require any justification or consideration of other factors at all. When it is questioned, the matter is often resolved by appealing to the facts that are brute in relation to the description: “how do you know that Anscombe bought potatoes?” “I saw the grocer taking them to her house”. It is only in necessarily exceptional cases that the normality of circumstances even enters the discussion — and, in fact, part of the point of distinguishing the levels of normality and of institutions is to limit the reach of counterexamples based on exceptional circumstances, showing that such special cases are not grounds for questioning the general validity of an institution. Thus, Anscombe's account isn't properly understood as a general theory of descriptions and their assembly from simpler elements, but rather an exploration of the complexity of normal procedure, and how much of it is, and must be, assumed.

References

Anscombe, G. E. M. (1981) [1962a] “Authority in Morals”, in: *Collected Philosophical Papers*, Vol. III: Ethics, Religion and Politics, Oxford: Basil Blackwell.

—. (2008)[1975a] “Faith”, in: *Faith in a Hard Ground*. Exeter: Imprint Academic, pp. 34-42.

—. (2000)[1957] *Intention*, Cambridge, MA: Harvard University Press.

—. (1981)[1958] “On Brute Facts”, in: *Collected Philosophical Papers*, Vol. III: Ethics, Religion and Politics, Oxford: Basil Blackwell, 22-5.

—. (1981)[1962b] “On Sensations of Position”, in: *Collected Philosophical Papers*, Vol. II: Metaphysics and the Philosophy of Mind, Oxford: Basil Blackwell, 71-4.

—. (1981)[1976] “The Question of Linguistic Idealism”, in: *Collected Philosophical Papers*, Vol. I: *From Parmenides to Wittgenstein*, Oxford: Basil Blackwell, 112-33.

—. (1981)[1979a] “Under a Description”, in: *Collected Philosophical Papers*, Vol. II: *Metaphysics and the Philosophy of Mind*, Oxford: Basil Blackwell, 208-19.

—. (1981)[1979b] “What is it to believe someone?”, in: Luke Gormally and Mary Geach (eds.) *Faith in a Hard Ground*, Exeter: Imprint Academic, 24-33.

—. (2015) “Grounds of Belief”, in: Luke Gormally and Mary Geach (eds.) *Logic, Truth and Meaning*, Exeter: Imprint Academic, 182-189.

Baker, G. P.; Hacker, P. M. S. (2009) *Wittgenstein: Rules, Grammar and Necessity. An Analytical Commentary of the Philosophical Investigations*, v. 2, 2nd edition. Oxford: Wiley-Blackwell.

Bloor, David (1997) *Wittgenstein, Rules and Institutions*, London: Routledge.

Dörge, Friedrich Christoph; Holweger, Matthias (2021) “The non-existence of institutional facts”, *Synthese* 199:4953-4974.

Greco, John (2021) *The Transmission of Knowledge*, Cambridge: Cambridge University Press.

Lackey, Jennifer (2008) *Learning from Words*, Oxford: Oxford University Press.

Searle, John (1995) *The Construction of Social Reality*, New York, NY: The Free Press.

—. (1965) “What Is a Speech Act”, in: Black, Max (ed.) *Philosophy in America*, Ithaca, NY: Cornell University Press.

Acknowledgement

This work is funded by national funds through FCT – Fundação para a Ciência e a Tecnologia under the scholarship reference UI/BD/154583/2023. Thanks to Nuno Venturinha and Samuel Laves for their comments.

The origin of language in Cavell, Mulhall, and Wittgenstein

So Kimoto (Kyoto, Japan)

Abstract

In this essay, I will explore Ludwig Wittgenstein's view on the origin of language, while engaging with two other thinkers, Stanley Cavell and Stephen Mulhall, who offer their own distinctive ideas on this subject. While their ideas are closely linked to Wittgenstein's, I believe his work allows for an alternative picture of the origin of language. First, drawing primarily from *A Pitch of Philosophy*, I will analyze Cavell's notion of a "voice that precedes language and gives itself language", arguing that this picture means taking the given language as one's own. Second, I examine Mulhall's reading in *Wittgenstein's Private Language*, showing that his picture of language acquisition presents the child as passively exposed to otherness while at the same time retaining subjectivity. Finally, I compare *The Blue and Brown Books* with *Philosophical Investigations* regarding Wittgenstein's evolving conception of teaching. I argue that his later work increasingly emphasizes the past-tense of teaching, leading to a picture of language's origin that fully foregrounds otherness.

In Wittgenstein's philosophy, the question of *the givenness of language* arises. Analyzing the notion of 'forms of life' (*Lebensformen*) in *Philosophical Investigations*, Juliet Floyd argues that this concept functions not as an "object of investigation, but rather a norm of elucidation" (Floyd 2020: 118). The 'forms of life' provide a structural foundation for language. However, Wittgenstein does not seek to "offer a description of the 'given'" or to assert "that *Lebensformen* should be taken to *be* this" (ibid.), but rather he aims to illuminate our relationship to the givenness.

What is my relationship to my language? In general, a *picture* of receiving a given varies depending on what is emphasized: one may claim to have *taken* the given powerfully and on one's responsibility; or to have *been compelled* to receive it passively; or to have accepted it as a *gift*. These different pictures of the given may correspond to different conceptions of its origin.

In this essay, I explore Wittgenstein's view on the origin of language, while engaging with two other thinkers, Stanley Cavell and Stephen Mulhall, who offer their own distinctive ideas on this subject. Although their ideas are closely linked to Wittgenstein's, I believe his work allows for an alternative picture of the origin of language. The essay is structured as follows: First, drawing primarily from *A Pitch of Philosophy*, I analyze Cavell's notion of a "voice that precedes language and gives itself language", arguing that this picture means taking the given language as one's own. Second, I examine

Mulhall's reading in *Wittgenstein's Private Language*, showing that his picture of language acquisition presents the child as passively exposed to otherness while at the same time retaining subjectivity. Finally, I compare *The Blue and Brown Books* with *Philosophical Investigations* regarding Wittgenstein's evolving conception of teaching. I argue that his later work increasingly emphasizes the past-tense nature of teaching, leading to a picture of language's origin that fully foregrounds otherness.

1. Cavell: Subjectivity of Responsible Agent

The discussion on the origin of language can be most clearly approached through Cavell's perspective, as he explicitly articulates his position. In *A Pitch of Philosophy*, he speaks of "a fantasy of a voice that precedes language, that as it were gives itself language" (Cavell 1994: 69). Although he uses the term "fantasy", he does not mean that this idea is illusory. Rather, Cavell often employs 'fantasy' in a positive sense – particularly to denote a formless yet foundational thought that, though lacking objective grounding, underpins his other ideas. In this context, 'fantasy' roughly corresponds to 'picture' discussed earlier in this essay. An alternative picture he presents, though does not endorse, is the fantasy of "coming late to language" and thus acquiring it "by stealing" (ibid.). In this case, he states:

[...] there always remains a problem whether language is mine, something that giving myself language should precisely settle. (Unless, among other possibilities, I can, with Augustine, thank God for giving it to me.) (ibid.)

Cavell identifies several pictures concerning the origin of language: 1) A voice (my voice?) gives itself language (Cavell's picture); 2) One steals language from others (e.g., elders), living with perpetual doubt about its ownership; 3) Language is a divine gift, received with gratitude. The third view aligns with Augustine's thought, which Cavell discusses in *Philosophical Passages*. There, Cavell quotes from *Confession*, where language acquisition is depicted as self-teaching, enabled by virtue of "the mind, O my God, which thou gavest me" (see Cavell 1995: 176). Here, the child learns language independently but does so through a faculty bestowed by God.

Famously, Wittgenstein opens *Investigations* with a quote from *Confession*, though from a different passage than cited above. Cavell highlights the key distinction: the "difference [...] lies in this one's appeal to God" (Cavell 1995:

177). According to Cavell, Wittgenstein deliberately chose not to quote the picture of language's origin that grants the presence of God. To Cavell, this choice indicates the nature of Wittgenstein's philosophical engagement.

I take Wittgenstein's quoting of his Augustine passage, after the death of God (hence after it became impossible to quote, to similar effect, Augustine's previous passage), to bring the questions it contains into the ensuing (modern) realm of silence, of sourcelessness. (ibid.)

In the age after "death of God", the problem about the legitimacy of inherited language can no longer be settled by appealing to an external source. Cavell sees Wittgenstein as recognizing this problem and addressing this issue. Consequently, from Cavell's perspective, the picture of a "voice giving itself language" continues the very investigations that Wittgenstein pursued.

Now, is it odd to conceive of something that precedes language and gives itself language? If that implies imagining a substantial entity existing before birth, it would be nonsensical. However, Cavell does not argue for this entity, nor that we invent language independently, without learning it from our parents. Interpreting Thoreau's *Walden*, he states that the term "writer's mother" or "mother tongue" implicitly contains "the recognition of endless indebtedness, which others imagine can be paid off" (Cavell 1994: 9). If Cavell does not deny the biological precedence of parents, or a temporal precedence of parental language, what does this fantasy of voice mean and seek to illuminate?

One way to investigate this question is through Cavell's notion of 'inheritance'. In *A Pitch of Philosophy*, subtitled "Autobiographical Exercises", he recounts his personal life history and intertwines this with the history of philosophy. As his telling unfolds, the fact that Cavell stands in a "split" between two philosophical traditions – the German-French and the English-American, exemplified by the miscommunication between Derrida and Austin – trying to voice it (Cavell 1992a: 233), but ultimately coming to be perceived as "an alternative voice" from both sides (Cavell 1994: 13), strikingly mirrors how Stanley is raised by two parents of different gifts – a musically talented mother and a father skilled in story-telling, who had however an incessant "quarrels", "mutual destruction of interest in the world" (ibid.: 21-22) – now seeking to "put together the Segals and the Goldsteins" (ibid.: 29-30), both sides of his

families, and yet how this attempt ends up being perceived as a “mysterious” presence by both parents (ibid.: 17). Writing an autobiography in both micro and macro levels, a history of his family and philosophy, thus reflects his engagement with the feeling that “philosophy is not *mine*” (Cavell 2002: xxxvii) and his attempt to inherit his ‘identity’ (cf. Cavell 1994a: 136). Through this path, the question of ‘inheritance’ finds its way into, and runs through, Cavell’s reflection on language acquisition:

This way of formulating the birth of the human is the product of two figures or scenes of inheritance that have recurred in my thoughts since my decision to incorporate autobiography into these lectures: the child in Augustine’s account of his acquiring of language, with which Wittgenstein opens his *Investigations*, and – I guess the most famous of our texts of inheritance – Jacob’s presentation of himself for Isaac’s blessing. (ibid.: 36)

Insofar as the inheritance (of traditions, of families) and finding its voice is the essential attempt in *A Pitch of Philosophy*, the inheritance and finding one’s voice in the origin of language are intrinsically connected. To inherit is to *take possession* of something *as one’s own*. If I inherit a house, it belongs to me and I assume authority over it. If I inherit a debt, I become responsible for its payment. Cavell’s picture underscores, I suggest, this sense of ownership: the language is *mine* and *I* am responsible for every word I speak.

This responsibility in language use aligns with Cavell’s reading of Wittgenstein’s later philosophy. In *The World Viewed*, for instance, Cavell writes:

Wittgenstein is known for his emphases upon the publicness of language. But his emphasis falls equally upon the absoluteness of my responsibility for the meaning I attach to my words. (Cavell 1979: 127).

In this section, I have examined Cavell’s picture of language’s origin. His fantasy of ‘a voice preceding language, and gives itself language’ reveals language as fundamentally *mine*, as demonstrated through his concept ‘inheritance’. The emphasis on *responsibility* is central to Cavell’s interpretation of Wittgenstein, and thus, this picture is seen as a continuation of Wittgenstein’s philosophical inquiries. In sum, Cavell’s picture highlights the *subjectivity of a responsible agent*.

In the next section, I will turn to Mulhall's interpretation of Wittgenstein, examining how his account foregrounds an *otherness* as deeply embedded within the language user, contesting with its subjectivity.

2. Mulhall: The tension between Subjectivity and Otherness

In *Wittgenstein's Private Language*, Mulhall gives thorough interpretations of *Investigations*, focusing exclusively on the sections from 243 to 315, in a manner largely aligned with Cavell's thinking. What distinguishes Mulhall's reading, I suggest, is his emphasis on the role of elders in a child's language acquisition. For instance, analyzing section 268, he states: "[T]o be in a position genuinely to give my word, and so to think of words as mine to give [...], I must first receive them, take them from others" (Mulhall 2007: 114).

His sense of receiving language from others also emerges in his discussion of section 244 of *Investigations*, where Wittgenstein describes how a child, crying in pain without articulated language, learns the word 'pain' from an elder (PI 2009: §244). Mulhall points out that "the key connection (between natural and verbal expressions of sensations) is not itself seen and acted upon by the person whose expressions these are, but rather by those around him" (Mulhall 2007: 29). He further notes that for a child to learn the word 'pain', the elders must *acknowledge* her behaviour as pain behaviour (cf. *ibid.*: 34). Language acquisition, then, is inherently dependent on others. Here, a subtle distinction emerges between Cavell and Mulhall regarding the concept of 'acknowledgement'. Cavell famously argues that philosophy's concern with the limit of knowledge of the world suggests our relation toward the world to be beyond knowledge, the relation of acknowledging it (e. g. Cavell 2002: 237; Cavell 1992b: 133). While Cavell emphasizes the subject's act of acknowledging others, Mulhall here focuses on *being acknowledged by* others.

Mulhall speaks of "idea of children acquiring language, understood as in the gift of one's elder" (Mulhall 2007: 30), and in this respect Mulhall's picture of the origins of language – namely language acquisition as a gift – differs from Cavell's. Yet Mulhall's picture does not present language acquisition purely as a result of otherness. His view is not the same as, for example, Foucault's picture of our identity that subjectivity is formed entirely through internalizing the social norm. In Mulhall's account, subjectivity neither consists wholly of otherness nor is extinguished in the process of entering into

a linguistic society. There is a conflict between the subjectivity and otherness, and this tension is reflected in his concept of ‘self-differentiation’.

For if it is the introjection of language and society that makes it possible for the child to avoid being lost in, and hence to, his experience, if without that internalization he would lack the capacity to articulate and hence acknowledge the state he is in, then it becomes tempting to say that the child’s achievement of self-awareness and selfhood here appears as a matter of achieving a certain kind of internal self-differentiation. In internalizing his elders’ gift of language, he internalizes (its and their) otherness: he gains the capacity to distance himself from himself [...]. (ibid.: 37)

Although Mulhall foregrounds the otherness of language, his account retains a claim to subjectivity. There remains still “he” who *performs* the act of internalization. The distance in ‘self-differentiation’, I suspect, lies between the self of others and the *self of the child*. This tension may explain why Mulhall describes language acquisition not only as a “gift” (ibid.), but also as “an intervention from the very social world” (ibid.: 34) – a term that seemingly contradicts the notion of a gift. This ambivalence, in which language is both “gift” and “intervention” from outside, characterizes Mulhall’s picture as a *tension between subjectivity and otherness*.

In the next and final section, I will explore another possible picture of language’s origin within Wittgenstein’s philosophy, comparing it with perspectives of Cavell and Mulhall while tracing the development of Wittgenstein’s thought from *The Blue and Brown Books* to *Investigations*.

3. Picture of Otherness in Wittgenstein

We have examined two philosophical pictures of language’s origin: one emphasizing subjectivity and responsibility, and another balancing subjectivity with otherness. I now turn to a third picture, drawn again from Wittgenstein, which foregrounds *otherness entirely*.

To motivate this exploration, I first express my reservations about the previous two pictures. Regarding Mulhall, I have already suggested that his account

presupposes the child already as, to some extent, a subject, one *who* internalizes their surroundings and upon *whom* society imposes itself. But should we consider a being before the ‘language-game’ as already a subject?

Regarding Cavell, I note two key differences between his view of teaching and Wittgenstein’s. According to an idea that Cavell develops under the name of *Emersonian Perfectionism*, each person represents their own voice while confronting others with different voices. This confrontation is conceived as a kind of teaching – I teach you with a different partiality; I learn from you who possess a different voice. Cavell clarifies this idea by contrasting Plato’s and Emerson’s conceptions of selfhood, describing the path from Republic to “democratic need for perfection” as “a path from the idea of there being one (call him Socrates) who represents for each of us the height of the journey, to the idea of each of us being representative for each of us” (Cavell 1990: 9). The first feature of teaching in Perfectionism is an *symmetrical structure*. Unlike Platonic pedagogy, there is no Socratic figure who is “intellectually authoritative” (ibid.: 6). While language acquisition and political education are distinct, Cavell treats Wittgenstein’s concern similarly, for instance, in his discussion of section 217, which he calls “scene of instruction”, (ibid.: 71). From Cavell’s perspective, the child is as authoritative as the teaching adult. Another key distinction is the temporal orientation of teaching. According to Perfectionism, teaching is directed toward a yet unattained future. Wittgenstein, by contrast, asks: How *did* we learn this word? This suggests that, for Wittgenstein, language acquisition belongs to the *past*.

The picture that I draw from Wittgenstein emphasizes the following aspects: Language acquisition is perceived as something that has occurred in the *past*; and it involves an *asymmetrical* teacher-learner relationship in which the child is passive and *not yet a subject*.

I take the first aspect – namely, the past-tense nature of language acquisition – to be particularly worth considering here. Can we recall acquiring language (by putting ourselves, as it were, in the position of Augustine’s child cited at the beginning of *Investigations*)? Can we talk about another person’s language acquisition in the present tense, specifying the moment when a child has *just* learned it? Although a discussion on this exact theme may not be found in *Investigations*, some passages can be read as reflecting Wittgenstein’s ideas on

this topic. In sections 156 to 178 in *Investigations*, Wittgenstein discusses the act of reading. His primary concern is whether we have a certain feeling when we are reading that distinguishes it from non-reading. However, in the first few sections (§§156-158), he also examines the case of a child learning to read.

A third person hears this pupil on such an occasion and says, “He is reading”. But the teacher says, “No, he isn’t reading; that was just an accident”. – But let’s suppose that this pupil continues to react correctly to further words that are put before him. After a while, the teacher says, “Now he can read!” – But what of that first word? Is the teacher to say, “I was wrong, he *did* read it after all” – or, “He only began really to read later on”? – When did he begin to read? Which was the first word that he *read*? This question makes no sense here. (PI 2009: §157)

“At what point has the child learned to read?” – this is the question that Wittgenstein rejects here. The same thing may perhaps be said about the language learning as a whole. After all, we can only say *retrospectively*, whether about others or ourselves, that one has learned language.

This interpretation aligns with the development of Wittgenstein’s thought from *The Brown and Blue Books* to *Philosophical Investigations*. First, this work abandons the linear progression from a primitive to complex language-games in the earlier work (cf. Pichler 2013: 74). Wittgenstein increasingly refrains from specifying the *moment* a child learns a language, treating rather the process as something that has *already happened* (of course, not exclusively – since *Investigations* still begins with a description of language acquisition). This shift coincides with a decline in his use of the term ‘training’ (*Abrichtung*). Second, Wittgenstein no longer refers to an elder’s act of pointing as an ostensive “definition” or “explanation”, as he did in *The Blue and Brown Books*. Instead, he speaks of “ostensive teaching of words” (PI 2009: §6). If pointing to something were an explanation, the child could react to it – asking a question, evaluating, or rejecting the definition. Hence, these textual developments suggest an alternative picture of language’s origin: language was *given* in the past, at a time when the child was not yet a linguistic subject.

This essay has examined Cavell's and Mulhall's views on language's origin and proposed an alternative. However, further investigation is needed to explore the full implications of the alternative. One unresolved issue is how givenness of language (or existence, as Mulhall suggests) can be understood as a gift. As noted in the second section in relation to Foucault, absolute givenness can also manifest as *oppression* – as something *imposed* upon the recipient, upon the subject. From this perspective, the private linguist in *Investigations* can be understood as regarding the givenness of social language not as a *gift*, but as a *curse* – not as a *freedom* to participate in the language-game, but as *confinement* within it. How, then, might one draw a picture in which the givenness of language – or my very existence – becomes a *blessing*?

References

Cavell, Stanley (1979) *The World Viewed: Reflections on the Ontology of Film*, enlarged edition. Harvard University Press.

Cavell, Stanley (1990) *Conditions Handsome and Unhandsome – The Constitution of Emersonian Perfectionism*, Chicago: The University of Chicago Press.

Cavell, Stanley (1992a) "In the Meantime", *The Yale Journal of Criticism* 5, no. 2, 229-237.

Cavell, Stanley (1992b) *The Senses of Walden*, expanded edition. Chicago: The University of Chicago Press.

Cavell, Stanley (1994) *A Pitch of Philosophy – Autobiographical Exercises*, Cambridge: Harvard University Press.

Cavell, Stanley (1995) *Philosophical Passages: Wittgenstein, Emerson, Austin, Derrida*, Cambridge: Basil Blackwell Ltd.

Cavell, Stanley (2002) *Must We Mean What We Say?: A book of essays*, 2th. updated edition. Cambridge: Cambridge University Press.

Floyd, Juliet (2020) 'Wittgenstein on Ethics Working Through Lebensformen', *Philosophy and Social Criticism* 46, no. 2, 115-130.

Mulhall, Stephen (2007) *Wittgensteins Private Language. Grammar, Nonsense, and Imagination in Philosophical Investigations §§243-315*, Oxford University Press.

Pichler, Alois (2013) "The Philosophical Investigations and Syncretistic Writing", in Nuno Venturinha (ed.) *The Textual Genesis of Wittgenstein's Philosophical Investigations*, New York: Routledge.

Wittgenstein, Ludwig (1969) *Preliminary Studies for the “Philosophical Investigations”*. Generally Known as *The Blue and Brown Books*, 2th. Oxford: Basil Blackwell.

Wittgenstein, Ludwig (2009) *Philosophical Investigations*. 4th, revised edition, P.M.S. Hacker and Joachim Schulte (eds.), G.E.M. Anscombe, P.M.S. Hacker and Joachim Schulte (trans.), Oxford: Wiley-Blackwell.

Acknowledgement

This work was supported by JSPS KAKENHI Grant Number JP24KJ1356.

"Doing things with names" Recognising and Resisting the Harms of Deadnaming

Anna Klieber (Cardiff, United Kingdom)

Abstract

The phenomenon of deadnaming – calling a trans person by a name they have discarded – is a common issue for trans people, and generally considered to be very harmful. Given the pervasiveness and harmfulness of deadnaming, it is tempting for the feminist philosopher of language to focus primarily on the (linguistic) harm it gives rise to. This essay will take this harm seriously, but not limit its analysis to it: On the one hand, I show that deadnaming harms people by denying their first-person authority over their names, and that, as a practice, it is tied to certain ideological commitments towards sex and gender. On the other hand, however, I focus on trans linguistic practices that allow to reclaim this authority. More specifically, I highlight how trans linguistic practices surrounding naming are importantly connected to trans joy and playfulness. I have in mind practices such as "trying out a new name", playing with the "fittingness" of a name, frequent name changes, or the use of multiple names. In playing with (gendered) names in this way, trans people establish a new idea of what it means to have authority over them, resisting the transphobic ideology underpinning practices of deadnaming.

Feminist philosophers of language have shown us time and time again that speaking does not take place in a vacuum, but is crucially situated in a political context, saturated by power structures and imbalances. Gender is an important component of these (linguistic) power struggles, and feminist philosophers have provided us with a plethora of resources for analysing sexist, misogynistic, but also *transphobic*, speech: We see discussions on pronouns, misgendering, and slurs (Kapusta 2016; Dembroff and Wodak 2018; McNamara 2021), all pointing out how our social and feminist analyses of communicative enterprises ought to take seriously the harm done to trans individuals through transphobic linguistic practices. What has received less attention (with the exception of Koles 2024; Klieber and Bolton MS) is the phenomenon of *deadnaming* – calling a trans person by a name they have discarded. Deadnaming is a common issue, and generally considered to be very harmful.

Given the pervasiveness and harmfulness of deadnaming, it is tempting for the feminist philosopher of language to focus primarily on the (linguistic) harm it gives rise to. Klieber and Bolton (MS) in particular, argue that deadnaming is a form of subordinating speech because it deprives trans individuals of first-person authority (following Bettcher, 2009).

In this essay, beyond just describing how the harm done *in* deadnaming is a denial of first-person authority, I wish to focus on trans linguistic practices that allow us to *reclaim* this authority. My two core questions are the following: Firstly, in what ways does deadnaming constitute a denial of first-person authority? I will follow Klieber and Bolton's account, but further add it by highlighting that the practice of deadnaming is intimately tied to a certain kind of (transphobic) ideology. My second question is about practices that trans people have developed to *reclaim* this denied authority. I show that, beyond directly countering such harmful speech, there are linguistic practices within the trans community surrounding names that ultimately go towards challenging the ideology underlying deadnaming practices more broadly.

I will proceed by first introducing what deadnaming is, and outline in a second step why it constitutes a deep harm and violation. Third, I highlight to how resistant linguistic practices surrounding naming are importantly connected to trans joy and playfulness, drawing on Hernandez and Crowley (2024). I have in mind practices such as "trying out a new name", playing with the "fittingness" of a name, frequent name changes, or the use of multiple names. In playing with (gendered) names in this way, trans people establish a new idea of what it means to have authority over them, resisting the transphobic ideology underpinning practices of deadnaming.

1. What is deadnaming?

The term *deadname* first emerged in online trans spaces to refer to a name that has been discarded by trans people for reasons related (but not limited) to gender; commonly the name given at birth. The reasons for name changes are sometimes less straightforward than popular discussions would suggest: By no means every trans person rejects their given name, and not all trans people aim to take it out of circulation completely (though this is a goal for many). The notion "deadname" can also apply to other kinds of name changes – e.g., Koles (2024: 716) discusses Muhammad Ali's rejection of what he called his (former) "slave name". To simplify things, this essay will focus on cases that surround gender specifically. I will understand the practice of *deadnaming* (for trans people) as one that reinforces, assigns, or refers to a trans person using a name they have discarded as part of their gender expression and identity.

Importantly, I consider utterances as deadnaming *even if* they occur in absence of the person being deadnamed. For instance, while the direct case may be one of a parent continuing (despite being corrected) to call their child their deadname, in the indirect case, the parent may use it without said child being present. Both are cases of deadnaming, and both, as I will outline next, constitute harmful speech.

2. How does deadnaming harm?

Ishani Maitra (2012) puts forward a pervasive account of how "ordinary" cases of racist hate-speech – where the speaker doesn't hold an exalted position of authority (as a lawmaker or judge) – can subordinate. Her famous target example is a white man shouting racist abuse at an Arab woman in a crowded subway car (100, 115).

Maitra explains hate speech of this kind as subordinating by utilising Austin's (1962) conception of speech acts, understanding the *illocutionary* as the dimension of our speech acts that does something *in the very act* of saying it, and the perlocutionary as the (*causal*) *effects*. Accordingly, hate speech has *perlocutionary dimensions* (such as *causing* distress, or legitimising such speech), but also has *illocutionary dimensions*: In the case she discusses, the racist hate speaker succeeds in subordinating the woman *in speaking* – his utterances are *subordinating illocutions* that rank her as inferior (Maitra 2012, 98).

According to Maitra, authority still plays a role here. Rather than already holding "official" authority, a speaker may *derive* their power from a formal authority (e.g. a child being instructed by a teacher to distribute tasks among their classmates), or they may gain it through *context*. In the latter case, someone is *licensed* when nobody resists someone's *assumption of authority*. Both of these means allow for the possibility of ordinary speakers subordinating others. As in Maitra's target example (105, 115), the act of doing this can be subordinating *in itself* – in this case by being *licensed* through individuals in the immediate context.

Now, what does this have to do with deadnames? I here pick up the already mentioned account by Klieber and Bolton (MS). They argue that deadnaming can constitute harm in ways parallel to what Maitra describes. In particular, they argue that *the illocution of deadnaming* denies a trans person's authority

over their names (and gender), drawing on Bettcher's (2009) account of denials of first-person authority (FPA).

According to Bettcher, such "Basic Denials of Authenticity" are hostile to people's self-identification (e.g. if a trans man is told that he is "really" just a woman dressing up). As will be important again shortly, this denial is embedded in broader societal discourses about appearance, "biological" reality, discovery of trans-status, and deception (99). To deny a person's avowal of gender identity, then, is to deny what Bettcher calls their *existential* self-identity – a conception of self-identity that concerns all the things that make us "*who we really are*": the beliefs we hold, our attitudes and commitments (110). A denial of such identity – and our first-person authority over it – reinforces a certain transphobic reality, contesting not only people's gender, but also *who they are*. FPA is, according to Bettcher, an inherently *ethical* conception, making its denial an inherently ethical issue.

Klieber and Bolton argue that in cases of deadnaming something equivalent happens: If a person insists that your name is not really your name, that *they* get to decide *who you are*, they make a judgement about your *existential identity*. If announcing one's new name is a public avowal – an illocution that *authorises* how we see ourselves – a denial of this act is a denial of our authority over our own names, an attempt to overrule our avowal (see Klieber and Bolton MS). This is more than just an effect of deadnaming: the denial of FPA manifests in the act of deadnaming itself, constituting subordination.

Bettcher's notion of FPA is crucial going forward. However, I wish to highlight a harm that goes beyond Klieber and Bolton's original analysis. In particular, I want to pinpoint the *broader harmful ideology* underlying deadnaming as a (linguistic) practice, which is intimately tied to (transphobic) ideas of sex and gender. Bettcher (2007) is helpful in illuminating this: In this text, she highlights two common transphobic stereotypes – the trans person as a *pretender* or *deceiver*. This idea, e.g. that a trans man only "pretends", or that he is deceiving others about his "true sex", reveals a deep ideological commitment: the "sexed body constitutes the hidden, sexual reality" (48), and gender presentation is mere "appearance", something trans people frequently deceive others about. On this view, genitalia are taken to be "the essential determinants of sex" (48). Much more can be said about Bettcher's critique of

this, but for now just two points are important: First, transphobic reality reinforcement is underpinned by a strong *essentialism*, including that sex assigned at birth is "fixed", "determined" and "unchangeable". Second, this transphobic reality gets enforced by (a) labelling trans people as deceivers and pretenders, and (b), ultimately emphasising *that it is not they who have first-person authority over their genders*.

So, how does this connect to names? Of course, names are different from gender as a general social category. Nonetheless, how deadnaming is practiced in our society resembles the just-described essentialism, deception/pretending discourse, and ultimately *structural* denial of FPA.

For one, first names are often treated as "fixed", mirroring what we heard above. Beyond the difficulties of getting people to actually adopt a name change, many trans people will be familiar with social media comments demanding the disclosure of someone's "real" or "biological (!) name" (e.g., see Reed (2024) on a recent court case in Ohio surrounding a teachers refusal to use pupils' correct names, claiming that "she should be allowed to use a student's 'biological' name"). Of course, no names are "biological", in that sense. All first names are chosen, tied to cultural and family conventions, and frequently simply down to parental taste. The transphobic speaker urging a trans person to identify themselves by their "biological name" is not only making a category mistake; they also make a statement about how their attitude to names is tied up with their attitude towards *gender itself*. They may as well be asking, "What are you, *really*?"

Likewise, the disclosure of deadnames is linked to the ideological assumptions surrounding the "pretenders and deceivers"-discourse. Deadnames are often used in contexts that are aimed at "revealing" or emphasising that someone is trans. Klieber and Bolton (MS) discuss the disturbing cases of the murder of 16-year old Brianna Ghey, where numerous newspapers disclosed her deadname without any relevant reason (e.g. Tozer (2023) in the *Daily Mail* and Hurst (2023) in *The Independent*), and that of 21-year old Brandon Teena, who was murdered after being "discovered" as trans following a police report deadnaming him (Minkowitz 2018). Of course, deadname-use doesn't always

happen with the *intention* to demean. But even "accidental" deadnaming can have the effect of outing someone as trans, and disclosing their old name to others – marking the person's status *as trans*.

My point here is ultimately a, hopefully, straightforward one: Drawing on Bettcher (2007, 2009) and Klieber and Bolton (MS), we can see that FPA over names is not only denied in direct instances of deadnaming, but enshrined in the practice more broadly.

I now want to ask: If deadnaming denies trans people's FPA over their names, what practices exist to reclaim this authority? What I aim to highlight here is not a comprehensive list, but an outline of an overall linguistic practice: even if trans people are disenfranchised on a societal, broader level, the community is still able to use gendered words – and names – in ways that allow them to regain power, while *also* questioning the deeper ideology standing behind deadnaming more broadly.

3. Doing things with names

There are some obvious ways trans people and their allies may resist instances of deadnaming. One would be direct *calling-out*. How this plays out will differ (e.g. on whether it is you, or another person, that is being deadnamed, and whether it is outright malicious or down to a lack of knowledge and understanding of the issue). We may class this under the broad umbrella of counterspeech: Communication that has the goal of counteracting the potential harm brought about by other speech. For instance, we might do what Langton (2018) calls *blocking* – responding to an instance of deadnaming by asserting the correct name, refusing to let the deadname enter the common ground as an acceptable name to use. Another strategy may be the production of educational materials: This may take the form of spreading awareness via social media and advocating for an improvement of institutional name-change procedures.

While much more can be said about this, I want to consider, in the remainder of this space, in the spirit Hernandez and Crowley (2024: 294), *how trans people use language for liberation*. They highlight a number of different ways in which gendered language is used in liberatory and resistant ways, including their "playful" and "joyous use". I want to make the case that naming practices in the trans community fall into these categories as well, and thereby

structurally reclaim the authority denied through restrictive assumptions underlying practices of deadnaming more broadly.

How might trans people use names *playfully*? Hernandez and Crowley (2024) borrow the notion of *playfulness* from María Lugones (2003). She sees playfulness as an *attitude* carrying through an activity, which turns that very activity into play. Lugones emphasizes the importance of this for marginalized people to explore a world where it is others who hold power, without "absorbing" or "being taken in" by the perspective of the powerful or oppressor. According to Hernandez and Crowley (2024: 303–307), playful language can involve (a) taking up a playful attitude when defining and creating terminology, or (b) using creative terminology to cultivate a playful attitude. While their examples focus on terminologies developed and used about trans bodies in the trans community, reappropriating language (e.g. "transsexual") and playing with medicalised terminology, an equivalent attitude can be seen in certain naming practices within the trans community.

One crucial example is the play with the "fittingness" of a name. Many trans people simply think about what they want their name to be, and discard the old one – but equally many discard the old one, and *play* with a number of names, such as using multiple names at the same time, to evaluate which one may work best for them. This playful attitude can also include new name creation. Trans people frequently quite literally "make a name for themselves" (e.g., a thread on the subreddit r/NonBinary (2024) contains in-depth elaborations how people quite literally created their own name) – finding which names "feel good". While this can, and usually does, involve a clear rejection of deadnames, this rejection is simultaneously treated as opening up new possibilities: Finding a name that fits, using different names in different contexts – and, more broadly, finding ways to asserting and avowing that a particular name is *theirs*.

This brings us to the next usage Hernandez and Crowley (2024: 307–311) outline – joyous use. Joyous use of gendered language can, ultimately, help trans people fight marginalization and build resilience. As seen in the example above, building and trying out a new name is crucially about finding the language that "accurately describes how a trans person experiences gender" – again, it's about finding what language "*feels good*" (308). Playfulness with

names, I would say, is directly connected to this joyous use: Using your new name, and getting others to use it, can be part of the *playful attitude* that brings about *trans joy* and allows trans individuals to discover *what name works*.

I already mentioned above how both of these elements help build resilience and create space in a world that is often hostile to trans people. Deadnaming is part of this hostility, and so is the broader ideology underlying it. So, let's make it more explicit how this playful and joyous approach counters this ideology: One crucial aspect here is that, again to echo Hernandez and Crowley (2024: 309), trans language usages create *definitional power*. It is about creating social meaning within communities and for oneself. This makes linguistic self-determination an important underlying feature of the joyous (and playful) use of trans language.

More specifically, the playful approach (e.g. of trying out names, and building new names) directly counters the idea that names are fixed, unchangeable, and non-fluctuating. The joyous use of these names (enthusiastically using various names all at once or changing your name frequently) allows for the rejection of some names (e.g. a deadname, or names with a particular gendered connotation in general) while allowing for the joyous affect following from affirmative naming.

As a practice, this gives people *power*, *control*, and explicit *authority* over their names. To put it in the words of non-binary writer Erin Paterson (2021): "Renaming myself was a gift to myself; to claim that agency over my life was incredibly empowering, but it was also political". The ideological move underpinning deadnaming as subordinating speech is to deny that trans people have this authority. What the playful and joyous use can do, on the other hand, is – at least within one's (trans) community, but ideally more broadly than that – to (a) make room to reassert this authority when finding one's new name, and (b) give one the sense of your name as specifically *yours*.

4. Conclusion

This discussion is not exhaustive. Much more can and needs to be said about gendered naming practices in our society as well as in specific sub-cultures. My point was a general one: To highlight how deadnaming, as a specific linguistic practice, is underpinned by harmful ideology that goes beyond the issues seen in individual instances of deadnaming, and how, on the other

hand, there are linguistic practices in the trans community that go towards challenging that very ideology. Whether or not this latter practice – right now – has the power to change the more structural issue underlying deadnaming is a much bigger and broader question. While, at the time of writing, it has not yet undermined the harmful deadnaming practice, what it can do, at least, is create a *space* to regain and reassert power and self-definition. This will not render deadnaming *harmless*, just like that. My point is rather that certain *naming practices in the trans community* succeed in giving people a renewed sense of authority over their names, one that is all too often denied elsewhere.

References

- Austin, John Langshaw (1962) *How do to Things with Words*. Oxford/New York: Oxford University Press/Clarendon Press. doi: 10.1093/acprof:oso/9780198245537.001.0001.
- Bettcher, Talia Mae (2007) "Evil Deceivers and Make-Believers: On Transphobic Violence and the Politics of Illusion", *Hypatia: A Journal of Feminist Philosophy*, 22(3), 43–65.
- Bettcher, Talia Mae (2009) "Trans Identities and First-Person Authority", in Laurie Shrage (ed.) *You've Changed! Sex Reassignment and Personal Identity*. Oxford University Press, 98–120.
- Dembroff, Robin; Wodak, Daniel (2018) 'He/She/They/Ze', *Ergo, an Open Access Journal of Philosophy*, 5(20190926), 371–406.
- Hernandez, E. M.; Crowley, Archie (2024) "How to do things with Gendered Words", in Luvell Anderson and Ernie Lepore (eds.) *The Oxford Handbook of Applied Philosophy of Language*. Oxford/New York: Oxford University Press, 293–320.
- Hurst, Pat (2023) "Potential jurors given questionnaire ahead of Brianna Ghey murder trial", *The Independent*, 27 November. Available at: <https://www.independent.co.uk/news/uk/crime/warrington-transgender-manchester-crown-court-cheshire-justice-b2454180.html>.
- Kapusta, S. J. (2016) "Misgendering and Its Moral Contestability", *Hypatia*, 31(3), 502–519.
- Klieber, Anna and Bolton, Emma (MS) "'Call me by my name". Investigating the Harms of Deadnaming". *Under Review*.
- Koles, Taylor (2024) "The Semantics of Deadnames The Phenomenon of Deadnaming", *Philosophical Studies*, 181, 715–739.
- Langton, Rae (2018) "Blocking as Counter-Speech", in Daniel Fogal, Daniel W. Harris, and Matt Moss (eds.) *New Work on Speech Acts*. Oxford/New York: Oxford University Press.
- Lugones, Maria (2003) *Pilgrimages/Peregrinajes: Theorizing Coalition Against Multiple Oppressions*. London/New York: Rowman & Littlefield Publishers.

Maitra, Ishani (2012) 'Subordinating Speech I', in Ishani Maitra and Mary Kate McGowan (eds) *Speech and Harm: Controversies Over Free Speech*. Oxford University Press, 95–121.

McNamara, Chan Tov (2021) "Misgendering", *California Law Review*, 40(2020), 2227–2322.

Minkowitz, Donna (2018) "How I Broke, and Botched, the Brandon Teena Story", *The Village Voice*. Available at: <https://www.villagevoice.com/how-i-broke-and-botched-the-brandon-teen-story/>.

n.A. (2024) *How did you guys come up with a new name?*, r/NonBinary Reddit. Available at: https://www.reddit.com/r/NonBinary/comments/189eqxw/how_did_you_guys_come_up_with_a_new_name/ (Accessed: 28 March 2025).

Paterson, Erin (2021) "It Was A Line In The Sand For Everyone, Including Myself": One Vogue Writer Reflects On Changing Their Name After Identifying As Trans", *British Vogue*. Available at: <https://www.vogue.co.uk/arts-and-lifestyle/article/choosing-name-trans>.

Reed, Erin (2024) *Trump Judge Rules Deadnaming Trans Students Teaches 'Tolerance' In Bizarre Opinion*, *Erin in the Morning*. Available at: <https://www.erininthemorning.com/p/judge-rules-teacher-deadnaming-trans> (Accessed: 28 March 2025).

Tozer, James (2023) "I want to know if Brianna screams like a man or a girl": Teens accused of murdering transgender schoolgirl Brianna Ghey swapped messages as they plotted to lure her to a park where one of them stabbed her 28 times, murder trial hears" *The Daily Mail*, 27 November. Available at: <https://www.dailymail.co.uk/news/article-12796553/Teens-accused-murdering-Brianna-Ghey-talked-watching-torture-vids-buying-hunting-knife-weeks-transgender-schoolgirl-stabbed-28-times-murder-trial-hears.html>.

L. Susan Stebbing: The Art of Thinking Clearly

Silke Körber (Berlin, Germany)

Abstract

In the first half of the 20th century, debates about the status and transfer of scientific knowledge as well as the role of specific scientific disciplines played a crucial role in academia, including for the British analytical philosopher L. Susan Stebbing. In times of crisis and war she insisted that logical-analytical thinking could be useful to support modern societies, individual freedom and the development of liberal democracies and in contributing to collective well-being.

Using sources from Stebbing's collaboration and correspondence with the Vienna Circle member Otto Neurath from the 1930s onwards, as well as her own publications, this paper attempts to illuminate Stebbing's understanding of a general 'scientific attitude' and the role of modern science (and the scientist). This is particularly evident in the ideal of clear thinking and acting as purposive or directed to an end – considering the choice of suitable means for successful implementation in terms of empirical basis, causality and the complex real-life situations that influence the overall usefulness to society. The relevance of making these processes transparent lies in the examination of the possibilities and limits of scientific explanations in times of crisis and also in further nuancing the debate on what ideals and values can or should play a role in order to support emancipatory democratic practice – also in the context of more recent discussions about the purpose and value of scientific expertise.

1. The 'Scientification' of Philosophy

At the beginning of the 20th century, leftist and liberal intellectuals, philosophers, and scientists were engaged in the comparative study of appropriate social systems, ideologies and the question of how social collectives could be organised more efficiently and justly, not least in the face of political and economic crises, the First World War and emerging totalitarian regimes. It seemed essential to develop methods for rational planning and a solid scientific foundation without a rigid, authoritarian character or system and also protect human rights and the freedom of individuals. In other words, to achieve 'scientification' without succumbing to the temptation of assuming a simple linearity in the development of science, to believe in never-ending progress and to join the chorus of those who have labelled science a 'new religion'. The aim was to promote a new and better 'scientific attitude' in dealing with current problems in order to support the development and possibilities for action of citizens as social beings.

Central themes of modernity emerged in these discourses and led to programmatic consequences: from the influences of the new mass media and

processes of individuation vs. collectivisation to hopes and doubts about the latest achievements of science and technology. The latter had permanently changed Europe in the 19th century, the ‘scientific age’, as Werner von Siemens had called it. In this context, the ongoing process of professionalisation, standardisation and specialisation in the sciences and humanities was significant and at the time promoted both the crisis and the reorganisation of the academic system itself. The enormous technical progress and gain in knowledge in the ‘positivist’ sciences, from the natural to the technical sciences, especially physics with quantum mechanics and Einstein's theory of relativity, and applied mathematics, was not matched by comparable achievements in the humanities. This applied in particular to philosophy, which Kant and Wilhelm von Humboldt saw as the leading discipline or foundation of the sciences, and the associated neo-humanist ideal of education. It was not just a generational shift between senior scholars, who were brought up in a more humanist tradition and understanding of philosophy, and their younger peers, but a fundamental change and revision of the role of philosophy. Nevertheless, the accessibility and usability of scientific knowledge in the technical and natural sciences (as in the humanities) was still limited for certain social groups due to very different levels of education, the availability of information and its lack of transparency.

These developments formed the background for a ‘scientification’ of philosophy, a revised philosophical methodology represented by Ludwig Wittgenstein and the members of the Vienna Circle in the Manifesto of the *Wissenschaftliche Weltauffassung* and associated protagonists such as from the Lviv-Warsaw School (LWS): from rationalism to the postulate of linguistic clarity and exactness, to the strong influence of symbolic logic, to metatheoretical research, i.e. questions about epistemology that were primarily seen as clarifiable through the medium of language.

2. L. Susan Stebbing – Philosopher and Public Intellectual

As a growing number of works and recent studies show, L. S. Stebbing was closely associated with these movements, especially with the members of the Vienna Circle, even before Alfred J. Ayer introduced Logical Empiricism to British philosophy. She had corresponded and met with Moritz Schlick since 1930, invited Carnap to give a lecture in London in 1934, then also others, and became well acquainted with the various positions of the circle and the

analytic method. Thus, when Neurath wrote in a letter to Stebbing in 1938 “[...] you are flesh of the flesh of Logical Empiricism” it was by no means out of the blue. Moreover, according to Sander Verhaegh's research on 20th century philosophers, Stebbing was one of the most frequently cited authors in British philosophy journals (e.g. *Philosophy*, *Analysis*, *Mind*). She also influenced students like Max Black or Margaret MacDonald, and according to the Polish logician Maria Ossowska (LWS), who lived in London during the 1930s, Stebbing assembled an illustrious group of philosophers for regular meetings: “I attend private meetings of Prof. Stebbing, around whom the most interesting and most reasonable people gather. This group publishes the journal ‘Analysis’ [...]”.

Stebbing's position is all the more significant as she was the first woman to hold a professorship in philosophy in Great Britain at a time when the admission of women to university (also against the background of different curricula in school education for boys and girls) and the possibility of doctorates, habilitations and academic careers was still an exception in European countries: She had studied History at Girton College in Cambridge and went on to study Philosophy, graduating with a ‘Master of Moral Science’ from Kings College in 1912. In 1933, she became the first woman to be appointed Professor of Philosophy at Bedford College. She was chairwoman of the Aristotelian Society (1933/34), president of the Mind Association (1931/32), a honorary associate of the Rationalist Press Association, and a supporter of the League of Nations Union. During her studies, she was interested in G. E. Moore's analytical and common-sense philosophy, studied modern logic, philosophy of science and philosophy of language as represented by Bertrand Russell and Alfred N. Whitehead. According to Beaney and Chapman, Stebbing can be regarded as a proponent of what later became known as the Cambridge School within Analytic Philosophy, which was committed to logical analysis and interpretation of the ‘common language’ for solving scientific problems even before the rise of ‘ordinary language philosophy’ in Oxford after the end of the Second World War.

Nevertheless, and this made her stand out in the British academic environment of the time, Stebbing endeavoured to apply her philosophical work and humanist principles not only in theory, by examining the assumptions and principles underlying science, morality or society, but also in

practice, by enquiring into the specific purpose of clear thinking. On a personal level, this meant that she supported refugees at her own expense, helping scientists (including Otto Neurath, who fled to England during the war) and taking part in anti-war activism and on behalf of the League of Nations Union. In 1941 she became the first female president of the Union of Ethical Societies, the forerunner of Humanists UK, which campaigned for secularisation, human rights, democracy, and equality. Moreover, she intended to transfer the 'scientification' of philosophy and the sciences into the public sphere and support citizens to think more clearly and lead a better life with and through science. In her opinion, not religion or ideologies, but philosophy and the sciences should provide orientation and help to answer the question of a 'worthwhile life', as she expressed 1939 in *Ethics and Materialism*. This dual focus on providing standards for logically analysing, interpreting and clarifying statements to solve scientific problems, eliminating difficulties by showing how to formulate the language of science as she put it in *Language and Misleading Questions* in 1938 – and on the practice of logical-critical thinking played a central role in her exchange with Otto Neurath.

3. Stebbing and Neurath: Thinking and Acting for the People

It is noticeable that both Stebbing and Neurath were equally interested in science and its role and usefulness for general education. Both criticised scientists as well as representatives from politics and the media, who were dedicated to providing unobjective and emotionalised information for the common reader. Otto Neurath had already established relationships with representatives of a 'utopian internationalism' in Vienna and tried to develop ideas for better communication of scientific knowledge in a systematic, rational methodology and according to circumstances and contexts. And although all members of the Vienna Circle recognised the epistemological change, they developed very different approaches to dealing with the resulting challenges – as we understand today. In Otto Neurath's case, these were more radical scientific and socio-political ideas that combined the philosophical agenda of 'Red Vienna' and the Scientific World Conception with his understanding of education as a prerequisite for modern democratic societies. This led, for example, to his involvement in the Munich Soviet Republic (Münchener Räterepublik) 1919 or the development of the Method of Pictorial Statistics at the Museum for Social and Economic Affairs (Gesellschafts- und

Wirtschaftsmuseum) in Vienna in the 1920s: Together with his team, he developed rules for the systematic transformation of data, primarily for the German speaking market.

In 1934, Neurath, his future wife Marie Reidemeister and others were forced to emigrate after the rise of National Socialism, initially to The Hague. Stebbing later helped Neurath and Reidemeister to escape the Nazi occupation in the Netherlands and settle in Oxford. Neurath and Stebbing corresponded and worked together for over a decade from the 1930s: She became influential as a member of the Unity of Science movement led by Otto Neurath and others, and as an organiser of one of its conferences in Cambridge in 1938. In addition, she was a member of the board of directors of the International Foundation for Visual Education, which Neurath and his team founded in The Hague, and the first chairwoman of the Isotype Institute in the United Kingdom.

Neurath's interpretation of the 'visual auxiliary language' Isotype (International System of Typographic Picture Education), developed to provide education and information for an international mass audience, a specific 'picture-text style' derived from it, and the scientific publishing project *Encyclopedia of Unified Science* can be understood in the context of his preoccupation with modern symbolic languages and the utopia of a unified communication that transcends social and geographical boundaries for the purpose of disseminating knowledge for modern societies. In a letter to Stebbing he explained his approach as the "Unification by Logicalisation" and "Unification by Visualisation" in 1939. For Neurath, these were two sides of the same coin and they characterized the scientific agenda of the Unity of science movement and the *Encyclopaedia* as well as his educational efforts, including exhibitions, books, film etc.

For both, Stebbing and Neurath, changes in the scientific and philosophical world view, just like linguistic changes, should have a direct impact on people's everyday lives. The high standards of adequate formalisation of scientific knowledge should also be introduced into everyday language. And one way to achieve this was to develop a suitable analytical set of rules and a more precise linguistic expression.

Despite some programmatic differences regarding the method of scientific analysis (in contrast to the logical analysis of the Vienna Circle, Stebbing opted for a 'directional analysis', a gradual form of clarification of the structure of statements) they agreed that the role of philosophy, science and expertise in general needed to be reconsidered and that the formation of an egalitarian global society had to be actively promoted. This was understood to be part of the educational duties and upbringing in democratic societies, as Neurath claimed: "The transfer of looking at more than one possibility, to be prepared to alter statements, is the principle of the scientific attitude. The social pattern, which permits more than one opinion etc. is the 'democratic pattern'." The focus was on demonstrating scientific virtues and values in dealing with knowledge and showing what good scientific practice should look like, by practical example. He argued against a 'popularisation' of content, particularly in the service of entertainment and without using a well thought-out methodology. Instead, it was important to train critical thinking and the visually guided, stepwise comparative assessment of facts and data with a didactic structure from the simple to the complex – as an independent, questioning attitude towards institutions, politics, science or the media. The requirements on the knowledge production side had to be reflected upon and revised accordingly. And in this respect, the intentions and approaches of Neurath and Stebbing seem to have complemented and enriched each other.

4. Envisioning a Better Life

In 1943 before her death from cancer, Stebbing gave the 13th L. T. Hobhouse Memorial Trust lecture *Man and Morals Principles*, in which she contrasted moral or ethical statements with scientific statements and examined them as 'isolates', such as those used in economics or as 'abstractions', as used in the sciences. For her, the comparison of moral statements with scientific statements makes clear that they follow different rules and belong to different areas of human interaction: moral judgements, which are concerned with 'ethical situations', with 'good' and 'evil', are always individual and circumstance- or context-dependent, scientific judgements about the physical world can be objectified and abstracted. Nevertheless, the distinction between the sciences should be made according to the purpose and situation that the respective experts are dealing with – and according to the internal complexity of a subject or field. Stebbing's understanding of science is – in agreement with

Neurath – that of a collective endeavour, not to be confused with the intentions of an individual scientist who is not free from interests, values or the circumstances of empirical experience, as she explained in *Philosophy and the Physicists*, 1937:

“Science is not a goddess or a woman. We cannot ask science, but only scientists. Moreover, we must ask our questions of the scientist at a moment when he is in a scientific temper, capable of giving us ‘the ascertained facts and provisional hypotheses’ without any admixture of the emotional significance which he reads into these facts in his scientific moods.”

In her writings, Stebbing distinguishes between the ‘ethical neutrality’ of science and that of scientists. In her opinion, scientists play an important role within society as providers of information on the basis of which decisions can be made for society and in the political sphere, but as individuals they are bound to socio-political, cultural or ethical contexts and convictions. In this sense, the individual scientist should not be confused with science itself: the pursuit of scientific goals is in itself something good and worthy of recognition. But the values involved are those of the individual actors, not those of science itself. This seems to be a rather modern understanding of science and the role of the scientist. Even if it is not really clear or likely whether she would ultimately integrate these ‘value-laden’ decisions by scientists – to use a modern term – into an overall more pluralistic concept of objectivity, as has been repeatedly discussed within the discipline since then. Especially as it was important at the time to defend the idea of science to support a free, equal democratic society against its enemies and not to weaken it by relativizing, but instead to help improve the lives of as many people as possible and make them less susceptible to propaganda and supposedly simple solutions.

For Stebbing, rational, purposeful thinking can or should lead to rational action on an individual and collective level. In her view, the principles of logical analysis are best associated with thinking ‘towards a purpose’, as ‘directed’ thinking aimed at answering a question or solving a problem. She explained what this meant in practice in her bestseller, which was published by Penguin Books in 1939. This text was based on a proposal for some talks Stebbing wanted to give on the BBC radio programme. *Thinking to Some*

Purpose is a philosophical text directed at the general public. Its objective was to educate people about how the philosophy could benefit society and at the same time train readers to practise it, as described on the cover: “A manual of first-aid to clear thinking, showing how to detect illogicalities in other people’s mental processes and how to avoid them in our own.”

She illustrated her arguments with a series of analyses of contemporary texts from newspapers, political speeches and advertising. In the process, she analysed how language was used according to the situation and in this type of interaction. As a result, she called for a more conscious and clearer use of language on the part of authorities and experts, as this was often worthy of criticism, confusing and needed to be scrutinised more closely. On the part of citizens, she called for alertness and attentiveness in order to fulfil civic duties - clarity and clarification are therefore prerequisites on both sides for better communication and understanding, especially of complex contexts. Therefore, the objective would be to provide clear, non-manipulative, non-emotional and non-simplistic information that enables people to form their own judgements and act rationally and practically. If citizens had a clear and coherent picture of the world with and through science, this would ultimately lead to increased empowerment and participation. As Otto Neurath put it in a nutshell, the common man should have “enough information to keep the power of the expert in check“.

It seems that the scientification of public life embracing the democratic idea of a happier and better future was a vision that Stebbing and Neurath shared in times of war and crisis, trying to develop a rational vision for a better future beyond religion or totalitarian ideologies. This also becomes clear in Stebbing's book *Ideals and Illusions* from 1941. Here she emphasises not only the republican understanding of life in modern democracies, a constant questioning and evaluation on an empirical, scientific basis, but also the importance of utopian ideals for social change, such as those represented by the internationalism of the League of Nations – a hope that Otto Neurath also shared and which motivated his work.

All in all, the relevance and modernity of Stebbing and Neurath's approach, even from today's perspective, lies in the focus on the ordinary use and dissemination of information, in the questioning of the unlimited authority of

scientific experts and the ideal of science that was still predominantly held at the time – in favor of a process-based and ultimately pluralistic approach to combating propaganda and disinformation and supporting social change. This was based on an essentially unbroken trust in the possibility and effectiveness of one's own clear thinking and judgment and the confidence in science that could and would be tested again and again.

References

Beaney, Michael (2016) "Susan Stebbing and the Early Reception of Logical Empiricism in Britain", in: Ch. Damböck (ed.), *Influences on the Aufbau*. Dordrecht: Springer.

Bowler, Peter J. (2009), *Science for All: The Popularization of Science in Early Twentieth-Century Britain*. Chicago: University of Chicago Press.

Carnap, Rudolf (1963), *Intellectual Autobiography*, In P. A. Schilpp (ed.), *The Philosophy of Rudolf Carnap*, LaSalle, Illinois: Open Court, 3–84.

Cat, Jordi, and Tuboly Adam T. (eds.) (2019), *Neurath Reconsidered: New Sources and Perspectives*, Cham: Springer.

Chapman, Siobhan (2013), *Susan Stebbin and the Language of Common Sense*, Basingstoke: Palgrave Macmillan.

Cunha, Ivan Ferreira da; Linsbichler, Alexander (2024), "Science, Imagination and Values in the German Energy Turn: An Example of Neurath's Methodology for Social Technology", (Engl. Translation, <https://www.researchgate.net/publication/378768837>), *Revista Kriterion*, 65 (156), 673-700. doi: 10.1590/0100-512X2023n15603ifc.

Doulas, Louis (2022), "Making Sense of Stebbing and Moore on Common Sense", in: A. Coliva, L. Doulas (eds.), *Analysis, Common Sense, and Public Philosophy: Themes from the Philosophy of Susan Stebbing*, Oxford University Press.

Franco, Paul L. (2023), "Susan Stebbing on Logical Positivism and Communication", *Ergo*, Vol. 10 (48), 1378-1402.

Janssen-Lauret, Frederique (2022), *Susan Stebbing*, Cambridge: Cambridge University Press.

Körber, Silke (2019), "Thinking about the 'Common Reader': Otto Neurath, L. Susan Stebbing and the (Modern) Picture-Text Style", in : Cat and Tuboly (2019), 451-470.

Körber, Silke (2023), *Die Visualisierung von Wissen im "Jahrhundert der Augen": Otto Neurath, Isotype und Adprint*, Cham: Springer.

Neurath, Otto (1931/1983), "Physicalism: The philosophy of the Viennese Circle", in: R. S. Cohen and Marie (eds.), *Philosophical Papers 1913-1946*, Dordrecht: D. Reidel, 52-57.

Neurath, Otto (1939), *Modern Man in the Making*, New York: Knopf.

Neurath, Otto (1996), "Visual Education: Humanisation versus Popularisation", in: E. Nemeth and F. Stadler (eds.), *Encyclopedia and Utopia: The Life and Work of Otto Neurath (1882-1945)*, Dordrecht: Kluwer, 245-335.

Schliesser, Eric (2023), "Stebbing on Clarity", unpublished manuscript, <https://philarchive.org/archive/SCHSOC-13>.

Stadler, Friedrich and Thomas Uebel (eds.) (2012), *Wissenschaftliche Weltauffassung. Der Wiener Kreis*, Wien and New York: Springer.

Stebbing, L. Susan (1933), *Logical Positivism and Analysis*. From the proceedings of the British academy, Volume XIX, London: Humphrey Milford Amen House, E.C.

Stebbing, L. Susan (1934/1959), *Logic in Practice*. Fourth revised edition, London: Methuen & Co.

Stebbing, L. Susan (1937/1944), *Philosophy and the Physicists*, London: Methuen & Co.

Stebbing, L. Susan (1939), "Ethics and Materialism", *Ethics* 50 (1), 35-44.

Stebbing, L. Susan (1939/1959), *Thinking to some Purpose*, London: Penguin Books.

Stebbing, L. Susan (1941), *Ideals and Illusions*. London: Watts & Co.

Tuboly, Adam Tamas (2020), "Knowledge Missemiation: L. Susan Stebbing, C.E.M. Joad, and Philipp Frank on the Philosophy of the Physicists", *Perspectives on Science* 28 (1), 1–34.

Van der Schaar, Maria and Eric Schliesser (2017), "Women in Early Analytic Philosophy: Volume Introduction", *Journal for the History of Analytical Philosophy* 5 (2), 1-5.

Verhaegh, Sander; Petrovich, Eugenio; Boš, Gregor (2025), The American Reception of Logical Empiricism: A Mention-Based Bibliometric Analysis, in: G. Schiemer (ed.), *The Legacy of the Vienna Circle*, Vienna Circle Yearbook, Vol. 30, Cham: Springer (forthcoming).

Vlasits, Justin (2022), "Margaret MacDonald's Scientific Common-Sense Philosophy", *British Journal for the History of Philosophy* 30 (2), 267-287.

West, Peter (2022), "L. Susan Stebbing Philosophy and the Physicists (1937): a re-appraisal", *British Journal for the History of Philosophy* 30 (5), 859-873.

Acknowledgement

This article is in parts based on an analysis of Stebbing and Neurath's correspondence and cooperation "Thinking Clearly for a Modern Democracy Susan Stebbing and Otto Neurath" that I co-authored with Adam Tuboly in a forthcoming volume on Susan Stebbing, edited by Siobhan Chapman and Brian Garvey (Palgrave Macmillan, 2025).

Power and the ethics of innovating: defining deliberative injustice for hermeneutic technology assessment

Benedict Lane (Delft, The Netherlands)

Abstract

This paper explicates a definition of “deliberative injustice” for hermeneutic technology assessment. Hermeneutic technology assessment is motivated by the recognition that modes of technology assessment based on either the consequences of a technology (e.g., Brey, 2012) or the moral properties of the technology itself (e.g., Winner, 1980; Latour, 1992) are not suitable for moral assessment of new and emerging technologies (Swierstra and Rip, 2007). Such technologies are at such an early stage in there is little realistic hope of predicting their moral consequences or properties (Nordmann, 2010). Rather, the most morally salient features of such technologies are the processes by which they are coming into being – specifically, the form of the discourses that are giving meaning to these technologies (Grunwald, 2020), and the dynamics of power operative in those discourses (Sand, 2019). I develop a theoretical framework for assessing such dynamics of power based on David Lewis’s (1979) notion of conversational score and the notion of uptake developed in feminist ethics and philosophy of language (e.g., Austin, 1975; Langton, 1993; Hornsby, 1995; Hornsby and Langton, 1998; Kukla, 2014). I argue that the development of 21st century new and emerging technologies embody a particular kind of injustice that I call *deliberative injustice*: those set to be most affected by a new technology are systematically disempowered to affect the conversational score surrounding it in ways that are in their interests, including offering their own “visions” for the future development of the technology in question (Ferrari and Marin, 2014).

Hermeneutic technology assessment

It is commonly recognised that modes of technology assessment based on the impact of a technology (e.g., Brey, 2012), or the moral properties of its artefacts (e.g., Winner, 1980; Latour, 1992), are not suitable for the ethical assessment of the nascent technologies of the 21st century. Such technologies are at such an early stage in their development that there are no clear consequences or moral qualities there to be assessed – and there is little realistic hope of accurately anticipating them in advance (Nordmann, 2010). In lieu of concrete technological artefacts and measurable impacts to assess, the most salient loci of ethical analysis are the hermeneutic processes by which societal meaning is assigned to the *idea* of these technologies. In recognition of this, a mode of technology assessment has emerged specifically for dealing with the nascent technologies of the 21st century: hermeneutic technology assessment (e.g., Grunwald, 2014, 2017, 2020).

Such societal meaning comes in the form of *collectively imagined futures* arising from the projected co-evolution of a technology and the society into which it will be embedded. Such futures have been considered from a number of different angles in the literature. For example, the concept of a “sociotechnical imaginary”, a development of Charles Taylor’s concept of a “social imaginary” (2004), was introduced by Sheila Jasanoff and Sang-Hyun Kim (2009) to describe the fuzzy, implicit societal understandings and expectations of technologies and their potential that are sharpened and made explicit, for example, in the form of national innovation policy agendas. Jasanoff later expanded this conception of sociotechnical imaginaries beyond nation-states, in such a way that explicitly acknowledges the role of differential power and influence in how sociotechnical imaginaries take root: in general, sociotechnical imaginaries “can originate in the visions of single individuals or small collectives” which are elevated to the status of imaginaries through the exercise of social power, thereby “according them a dominant position for policy purposes” (2015, p. 4).

The orienting role of visions in societal deliberations about nascent technologies

Visions, then, can be understood as the candidate entities that get elevated to the status of dominant sociotechnical imaginaries through the exercise of social power. They are the units of selection in the evolutionary process by which societal meanings are attributed to new and emerging technologies. This evolutionary process, and the underappreciated structures of social power that operate as its selection pressures, will be the focus of ethical analysis in this paper.

Competing visions vie to orient and scaffold societal deliberation about new technologies in at least four ways (Ferrari and Lösch, 2017; Lösch, Heil and Schneider, 2017):

Possibility. First, visions act as “interfaces” between the present and future. In this regard, the visions recognised and in circulation at a given moment constitute a modal space, and thereby “open up imaginative and practical possibilities” and “enable the identification of options for change” (Loesch, Heil and Schneider, 2017, p. 142).

Intelligibility. Second, visions function as a pool of shared communicative and interpretative resources on which deliberators can draw. In this regard,

recognised visions serve as interpretative touchstones, to which deliberators can appeal in the claims they make on one another and in terms of which they can intelligibly express themselves.

Coordination. Third, visions serve to coordinate the heterogeneous practical activities involved in the innovation process, by establishing certain lines of inquiry as worthy of pursuit (or not worth the risk). In this regard, recognised visions ground divisions of labour and keep actors practically “on the same page”.

Normativity. Fourth, visions function not only to organise but also to obligate and motivate action. In this regard, the positive and negative affective valences of recognised visions are a recognised source of normative force; they “determine and establish legitimate normative expectations and create rights and obligations” (Lösch, Heil and Schneider, 2017, p. 142); as a corollary, “they tend to exclude alternative solutions of the described problems” (Ferrari and Lösch, 2017, pp. 78–79).

Changing the conversational score

What kind of social power is necessary for an individual to meaningfully and impactfully contribute to this practice of exchanging and assessing visions? A helpful device for answering this question is David Lewis’s notion of the “conversational score”, which has been extremely influential in the fields of linguistics and philosophy of language. As I hope to demonstrate, the concept is also highly relevant to hermeneutic technology assessment.

Lewis (1979) introduced the concept to describe the path-dependency and context-specificity of various aspects of our deliberative practices. Lewis recognised that deliberation at a given moment is constrained by the actions of deliberators up to that point. He illustrates the point by analogy with baseball. The score of a game of baseball evolves in a rule-governed way constitutive of the game. The score at a given moment also determines which subsequent moves players are entitled to make, according to yet other constitutive rules. According to Lewis, deliberation resembles a game of baseball in important ways. Like baseball, the state of a discourse at a given moment is captured by a kind of score – the conversational score. The conversational score is the common ground generally shared by participants in a deliberative practice, and on the basis of which their deliberation proceeds. It consists of all the

various elements that must be in place for a deliberation to get going, as well as the roughly predefined rails along which deliberation runs to keep it moving forwards. The conversational score evolves according to roughly specifiable rules. What is an acceptable conversational “move” – i.e., what deliberative contributions deliberators are entitled to make – at a given point in a deliberation also evolves with the conversational score, also in a roughly rule-like way.

Lewis identifies several components of the conversational score that align in striking ways with the four orienting functions of sociotechnical visions delineated above:

Relative modality. First, the component of the conversational score that Lewis refers to as the “relative modality” of a discourse at a given moment maps neatly onto the role of visions as *an interface between the past and the future*. As Lewis points out, when we deliberate, “[n]ot all the possibilities there are enter into consideration”:

The boundary between the relevant possibilities and the ignored ones ... is a component of conversational score, which enters into the truth conditions of sentences with “can” or “must” or other modal verbs. It may change in the course of conversation. (1979, p. 354)

In the context of societal deliberation about new technologies, such a relative modality will affect the (logically) possible sociotechnical futures one can reasonably be expected to consider when deliberating about a new technology. The recognised possibilities of a line of technological research have previously been described by van Lente and Rip as belonging to the “rhetorical space” associated with it: this space “is objectively ‘there’, because it enables and constrains. But it cannot be pointed at or measured, except through the actions that are enabled and constrained” (1998, p. 223). In this sense the possibilities of a research line are an aspect of social reality that guides and constrains societal deliberation about new technologies (1998, pp. 234–235).

The boundary between relevant and ignored possibilities changes according to what Lewis calls the “rule of accommodation” (1979, p. 347); all things being equal, when an individual appeals to a possibility that had previously been ignored, “[t]he boundary shifts outward so that what he says is true” (1979, p.

355). This possibility space can be changed when influential individuals articulate previously unappreciated possibilities: consider the seismic impact of Richard Feynman's interventions on the future development of nanotechnology (Feynman, 1992) and quantum simulation (Feynman, 1982) and on the possibilities considered by engineers and computer scientists. However, the effect of such interventions on the relative modality of a discourse depends on its receiving uptake by one's deliberative community. If one's intervention is met by scoffs or rolled eyes, the old relative modality remains in place.

Relative saliences and points of reference. Second, the role of visions as *shared communicative and interpretative resources* connects with two other components of conversational score considered by Lewis: relative saliences (1979, pp. 348–350) and points of reference (1979, pp. 350–351). Lewis analyses relative saliences in terms of the contextually-specific referents of definite descriptions of the form “the F” – according to Lewis, “‘the F’ denotes x if and only if x is the most salient F in the domain of discourse, according to some contextually determined salience ranking” (1979, p. 348). In the context of deliberations about new technologies, such contextually determined salience rankings will affect who or what is being talked about when we talk about, for example, “the stakeholders in a new technology”, “the problems that a new technology should address”, or “the harms/benefits of a new technology”.

Lewis analyses the notion of points of reference in terms of the sense one can make of a simple story. However, the notion of a deliberative point of reference can be generalised to include those elements of our shared common ground that enable us to make sense of each other's deliberative contributions and how they fit into the quasi-narrative “arc” of what is being said or done (cf. Pesch, 2021, p. 261). In the context of societal deliberation about new and emerging technologies, Swierstra and Rip identify “tropes”, “motifs”, and “storylines” that function as “a repertoire that is available in late-modern societies, both as a framing of how actors view issues and expect others to view them, and as a kind of toolkit that can be drawn upon in concrete debates” (2007, p. 4). This repertoire serves to make arguments about new technologies intelligible to their audience.

Again, relative saliences and points of reference evolve according to a version of the rule of accommodation, such that (all things being equal) these components of the conversational score adjust to make what has just been said acceptable in the relevant sense. Once again, what is said has this effect only if it receives uptake. If others treat what you say as nonsense, or as about the wrong thing, or not about anything, the relative saliences and points of reference remain as they were.

Provisions within a plan. Third, the role of visions as devices for coordinating the heterogeneous practical activities involved in the innovation process corresponds neatly to the component of the conversational score that has to do with the provisions contained within a plan (Lewis, 1979, p. 357). When the goal of deliberation is to form a plan to coordinate action (rather than simply to describe the world), certain deliberative contributions require for their acceptability certain provisions to be in place within the plan. All things being equal, if my contribution to the process of deliberation by which we put our plan together does not already contain the necessary provisions, the conversational score will change so that it does. In the context of new technologies, these provisions serve to carve out worthwhile sub-domains of scientific and technological research as well as shaping policy agendas to enact a planned-for sociotechnical future; the entire field of post-quantum encryption, for example, exists as a provision to counteract imagined future malicious uses of RSA-encryption-breaking quantum protocols like Shor's algorithm (Shor, 1994, 1997).

Yet again, what I say has this effect only if it receives uptake. If my contribution to the plan is deemed by others to be far-fetched, impractical, or unnecessary, or it is overruled by another proposal, the conversational score does not change to accommodate my contribution.

Permission. Fourth, and finally, the conversational score establishes a normative space within which participants in a discourse are permitted, obligated, or prohibited from certain courses of action, both within the discourse and beyond it (Lewis, 1979, pp. 340–341). This component of the conversational score maps satisfyingly onto the function of sociotechnical visions as a source of normative and motivational force. The normative space within which societal deliberation about new technologies takes place shapes

and is shaped by deliberation in several ways. One way of affecting this normative space is through performative, or illocutionary, acts (Austin, 1975; Searle, 1992; Searle and Vanderveken, 2009). Such acts include granting or refusing someone permission, giving someone an order, or swearing someone into office. In the context of societal deliberation about new technologies, the illocutionary act of *refusal* is perhaps especially relevant. For example, Zong and Matias suggest that we view acts of technology refusal as contributions to the process of sociotechnical design. They write that “one must understand the nature of design as a form of sociotechnical power that people resist and also take up in the act of resistance” (2024, p. 5) and that “refusal is an important form of participation in the design process, which expert designers never fully control” (2024, p. 6). However, the fact that one’s illocutionary acts depend on *uptake* by others to have their normal conventional consequences makes one vulnerable to a form of injustice that has come to be known as “illocutionary silencing” (Langton, 1993; Hornsby, 1995; Hornsby and Langton, 1998) or “discursive injustice” (Kukla, 2014): if an act of refusal fails to receive uptake as an act of refusal, it does not have the conventional consequences it would otherwise have, and the normative space occupied by deliberators remains as it was.

What I hope to have demonstrated is that the kind of social power required to participate meaningfully and impactfully in societal deliberation about new technologies is precisely *the power to change the conversational score* – in particular, its presuppositions, relative modality, relative saliences, points of references, provisions, and normative spaces. The elements of our deliberative practice scaffold present and future deliberation and action in the domain of technological innovation. These components map neatly onto the different dimensions of the orienting role of sociotechnical visions in societal deliberation about new technology: as an interfaces between present and future, as a pool of communicative and interpretative resources, as devices for coordinating collective projects, and as a source of normative obligations and expectations. Furthermore, the power to change the conversational score in all of these ways, and thus to steer the course of societal deliberation about new technologies through the practice of exchanging and assessing sociotechnical visions, is conditional on *uptake* from one’s deliberative community.

Defining deliberative injustice for hermeneutic technology assessment

These insights motivate the definition of deliberative injustice I propose for hermeneutic technology assessment:

Deliberative injustice involves prejudicially withholding uptake of an individual's attempts to affect the conversational score, in a way that does them both intrinsic and extrinsic harm.

As my emphasis on uptake reflects, the power to change the conversational score is not possessed by any individual inherently but is an aspect of their context-specific standing in the social reality they occupy, which agents jointly and continuously create, sustain, and transform through their interactions with one another. In this sense it is a component of their "social power" generally, their capacity to affect their social world (Searle, 1995). A great deal of any individual's power to affect their social world depends on the (socially constructed) "kind of person" they are (Hacking, 1996). Miranda Fricker refers to such identity-based social power as "identity power" (2007, p. 14). Identity power is highly context-specific: it is the capacity (i) *of someone like this* (ii) *to do something like this* (iii) *in a context like this*.

An individual whose identity power is attenuated by prejudice is always liable to harm in a straightforward way: they are unjustly excluded from accessing certain goods. However, in deliberative contexts, distorted attributions of such power harm individuals in less tangible but no less significant ways. These harms are constitutive of deliberative injustice as I understand it.

First, the individual is harmed specifically *in their capacity as a deliberator*. Deliberation is foundational to the human form of life; the capacity to participate in a deliberative practice is part of what makes us human and belongs to the grounds of the value of being human. In this way, an individual who suffers a prejudicial distortion of their capacity to participate in a practice of deliberation in which they have a stake suffers an "intrinsic harm"; they are "symbolically degraded *qua* human" (Fricker, 2007, p. 44).

Second, as Sally Haslanger astutely observes, "[w]e are the individuals we are today at least partly as a result of what has been attributed (and self-attributed) to us" (2012, p. 88). Thus, by attributing an individual with a deflated capacity to contribute to deliberation, we may bring it about, through

a process of “causal construction”, that they are *in fact less capable of so contributing* – less situationally attuned, less informed, less articulate, less motivated, etc. – thus confirming our initial prejudicial stereotyping of them (Fricker, 2007, p. 55). In this way, an individual who suffers a prejudicial distortion of their capacity to deliberate also suffers and “extrinsic harm”: their capacity to deliberate is *damaged* by our prejudicial treatment of them. Deliberative injustice therefore both degrades and damages an individual’s deliberative agency, doing them both intrinsic and extrinsic harm.

With this definition in hand, I hope that the growing field of hermeneutic technology assessment can come to pay more attention to the moral harm being done in the present as a result of structurally unjust power dynamics operative in contemporary forms of deliberation surrounding new and emerging technologies. This should not replace the attention that is being paid to the future harms that individuals stand to suffer due to the sociotechnical visions that receive uptake in the present, but rather to complement it. Deliberative injustice in the present and moral harm in the future are two aspects of the problem that hermeneutic technology assessment was introduced to ameliorate.

References

- Austin, J.L. (1975) *How To Do Things With Words: The William James Lectures delivered at Harvard University in 1955*. Oxford University Press. Available at: <https://doi.org/10.1093/acprof:oso/9780198245537.001.0001>.
- Brey, P.A.E. (2012) ‘Anticipatory Ethics for Emerging Technologies’, *NanoEthics*, 6(1), pp. 1–13. Available at: <https://doi.org/10.1007/s11569-012-0141-7>.
- Ferrari, A. and Lösch, A. (2017) ‘How Smart Grid Meets In Vitro Meat: on Visions as Socio-Epistemic Practices’, *NanoEthics*, 11(1), pp. 75–91. Available at: <https://doi.org/10.1007/s11569-017-0282-9>.
- Ferrari, A. and Marin, F. (2014) ‘Responsibility and Visions in the New and Emerging Technologies’, in S. Arnaldi et al. (eds) *Responsibility in Nanotechnology Development*. Dordrecht: Springer Netherlands, pp. 21–36. Available at: https://doi.org/10.1007/978-94-017-9103-8_2.
- Feynman, R., P. (1992) ‘There’s plenty of room at the bottom’, *Journal of Microelectromechanical Systems*, 1(1), pp. 60–66. Available at: <https://doi.org/10.1109/84.128057>.

Feynman, R.P. (1982) 'Simulating physics with computers', *International Journal of Theoretical Physics*, 21(6), pp. 467–488. Available at: <https://doi.org/10.1007/BF02650179>.

Fricker, M. (2007) *Epistemic Injustice*. Oxford University Press. Available at: <https://doi.org/10.1093/acprof:oso/9780198237907.001.0001>.

Grunwald, A. (2014) 'The hermeneutic side of responsible research and innovation', *Journal of Responsible Innovation*, 1(3), pp. 274–291. Available at: <https://doi.org/10.1080/23299460.2014.968437>.

Grunwald, A. (2017) 'Assigning meaning to NEST by technology futures: extended responsibility of technology assessment in RRI', *Journal of Responsible Innovation*, 4(2), pp. 100–117. Available at: <https://doi.org/10.1080/23299460.2017.1360719>.

Grunwald, A. (2020) 'The objects of technology assessment. Hermeneutic extension of consequentialist reasoning', *Journal of Responsible Innovation*, 7(1), pp. 96–112. Available at: <https://doi.org/10.1080/23299460.2019.1647086>.

Hacking, I. (1996) 'The looping effects of human kinds', in D. Sperber, D. Premack, and A.J. Premack (eds) *Causal Cognition: A Multidisciplinary Debate*. Oxford University Press, p. 0. Available at: <https://doi.org/10.1093/acprof:oso/9780198524021.003.0012>.

Haslanger, S. (2012) 'Ontology and Social Construction', in *Resisting Reality: Social Construction and Social Critique*. Oxford University Press, p. 0. Available at: <https://doi.org/10.1093/acprof:oso/9780199892631.003.0002>.

Hornsby, J. (1995) 'Disempowered Speech', *Philosophical Topics*, 23(2), pp. 127–147. Available at: <https://doi.org/10.5840/philtopics199523211>.

Hornsby, J. and Langton, R. (1998) 'Free Speech and Illocution', *Legal Theory*, 4(1), pp. 21–37. Available at: <https://doi.org/10.1017/S1352325200000902>.

Jasanoff, S. (2015) 'Future Imperfect: Science, Technology, and the Imaginations of Modernity', in S. Jasanoff and S.-H. Kim (eds) *Dreamscapes of Modernity: Sociotechnical Imaginaries and the Fabrication of Power*. University of Chicago Press, p. 0. Available at: <https://doi.org/10.7208/chicago/9780226276663.003.0001>.

Jasanoff, S. and Kim, S.-H. (2009) 'Containing the Atom: Sociotechnical Imaginaries and Nuclear Power in the United States and South Korea', *Minerva*, 47(2), pp. 119–146. Available at: <https://doi.org/10.1007/s11024-009-9124-4>.

Kukla, R. (2014) 'Performative Force, Convention, and Discursive Injustice', *Hypatia*. 2020/01/01 edn, 29(2), pp. 440–457. Available at: <https://doi.org/10.1111/j.1527-2001.2012.01316.x>.

Langton, R. (1993) 'Speech Acts and Unspeakable Acts', *Philosophy & Public Affairs*, 22(4), pp. 293–330.

- Latour, B. (1992) 'Where are the missing masses? The sociology of a few mundane artifacts', *Shaping technology/building society: Studies in sociotechnical change*, 1, pp. 225–258.
- van Lente, H. and Rip, A. (1998) 'The Rise of Membrane Technology: From Rhetorics to Social Reality', *Social Studies of Science*, 28(2), pp. 221–254. Available at: <https://doi.org/10.1177/030631298028002002>.
- Lewis, D. (1979) 'Scorekeeping in a language game', *Journal of Philosophical Logic*, 8(1). Available at: <https://doi.org/10.1007/BF00258436>.
- Lösch, A., Heil, R. and Schneider, C. (2017) 'Responsibilization through visions', *Journal of Responsible Innovation*, 4(2), pp. 138–156. Available at: <https://doi.org/10.1080/23299460.2017.1360717>.
- Nordmann, A. (2010) 'A forensics of wishing: technology assessment in the age of technoscience', *Poiesis & Praxis*, 7(1), pp. 5–15. Available at: <https://doi.org/10.1007/s10202-010-0081-7>.
- Pesch, U. (2021) 'Imaginarities of innovation: Turning technology development into a public issue', *Science and Public Policy*, 48(2), pp. 257–264. Available at: <https://doi.org/10.1093/scipol/scab017>.
- Sand, M. (2019) 'On “not having a future”', *Futures*, 107, pp. 98–106. Available at: <https://doi.org/10.1016/j.futures.2019.01.002>.
- Searle, J.R. (1992) *Speech acts: an essay in the philosophy of language*. Cambridge: Cambridge University Press.
- Searle, J.R. (1995) *The construction of social reality*. London: Lane, The Penguin Press.
- Searle, J.R. and Vanderveken, D. (2009) *Foundations of illocutionary logic*. Digitally printed version. Cambridge: Cambridge Univ. Press.
- Shor, P.W. (1994) 'Algorithms for quantum computation: discrete logarithms and factoring', in *Proceedings 35th Annual Symposium on Foundations of Computer Science. Proceedings 35th Annual Symposium on Foundations of Computer Science*, pp. 124–134. Available at: <https://doi.org/10.1109/SFCS.1994.365700>.
- Shor, P.W. (1997) 'Polynomial-Time Algorithms for Prime Factorization and Discrete Logarithms on a Quantum Computer', *SIAM Journal on Computing*, 26(5), pp. 1484–1509. Available at: <https://doi.org/10.1137/S0097539795293172>.
- Swierstra, T. and Rip, A. (2007) 'Nano-ethics as NEST-ethics: Patterns of Moral Argumentation About New and Emerging Science and Technology', *NanoEthics*, 1(1), pp. 3–20. Available at: <https://doi.org/10.1007/s11569-007-0005-8>.

Taylor, C. (2004) *Modern Social Imaginaries*. Duke University Press. Available at: <https://doi.org/10.1215/9780822385806>.

Winner, L. (1980) 'Do Artifacts Have Politics?', *Daedalus*, 109(1), pp. 121–136.

Zong, J. and Matias, J.N. (2024) 'Data Refusal from Below: A Framework for Understanding, Evaluating, and Envisioning Refusal as Design', *ACM Journal on Responsible Computing*, 1(1), pp. 1–23. Available at: <https://doi.org/10.1145/3630107>.

Acknowledgement

Preparation of this work was supported by the Dutch National Growth Fund (NGF), as part of the Quantum Delta NL programme.

A Note on Moore's 'Combined Reading' of the *Tractatus*

David Lindeman (Washington, D.C., United States)

Abstract

After setting out the primary interpretive puzzle of the *Tractatus*, I contrast the orthodox and resolute readings of the work. I turn then to Adrian Moore's 'combined reading' and contrast this in turn with what I call the Heideggerian reading. There are a number of resonances between Moore's reading and my reading, but also some key differences. While I agree with Moore that the sense we can make of propositional sense-making is non-propositional, and that making sense of propositional sense-making is an aim of the *Tractatus*, there is also, I submit, an overarching concern for what is left out of propositional sense-making beyond the non-propositional sense we can make of this: a concern for what Wittgenstein believes even the non-propositional sense-making presupposed by propositional sense-making presupposes, namely what he labels 'what is higher', 'the mystical'. It is my position that Wittgenstein's working out the limits of propositional sense-making by—as he says in the preface—'working outwards' from what can be said leads him to this concern; and that this is the way in which—as Russell puts it in his introduction—the logic naturally leads to mysticism. What Wittgenstein finds is the need for a deeper logic in the sense of the study of *logos*, the world-disclosing event of sense required for the propositional sense-making codified in logic narrowly construed. This observation requires greater elaboration, and this is the primary lacuna in Moore's reading. I suggest that several ideas from Heidegger help us to spell out the details.

If what I want to teach is, not a more correct way of thinking, but a new movement of thought, then I am aiming for a 'revaluation of values' and come to resemble Nietzsche, also because in my view philosophers should be poets.
(MS 120, trans. Schulte 2013)

I

In the preface to the *Tractatus*, Wittgenstein informs the reader that "the aim of the book is to draw a limit to thought, or rather—not to thought, but to the expression of thoughts". The immediate amendment included in this statement of the aim is explained as follows:

for in order to be able to draw a limit to thought, we should have to find both sides of the limit thinkable (i.e. we should have to be able to think what cannot be thought). It will therefore only be in language that the limit can be drawn, and what lies on the other side of the limit will simply be nonsense.

Wittgenstein ends the preface by remarking that if there is any value in the book it consists principally in two things: (1) "that thoughts are expressed in

it", the "better expressed" the "greater" the value, and (2) that the book "shows how little is achieved" when "the problems"—concerning the relations between language, thought, and world—"are solved", as he takes them to be with the *Tractatus*, "the *truth* of the thoughts that are [there] communicated" being, so it seems to Wittgenstein, he tells us, "unassailable and definitive".

Initially, what the work seems to provide is a theory of the relations between language, thought, and world which divides sense from nonsense in such a manner as to solve the philosophical problems concerning such relations. According to this theory, the function of language is to picture states of affairs, certain arrangements of objects (the 'simples' arrived at by logical analysis); the meaning of a sentence is what it pictures; and the sentence is true just in case that state of affairs obtains, that is, just in case what is pictured is a fact; where the picturing relation obtains just in case there is a correspondence between elements of the picture and elements of the state of affairs pictured (TLP 2.1, 2.11, 2.12, 2.13, 2.131, 2.14, 2.15, 2.1511, etc.). A sentence with sense, a proposition or thought, is possibly true and possibly false, depending on whether the relevant state of affairs obtains (TLP 4, 2.201, 2.21, etc.). Nonsense (*Unsinnig*) is neither possibly true nor possibly false (TLP 4.003). Besides this, there are the 'senseless' (*sinnlos*) propositions of logic: tautologies which rule out no possible states of affairs, and contradictions which rule out all (TLP 4.461, 4.462). These are said to be true and false, respectively, but as limit cases. The propositions of science are the paradigmatic propositions (TLP 4.11).

The sentences (propositions) of the *Tractatus*, including those apparently laying out this theory, do not concern objects. They concern, as just indicated, logical form, the relationship between language, thought, and world, and also such matters as the metaphysical subject, value, and God. At TLP 5.632-5.6311, 5.641 we are told that the metaphysical subject is not an object and accordingly cannot be the subject of propositions. Propositions concern facts, the totality of which compose the world. They do not concern matters of value.

6.4 All propositions are of equal value.

6.41 The sense of the world must lie outside the world. In the world everything is as it is, and everything happens as it does happen: *in* it no value exists—and if it did exist, it would have no value. [...]

- 6.42 So too it is impossible for there to be propositions of ethics.
Propositions can express nothing that is higher.
- 6.421 It is clear that ethics cannot be put into words.
Ethics is transcendental.
(Ethics and aesthetics are one and the same.)
- 6.4312 [...] The solution of the riddle of life in space and time lies *outside* space and time.
(It is certainly not the solution of any problems of natural science that is required.)
- 6.432 *How* things are in the world is a matter of complete indifference for what is higher.
God does not reveal himself *in* the world.
- 6.44 It is not *how* things are in the world that is mystical, but *that* it exists.
- 6.45 To view the world sub specie aeterni is to view it as a whole—a limited whole.
Feeling the world as a limited whole—it is this that is mystical.
- 6.522 There are, indeed, things that cannot be put into words. They *make themselves manifest*.
They are what is mystical.

What is higher, the mystical, cannot be the subject of propositions. These sentences accordingly do not picture states of affairs and must therefore be nonsense. “The totality of true propositions is the whole of natural science (or the whole corpus of the natural sciences)” (TLP 4.11). Whereas: “Philosophy is not one of the natural sciences. (The word ‘philosophy’ must mean something whose place is above or below the natural sciences, not beside them.)” (TLP 4.111). Famously, then, we have the penultimate line of the *Tractatus*:

- 6.54 My propositions serve as elucidations in the following way: anyone who understands me eventually recognizes them as nonsensical, when he has used them—as steps—to climb up beyond them. (He must, so to speak, throw away the ladder.)

He must transcend these propositions, and then he will see the world aright.

Seeking to make sense of the text as a whole in light of this remark, two interpretations of the *Tractatus* have dominated discussion, the orthodox and resolute readings.

II

According to the orthodox reading (see Anscombe 1959, Hacker 1972/1986), the nonsensical but elucidatory sentences of the *Tractatus*—elucidations—intend but fail to *say* what sentences with sense and the senseless propositions of logic *show*. Wittgenstein draws the distinction in the context of discussion of logical form as follows:

4.116 Propositions can represent the whole of reality, but they cannot represent what they must have in common with reality in order to be able to represent it—logical form.

In order to be able to represent logical form, we should have to be able to station ourselves with propositions somewhere outside logic, that is to say outside the world.

4.121 Propositions cannot represent logical form: it is mirrored in them.

What finds its reflection in language, language cannot represent.

What expresses *itself* in language, *we* cannot express by means of language.

Propositions *show* the logical form of reality.

They display it.

4.1212 What *can* be shown, *cannot* be said.

The trouble is that the *Tractatus* is, apparently, full of propositions about logical form. But then, by the foregoing, these 'propositions' cannot be propositions proper, nor give expressions to thoughts if these are taken in the sense of TLP 4; they are nonsense. But not mere nonsense but a species of nonsense singled out as elucidatory. They have the significant function of indicating, by their failure to say what they intend to say, what propositions with sense show. They do not show what propositions with sense show, nor say what these say. The resolute reading (see Diamond 1988, Conant 1989)

opposes this notion of 'significant' or 'substantial nonsense'. Rather, on this view, all nonsense is on a par—mere nonsense. All but a few sentences of the work, if not indeed the whole work, the picture theory and the say/show distinction included, must be 'thrown away'. The reader is led to recognize nonsense for nonsense. There are no ineffable truths, nothing that cannot be put into words.

It is difficult to square the resolute reading with the relevant extra-textual evidence (see Monk 1990), but let us stick to the text. Again, we read: "There are, indeed, things that cannot be put into words. They *make themselves manifest*. They are what is mystical." (TLP 6.522) What we are told is that *there are* things of which we cannot speak, that is, things we cannot picture. That this is so is pointed out—and problematized—in the famous quips of Otto Neurath and Frank Ramsey. But if we are to make some sense of this, what are these things?

In addressing just this question, Adrian Moore (2012) in *The Evolution of Modern Metaphysics: Making Sense of Things* considers two possibilities: truths and states of understanding. He rules out the first, on the ground that *ineffable truths* are "anathema" to Wittgenstein (240); Moore cites TLP 4.002, 4.063, 4.5, 6.5-6.51. In support of the second, he notes TLP 6.54. Moore lays emphasis there on "understands me", the idea being that Wittgenstein's 'propositions', which are nonsensical (and so not propositions proper), are not understood; Wittgenstein is. This distinction in hand, Moore reconciles the two readings—which he calls the traditional and new readings—as follows:

Where advocates of the traditional reading hold that there are 'things' that cannot be put into words, and that the *Tractatus* conveys 'things' of that sort, we can construe the 'things' in question as states of understanding, including the state of understanding Wittgenstein. Where advocates of the new reading hold that there is 'nothing' that cannot be put into words, and that the *Tractatus* conveys 'nothing' whatsoever, we can construe the 'things' in question as truths. (241)

On Moore's 'combined reading', the two readings are seen to be "consonant". The aim of the *Tractatus* is to make sense of propositional sense-making and what is uncovered is that the kind of sense-making required, indeed the only kind possible, is non-propositional.

[The *Tractatus*] is designed to help us make maximally general sense of things. But since the sense that it is designed to help us make is non-propositional, the means that it uses are indirect. It works through a creative use of nonsense. It is more like a work of art than like a work of science. (254)

Coming to understand Wittgenstein is coming to have this non-propositional understanding, where this involves an ability to recognize nonsense for nonsense (241-2). I offer an alternative reading, in some respects consonant with Moore's.

III

While I agree with Moore that the sense we can make of propositional sense-making is non-propositional, and that making sense of propositional sense-making is an aim of the *Tractatus*, there is also in the *Tractatus* an overarching concern for what is left out of propositional sense-making beyond the non-propositional sense we can make of this—a concern for what Wittgenstein believes even the non-propositional sense-making presupposed by propositional sense-making presupposes, namely what he labels 'what is higher', the 'mystical'. Engelmann's (1967) understanding of Wittgenstein seems most appropriate here: "Positivism holds—and this is its essence—that what we can speak about is all that matters in life. *Whereas Wittgenstein passionately believes that all that really matters in human life is precisely what, in his view, we must be silent about.*" (97) "Even if language were nothing but a depiction of sensually perceptible reality together with the conclusions obtained from it by abstraction, and if accordingly it were impossible to speak in any language about the higher sphere (as science and a philosophy conducted by scientific means are indeed *unable* to do)—even then there *exists* the higher sphere, there is a *sense* in our existence, there exists that from which values derive their value (and which confers it upon them 'from outside the world')." (110) I believe that Wittgenstein's working out the limits of propositional sense-making by—as he puts it in the preface—"working

outwards" from what can be said leads him to this concern, and that this is way in which—as Russell puts it in his introduction—the logic naturally leads to mysticism. It is my contention that any adequate interpretation must give pride of place to this concern.

It is given voice by the tone of the work, a tone signaling the spirit in which it is written (the importance Wittgenstein attaches to this phenomenon is clear in remarks collected in *Culture & Value*), a spirit which Wittgenstein risks making more explicit in the 'Lecture on Ethics', where he speaks of wonder at the existence of the world. The mystical, again, is "not *how* the world is but *that* it is" (TLP 6.44).

5.552 The 'experience' that we need in order to understand logic is not that something or other is the state of things, but that something *is*: that, however, is *not* an experience.

Logic is *prior* to every experience—that something is *so*.

It is prior to the question 'How?', not prior to the question 'What?'

5.5521 And if this were not so, how could we apply logic? We might put it in this way: if there would be a logic even if there were no world, how then could there be a logic given that there is a world?

What Wittgenstein finds, I submit, is the need—a need also recognized by Heidegger—for a deeper logic in the sense of the study of *logos*, the world-disclosing event of sense required for the propositional sense-making codified in logic narrowly construed; and it is here, and for this reason, that—again not unlike Heidegger—Wittgenstein turns to poetry. "I believe I summed up where I stand in relation to philosophy when I said: really one should write philosophy only as one *writes a poem*." (CV 28e) It has been argued that Wittgenstein turned to poetry only later (see Klagge 2021), after his return to Cambridge, but I believe to the contrary that from the beginning Wittgenstein writes philosophy as poetry. (His turning his back to the members of the Vienna Circle, interested to understand the *Tractatus*, to read from the poetry of Rabindranath Tagore was, I think, a telling gesture.) So indeed I agree with Moore that the *Tractatus* is more like a work of art than like a work of science. Recognizing the art of the work is key to its understanding. But this is a fact not sufficiently reflected in Moore's reading.

The creativity of which Moore speaks is deployed in service of imparting a non-propositional understanding of propositional understanding (an understanding of logical form, the relationship between language, thought, and world), not a non-propositional understanding of what *that* presupposes: what is higher. What's more—and by Moore's admission; see n. 89, p. 254—there is no accounting in his account for Wittgenstein's remarks in the preface, that if the book has any value it consists in that thoughts are expressed in it; that these thoughts are, it seems to the author, *true*; and that perhaps *they* “will be understood only by someone who has himself already had [these] thoughts—or at least similar thoughts”—occasioned, we might add, by similar *experience*. If Wittgenstein is understood, it is by way of his ‘propositions’, that is, his elucidations, the ‘thoughts’ expressed in the book, which, to pick up on what was just quoted, require, for *their* understanding, appropriation on the part of the reader (compare Kierkegaard's *Concluding Unscientific Postscript*, another plausible origin of the image of the ladder)—require, that is, a certain experience which is tantamount to ‘seeing the world aright’: seeing it as a miracle. Again, we have a connection to Heidegger, who identifies artworks as sites not just of beauty but truth (see e.g. “On the Origin of the Work of Art”). And in this connection, a remark from Wittgenstein: “We might say: art *discloses* the miracles of nature to us.” (CV 64e) (The rest of the remark goes on to echo the mystic Angelus Silesius, echoed also in Heidegger.)

But in what sense, then, are these ‘thoughts’ *true*? They are not true in that they correspond to the facts. They do not concern facts, which are the province of the sciences. They concern what is higher, which is the concern of philosophy (being either above or below but not beside the sciences). They are, as we might put it, higher truths. But in what sense, again, are they true? It is here I believe that the work of Heidegger is most helpful. The project of metaphysics, according to both Heidegger and Wittgenstein, understood by both to have been a matter of making propositional sense of things, has been completed in science. What follows, if there is to be a post-metaphysical philosophy, is what Heidegger (1977) calls the ‘task of thinking’, which he likens in many places to poeticizing. What I suggest here is that those elucidations concerning the higher play the same role as Heidegger's ‘thoughts’. The truth of these thoughts is a matter of disclosure—that is, therefore, truth in the sense of *aletheia*, as Heidegger understands this. The

form this thinking takes—in one who has let it take hold, having been seized by the truth, by what has been seen—is that of a thanking. (*Denken heißt danken*, Heidegger (1968) says.) The form it takes, in a word, is poetry: grateful wondering at the “wonder [*Wunder*] of wonders: that what-is is”, to pick up a remark from Heidegger’s (1949) “Postscript to ‘What is Metaphysics?’”. These elucidations, or in Heidegger’s terminology, ‘formal indications’, ‘signposts’ (*Wegweiser*)—designed to show (*weiser*) the way (*Weg*)—were rather sayings in the sense of disclosures, language bent into a world-disclosing gesture, concerning, as they do, not how things are in the world but the world as a whole, *that it is*—and feeling at one with which, being its limit, is the mystical.

And so, there are the double meanings of ‘silence’ in the final line of the work: “What we cannot speak about we must pass over in silence” (TLP 7). It can be read as description or prescription. If speaking is picturing, one cannot say anything with language which does not picture. If being silent is not speaking, in this sense, one cannot but keep silent. But one should not try to say (in the sense of picture) what cannot be pictured. (Compare the biblical injunction against graven images.) Those elucidations concerning what is higher do not try to picture. In this sense, though they break the silence, they, too, maintain the silence. Just as music and poetry do. “The way music speaks. Do not forget that a poem, although it is composed in the language of information, is not used in the language-game of giving information.” (Z 28)

References

- Anscombe, G.E.M. (1959) *An Introduction to Wittgenstein's Tractatus*. London: Hutchinson University Library.
- Conant, James (1989) “Must We Show What We Cannot Say?”, in R. Fleming and M. Payne (eds.) *The Senses of Stanley Cavell*, Lewisburg: Bucknell University Press, 242–83.
- Diamond, Cora (1988) “Throwing Away the Ladder: How to Read the *Tractatus*”, *Philosophy*, 63, 5–27.
- Engelmann, Paul (1967) *Letters from Ludwig Wittgenstein, with a Memoir*, Oxford: Blackwell.
- Heidegger, Martin (1949) “Postscript to ‘What is Metaphysics?’”, in Werner Brock (ed.), *Existence and Being*. Chicago: Gateway.
- Heidegger, Martin (1968) *What is Called Thinking?* Trans. J. Glenn Gray. Harper & Row.

Heidegger, Martin (1977) "On the Origin of the Work of Art", in David Farrell Krell (ed.), *Basic Writings*, New York: Harper & Row.

Heidegger, Martin (1977) "The End of Philosophy and the Task of Thinking", in David Farrell Krell (ed.), *Basic Writings*, New York: Harper & Row.

Kierkegaard, Soren (1990) *Concluding Unscientific Postscript*. Trans. Julia Watkin. Princeton: Princeton University Press.

Klagge, James (2021) *Wittgenstein's Artillery: Philosophy as Poetry*. Cambridge: MIT Press.

Monk, Ray (1990) *Ludwig Wittgenstein: The Duty of Genius*. Oxford: Free Press.

Moore, Adrian (2012) *The Evolution of Modern Metaphysics: Making Sense of Things*. Cambridge: Cambridge University Press.

Schulte, Joachim (2013) "Wittgenstein on Philosophy as Poetry", in M Molder, D. Soeiro, and N. Fonesca. Bern: Peter Lang (eds.), *Morphology: Questions on Method and Language*. Bern: Peter Lang.

On Perceptual Change

Camila Lobo (Lisbon, Portugal)

Abstract

Standpoint theorists have long argued that the experience of marginalization does not automatically produce critical standpoints on the social world, although it offers epistemic advantages that are not similarly available to dominantly situated subjects. Some critics claim that this epistemic privilege thesis retains the problematic implication that, since dominant subjects have not experienced marginalization, they are prevented from developing a critical standpoint that adequately illuminates marginalized lives. Emily Tilton, for instance, has recently proposed a standpoint approach that rejects what she terms a “strong epistemic disadvantage thesis” — the implausible idea that only those with firsthand experiences of marginalization can cultivate critical standpoints. However, Tilton also contends that no necessary connection exists between marginalized perspectives and the development of genuinely critical standpoints. In this paper, I challenge this latter claim by arguing that dominant subjects can — and indeed must — cultivate particular ways of seeing that are not immediately available from within their subject positions if they are to develop critical understandings of the social world. This argument for perceptual transformation draws support from contemporary social epistemologists such as Charles Mills and Gayle Pohlhaus, who contend that resisting epistemic oppression requires more than the accumulation of knowledge — it demands a fundamental change in how we perceive the world.

Standpoint theorists have long argued that starting inquiry from the perspective of socially marginalized subjects enables the development of critical standpoints that adequately illuminate the structures of domination that permeate social reality. Importantly, these critical standpoints do not automatically emerge from marginalized lives but must be collectively struggled for on the philosophical and political level. As Sandra Harding puts it with regard to feminist standpoints: “This need for struggle emphasizes the fact that a feminist standpoint is not something that anyone can have simply by claiming it. It is an achievement.” (Harding 1991: 127). Despite their insistence on the collective, reflective, and achieved character of standpoints, standpoint theorists have often been accused of championing an “automatic privilege thesis” according to which subjects that occupy marginalized social locations “automatically know more, or know better” by virtue of being marginalized (Wylie 2003: 28).

Over the last decades, standpoint theorists' emphasis on the achieved character of critical standpoints effectively dispelled criticisms that charge them with upholding an automatic privilege thesis. However, some critics

claim that standpoint projects retain a problematic implication regarding the capacity of dominant subjects to achieve a standpoint insofar as they maintain that projects of critical understanding may only be carried out by marginalized subjects since dominant subjects are bound to ignorance about marginalized experience. This criticism has been recently articulated by Emily Tilton who argues that the widespread repudiation of the automatic privilege thesis among standpoint theorists has often given place to what she calls a “strong epistemic disadvantage thesis” according to which “dominant social positions impose strong, substantive limits on what the socially dominant can know about the oppression of others” (Tilton 2024: 1). The limits are substantive, according to Tilton, because dominantly situated knowers are said to be missing matters of great importance, and not just “minor or trivial details” (Tilton 2024: 1). They are strong because “the socially dominant cannot break free of them; their ignorance is the inescapable result of their dominant social positions” (Tilton 2024: 1). While the former conclusion is regarded as plausible, the latter is seen as “theoretically implausible and politically pernicious” since it represents an excuse for ignorance that can be mobilized by dominantly situated subjects against marginalized groups (Tilton 2024: 2). This move occurs, according to Tilton, as standpoint theorists strive to distinguish their view from the suggestion, arising from the automatic privilege thesis, that marginalization is a *sufficient* condition for the achievement of a standpoint. In response, they claim that “marginalization is not sufficient but *necessary*”, thus avowing a necessary connection between marginalized perspectives and political standpoints (Tilton 2024: 4). Tilton goes on to argue that standpoint theorists must reject this relation of necessity if they are to avoid ratifying a pernicious excuse for ignorance within supposedly progressive scholarship and emancipatory practice – what we could call, following Gayle Pohlhaus, a form of willful ignorance (Pohlhaus 2012).

In what follows, I will not strive to ascertain whether the charge leveled against standpoint theories is correct, but to inquire whether the claim that standpoint projects necessarily draw from marginalized perspectives commits us to the idea that the experience of marginalization is necessary for the development of a standpoint. While I maintain that dominantly situated subjects can participate in projects that adequately illuminate marginalized

experiences, I note that this may require them to undergo a perceptual transformation, which effectively sets high standards for epistemic justice. Hence, I want to introduce some nuance to Tilton's argument that "standpoint theorists must deny that there is a necessary connection between a marginalized social position and the possibility of achieving a standpoint" (Tilton 2024: 2). To defend that argument, Tilton's strategy involves clarifying the perspective-standpoint distinction at the heart of standpoint theories: "although everyone will have a *perspective*", she writes, "not everyone – not even all marginalized people – will have a *standpoint*" (Tilton 2024: 3). She goes on to explain that marginalized social positions provide perspectival advantages to the development of a critical standpoint – namely, by providing questions from which to begin our inquiries, as we learn from Harding –, but are not required for that purpose. While deprived of the same perspectival advantages, dominantly situated subjects may similarly engage in the hard work of developing a standpoint "by engaging in critical reflection that centers experiences that are not their own" (Tilton 2024: 14). This articulation of the capacity of dominantly situated subjects to develop a critical understanding of the forms of social organization that benefit them adequately avoids the implausible and politically pernicious claim that only subjects who *experience* marginalization can develop a critical consciousness about the social world. Since dominant subjects cannot, by definition, go through an experience of marginalization, Tilton is right to claim that this relation of necessity prevents us from holding dominantly situated subjects accountable because it characterizes their ignorance "as passive and inevitable" (Tilton 2024: 8). However, if the rejection of a relation of necessity between experience and reflection also implies the rejection of a necessary connection between perspectives and standpoints, we run the risk of understating the difficulty that characterizes the work of cultivating a critical consciousness – especially when we start from socially dominant positions. In fact, it seems possible that, at least in some cases, achieving a standpoint demands a transformation at the perspectival level, that is, a perceptual change.

We can reject the idea that a firsthand experience of marginalization is required to develop a standpoint, while maintaining that a different relation of necessity is implied by the achievement thesis we have inherited from standpoint theory. This relation may be understood as the idea that the

development of a standpoint *necessarily* taps into resources that are only available from the *perspective* of marginalized subjects. This perspective-standpoint interplay translates a philosophical commitment to epistemic non-neutrality understood as the idea that certain aspects of the world only come into view from particular situated locations (an idea, I must add, that can be defended by reference to Wittgenstein's philosophy). On this view, while dominantly situated subjects face perspectival disadvantages when it comes to making sense of the structures of domination that disempower marginalized subjects, they can overcome the epistemic limitations of their social position by incorporating the perspectives of the marginalized into their conceptual world. This understanding is, of course, congenial to Tilton's argument that "although dominant social positions confer epistemic disadvantages, they do not *determine* ignorance" (Tilton 2024: 9). However, my emphasis on the necessary relation between standpoints and perspectives is meant to make us grapple with the possibility that the achievement of an adequately reflective standpoint on the social world does not simply require an intellectual exercise of deriving knowledge from marginalized social locations. To the extent that some aspects of the world are not visible from the position of dominant subjects, achieving a standpoint may sometimes demand a transformation of our ways of seeing such that we come to acquire another subject's point of view.

In order to explain the pernicious character of the strong epistemic disadvantage thesis, Tilton resorts to the work of social epistemologists such as Charles Mills and Gayle Pohlhaus, who argue that the ignorance of dominantly situated subjects is an actively cultivated state, rather than a passive one. As Tilton puts it, the "socially dominant benefit from the social order, and it is not in their interest to put in the effort to identify and understand the mechanisms that uphold it, as doing so would threaten the privileges their place in society affords them" (Tilton 2024: 14). Indeed, the idea that dominantly situated subjects have an interest in protecting their ignorance about those parts of the world that would make visible the injustices that maintain their privilege is at the core of the work developed by authors working on "epistemologies of ignorance". However, in theorizing these non-accidental forms of ignorance,

these authors usually emphasize the impossibility that dominant subjects come to a critical understanding of social reality without first undergoing some form of perceptual transformation.

Writing about “white ignorance”, Mills speaks of a form of cognitive disablement that entails “the refusal to perceive systemic discrimination, the convenient amnesia about the past and its legacy in the present, and the hostility to black testimony” (Mills 2007: 35). White ignorance is one of many expressions of what Pohlhaus calls “willful ignorance”, whereby the dominant systematically neglect, ignore, or distort the conceptual resources of the marginalized. When dominantly situated subjects insist on using criteria, concepts, and words unsuited for making sense of marginalized experiences, they willfully refuse “to acknowledge and to acquire the necessary tools for knowing whole parts of the world” (Pohlhaus 2012: 729). This kind of situated ignorance is willful not because it is easily preventable – in fact, it may be extremely difficult to overcome – but because it is not necessary (Pohlhaus 2012: 729). While primarily affecting the cognitive lives of dominant subjects, willful ignorance reinforces oppression by perpetuating ideological imaginaries and compromising the epistemic agency of marginalized subjects.

White ignorance is particularly illustrative of this form of epistemic oppression since the geographical segregation that often accompanies racial domination seems to widen the gap between the situated perspectives of dominant and non-dominant subjects. As Mills explains, since dominantly situated white subjects have historically resisted integrating with Black communities, they have also actively rejected incorporating perspectives that would afford them a better understanding of the social world. This creates an “adversarial dynamic” that makes it impossible to resolve white ignorance by bringing Black perspectives “additively into relation with the ideology of white domination, since it is in fundamental opposition to it” (Mills 2017: 107). Hence, cases of white ignorance distinctly highlight the difficulties associated with achieving a standpoint by dominantly situated subjects. Although the dominant are not doomed to ignorance about marginalized lives, overcoming the perspectival disadvantages inherent in their social position requires more than cumulatively acquiring knowledge. In situations where marginalized perspectives pose “an ontological threat” to the hegemonic worldview, dominant subjects may need to undergo a radical transformation of their ways

of seeing if they are to achieve a standpoint that does justice to the lives of the marginalized (Mills 2017: 108). In cases of white ignorance, genuinely recognizing marginalized experiences often requires rejecting “whiteness” understood as a particular way of being in the world where white superiority is internalized (Mills 2017: 108). As Mills powerfully writes: “One could not continue to be ‘white’ in this way if one were to accept what is being said, but would have to reshape one’s life and existence in the light of this new understanding” (Mills 2017: 108).

The extent to which marginalized perspectives pose an ontological threat to the dominant ideology may be a matter of empirical debate. It is plausible to expect that, as the perceived threat increases, dominant subjects have less incentive to overcome ignorance about marginalized lives. However, the perceptual transformations required to set that process in motion are not always swift changes. In fact, dominant subjects may gradually become attuned to the world of the marginalized and learn to care for their experiences. In this respect, I agree with Tilton’s consideration that “standpoints come in degrees” and dominantly situated subjects “can make more or less headway toward achieving a standpoint” depending on the amount of work they are willing to put in their route out of “ideological ignorance” (Tilton 2024: 16). As she explains, this work involves more or less radical ways of acknowledging marginalized experience, including “engaging with work on oppression that is written by marginalized people; questioning preconceived or stereotypical notions of marginalized people; talking with marginalized people; and reflecting on the ways that dominant ideologies shape thought” (Tilton 2024: 16).

However, Tilton overlooks how even intellectual work – let alone the work of political praxis – may produce a genuine transformation at the level of perspective from which greater reflection may ensue. In her view, one does “not have to see *like* a marginalized person to have achieved a marginalized standpoint” and “indeed, there is no way to see *like* a marginalized person” (Tilton 2024: 16). The authors working on epistemologies of ignorance lead us in a different direction when they argue that overcoming willful ignorance may require us to challenge our most basic commitments. Where some marginalized perspectives can be additively integrated into the dominant worldview, others will be perceived as “incredible or even

incomprehensible” because they are conceptually incoherent with the dominant conceptual framework (Mills 2017: 108). Hence, overcoming willful ignorance in cases where the perspectives of the marginalized pose a threat to our inherited worldview demands a perceptual transformation that enables us to see aspects of the world that had previously been unavailable to us. At this point, Tilton might argue that the requirement to adopt marginalized perspectives makes it “*too difficult* for the socially dominant to achieve a standpoint” (Tilton 2024: 15). However, difficulty alone might not be a good reason to lower our standards. In fact, the difficulty of the task at hand may be what makes it necessary to uphold them and should not discourage us from honoring them.

Learning to see the world as others do requires a commitment that largely exceeds the theoretical pledge to acquire more and better information about a particular group of people and the material relations that subordinate them. It is a commitment of a practical order. As Pohlhaus explains: “Becoming attuned to when and how to use epistemic resources requires engagement with practitioners skilled in their use, placing oneself in encounters where it makes sense to use them, making mistakes and being corrected” (Pohlhaus 2012: 721). It is a kind of commitment that encourages us to take seriously the experience of others, to imaginatively or even physically put ourselves in their positions. In that sense, we do not simply achieve a standpoint, but effectively take up the perspectives of the marginalized. If this seems to create very high standards for epistemic justice, it is because it does.

These difficulties are all the more pronounced when it comes to dominant subjects who, as we have seen, have few incentives to develop a standpoint that adequately illuminates the forms of social organization that systematically benefit them at the expense of vulnerable others. Becoming attuned to the world of the marginalized is an extremely disturbing endeavor since it may open “one’s eyes to aspects of one’s situatedness with which it is not easy to contend (for example, one’s unearned privilege that cannot simply be disowned)” (Pohlhaus 2012: 721). Furthermore, even when they do set out to understand the world as experienced by marginalized people, they may find that these individuals and communities are unwilling to let them in and share their *arduously honed* epistemic resources (Pohlhaus 2012: 722).

The standards for achieving adequately critical standpoints are high indeed. However, while full epistemic justice may be impossible to achieve, having it as a horizon puts us in the way of fairer epistemic practices. As I have conveyed earlier, I agree with Tilton's claim that critical understandings of social reality come in degrees, which implies that dominantly situated subjects can make more or less headway depending on the amount of work they are willing to put into overcoming situated ignorance (Tilton 2024: 16). Dominant subjects do not have to see everything as marginalized people do, and they certainly do not have to live life as they do, to come upon social reality in a new light. Hence, overcoming situated ignorance about a given dimension of social reality can coexist with the perpetuation of willful ignorance about other dimensions. While this is true, it is also an improvement over a previous situation in which situated ignorance was possibly much more pervasive. It is an improvement not simply because some of the ignorance has been dissipated but because it effectively places dominant subjects in a more advantageous epistemic position that is more likely to expose them to the contradictions that emerge between hegemonic and counter-hegemonic modes of intelligibility.

History is filled with examples of individuals who developed a critical consciousness by imaginatively or physically putting themselves in the position of the least well-off. From Francis of Assisi to Karl Marx and Che Guevara, we find that significant transformations in political worldview are often grounded in profound changes in perspective, derived from serious engagement with marginalized experiences. We also find that even our most critical endeavors have blind spots and may harden into closed conceptual structures that prevent us from seeing any further. Hence, some of our greatest revolutionaries have historically been unable to speak, for instance, for the experience of feminized, racialized, and queer people. They have often held on to these failures of imagination and willfully refused to know *whole parts of the world*. In doing so, they have not only failed the most vulnerable among us but have also actively undermined their revolutionary theory and action.

Willful ignorance is, as Pohlhaus argues, the conceptual opposite of a critical standpoint since "standpoints and willful ignorances rely on the interest one takes in others and their experiences, not one's social position, to determine

what can be known” (Pohlhaus 2012: 731). In that sense, willful ignorance is a form of carelessness. Indeed, both epistemic states depend on the willingness of the subject to attend to others, and in particular, to vulnerable others, in order to move beyond the epistemic limits that characterize her social position. In hierarchically organized societies, where a minority prospers at the expense of the vast majority, virtually everyone begins their epistemic activity in locations that impose severe ideological constraints on their cognitive and political lives, that is, their lives with others. That some of us can challenge such constraints and adopt more critical outlooks on the world speaks to the contingent character of the epistemic limits inherent in the social positions we are born into. We move collectively, critically, and reflectively. But we also move caringly. Understood as a willingness to attend to the experiences of others, to see the world from their perspectives, even when they may profoundly challenge our worldview, care stands as a condition of possibility for knowledge.

References

- Harding, Sandra. (1991). *Whose Science? Whose Knowledge?: Thinking from Women's Lives*. Ithaca, NY: Cornell University Press.
- Mills, Charles Wade. (2007). “White Ignorance.” In *Race and Epistemologies of Ignorance*, edited by Shannon Sullivan and Nancy Tuana, 13–38. Albany, NY: State University of New York Press.
- Mills, Charles Wade. (2017). “Ideology.” In *The Routledge Handbook of Epistemic Injustice*, edited by Ian James Kidd, José Medina, and Gaile Pohlhaus Jr., 100–111. New York, NY: Routledge.
- Tilton, Emily. (2024). “That’s Above My Paygrade: Woke Excuses for Ignorance.” *Philosophers’ Imprint* 24 (8): 1–19.
- Pohlhaus, Gayle. (2012). “Relational Knowing and Epistemic Injustice: Toward a Theory of Willful Hermeneutical Ignorance.” *Hypatia: A Journal of Feminist Philosophy* 27 (4): 715–735.
- Wylie, Alison. (2003). “Why Standpoint Matters.” In *Science and Other Cultures: Issues in Philosophies of Science and Technology*, edited by Robert Figueroa and Sandra Harding. New York, NY and London: Routledge.

When Clocks Stop Working – Chrononormativity and Rhetorical-Epistemic Oppression

Flora Löffelmann (Vienna, Austria)

Abstract

This paper introduces the concept 'Rhetorical-Epistemic Oppression' (REO), a power mechanism I coin in my PhD project. REO describes situations where individuals are pressured, coerced, or manipulated to frame their embodied experiences in accordance with dominant societal norms and expectations. Experiences or feelings that do not conform to these normative expectations can not be articulated. Drawing on Foucault's concept of "productive power" (1976) and insights from social epistemology (Fricker 2007, Dotson 2014), REO provides a framework for understanding how this pressure operates in a way that is both epistemically oppressive, inhibiting the articulation of certain things, and epistemically productive, in that it incentivizes the articulation of experiences that accord to prevailing norms. In this paper, I focus on how REO enforces chrononormativity, the idea that life has to happen according to a specific normative "life schedule" (Halberstam 2005). I examine its effects by the example of the lived experiences of trans and queer individuals. I argue that REO obscures key "fleeting moments" (Amin 2014) that are vital to queer and trans lives. Finally, I discuss how the concepts of queer temporality (Halberstam 2005, Freeman 2006) and collective continuance (Whyte 2020) offer potential remedies to mitigate the harmful effects of REO.

1. An Introduction to Rhetorical-Epistemic Oppression

Imagine the following situation: A transgender man, let's call him Joe, wants to undergo chest surgery because having breasts does not conform with his wishes of embodiment. To do so, in Austria, he requires three psychologists' statements confirming that his wish is genuine, and should be granted. As a mechanism of this "medical gatekeeping" (Pearce 2018), he is quizzed about his childhood, sexuality, and embodiedness (Maier 2024). Yet, Joe cannot openly talk about his experiences, feeling compelled to conform to a predetermined narrative that emphasizes notions of binary gender and the stereotypes attached to them. He might highlight always having preferred stereotypically 'male' activities, or always having wished to dress like a man. Not wanting to fail the examination, he feels obliged to hide parts of himself that could be considered 'too feminine'.

My PhD project examines situations where individuals are pressured into making statements about their embodied experience in accordance with socially dominant norms. Drawing on Michel Foucault, I argue that "forced or provoked confessions" (1978) function as power mechanisms that produce

what is then taken to be knowledge. I propose that the power mechanism I call 'rhetorical-epistemic oppression' (REO) constitutes a form of "epistemic oppression" (Dotson 2014) fueled by productive power. My contribution to the literature on epistemic injustice (Fricker 2007) and oppression is to highlight that productive mechanisms of power entail forms of oppression that lead to individuals only being believed if they conform to specific modes of self-presentation. In such cases, individuals are compelled to produce 'knowledge' that ultimately reinforces their oppression. Building on a interview based case study of trans persons in Austria, I explore the material and phenomenal effects of REO on those compelled to confess.

The novelty I introduce is that the power mechanism REO is both epistemically productive and oppressive. REO works productively in compelling subjects to share information that reinforces their own marginalization, p.e. by reproducing stereotypical notions about their group. Trans people may feel pressured to testify only stories related to their gender identity or body that align with dominant cis-normative narratives. REO is oppressive as knowledges that do not confirm to these norms are not articulated, and thus do not enter collective knowledge.

I distinguish three harms of REO, namely epistemic oppression, epistemic domination, and affective injustice. As epistemic oppression (Dotson 2014), REO restricts marginalized subjects' ability to articulate a full range of truths about themselves, thereby infringing upon their epistemic agency. Joe might not tell the psychologist that he enjoys his vulva and vagina, since this does not conform to cis-normative male embodiment. Instead, Joe is forced to use dominant hermeneutical resources, such as male stereotypes, to make himself intelligible as a man. I frame this as epistemic domination (Alatas 2000; Dular 2023). Phenomenally, epistemic oppression and epistemic domination impact individuals' embodied experiences, causing, as I claim, affective injustice (Srinivasan 2018). Joe might feel ashamed for having to hide a part of himself from the psychologist, and experience other negative affects, such as anger at being subjected to medical gatekeeping.

In this paper, I examine how REO imposes chrononormativity on the experiences of transgender individuals who wish to change their bodies in a way that aligns with their gender identity. I argue that REO works to render

certain “fleeting moments” (Amin 2014) which are formative for queer and trans experiences, and experiences of non-linear “queer temporality” (Halberstam 2005), inarticulable in situations permeated by a strong power differential. Following this, I propose that the concept of “collective continuance” (Whyte 2020) could mitigate some of the negative effects of REO if it entered into collective hermeneutical resources.

2. Chrononormativity and trans experience

In this paper, I focus on the temporal aspect of trans persons’ narratives with the psychologists. I claim that they are forced to present their gender identity in accordance with chrononormative ideals of identity. “Chrononormativity” was first coined by Elizabeth Freeman (2010) as “a social patterning of experiences of time in conformity with normative frameworks” (Amin 2014: 220). Chrononormativity describes how temporality is valued within the Western context. It was established through settler colonialism, which imposed its norms of space, time, and gender through violence and elimination. Importantly, settler colonialism is not of the past, but a continuing force. Its effects are still present, reflected in settler laws, policies, and practices, as well as the ongoing elimination of Indigenous peoples and their cultures (Wolfe 2006: 387).

Trans studies highlight the often non-chrononormative nature of trans experiences. For instance, the term “second puberty” (Hayward 2017: 268) is commonly used to describe how trans individuals feel once they begin their transition and start living as their ‘real’ gender. This experience frequently disrupts the normative expectations that Jack Halberstam refers to as a “life schedule” (2005). Halberstam argues that, in heteropatriarchal societies, certain life events, such as first love, graduation, stable relationship, marriage, stable employment, and having children, are expected to occur at specific ages. They are supposed to follow a linear, teleological progression. Being trans or queer often upends this normative schedule. I propose that chrononormativity contributes to the REO imposed on trans individuals during psychological assessments, and in everyday life. The epistemically productive dimension of REO forces people to only share experiences that conform to normative ideas of linearity and timing, and to exclude, by virtue of REO’s epistemically oppressive dimension, experiences that do not align with normative temporal expectations.

Many cis persons assume that trans persons view gender transitions as a linear process with a defined endpoint, which is to be undetectably trans, thus passing as a cis person. The underlying premise is that, in transitioning, one stable gender identity, the one held prior to transition, is replaced by another stable gender identity, the one assumed post-transition. Transition is thus conceptualized as a linear progression from one fixed gender identity to another (Bex-Priestley 2022). This perception reinforces the underlying assumption that gender is a fixed, unchanging characteristic.

Public discourse surrounding transition routinely frames individuals' efforts to transition as occurring at the “wrong time” (Goetzke 2024). As children, individuals are not trusted to be able to recognize their trans identity, while as adults, they are expected to have understood their identity earlier. This creates an artificial dichotomy, rooted in transphobic assumptions, which positions trans individuals as “innocents to be protected or as dangers to be feared” (Faye 2021: 21). Thus, transitions are consistently perceived as never having the 'right' timing.

Questions trans persons are asked in psychological evaluations aim to judge whether someone is a “good transsexual” (Preciado 2021). Often, they target the gendered nature of toys used in childhood (Maier 2024: 96), or other childhood experiences (Spade 2006: 319). In situations permeated by REO, statements necessary to be labelled a “good transsexual” are intended to affirm that transness is a fixed state assumed early on, rather than the dynamic, unfolding process that many trans individuals describe (Maier 2024). Ideally, trans persons can share a convincing story of origin that implies that they have early on felt trans. REO's productive dimension is thus that people articulate narratives that emphasize the “always having known”, sometimes even fabricating memories that never really happened, while accounts of non-linear developments of gender identity are withheld (Maier 2024).

Jules Gleeson (2021) highlights that trans communities perform the socially reproductive work of reciprocal recognition and information-sharing, providing formative relationships and bonds that help to shape transitioning subjectivities. In these communities, people frequently share that they did not experience that singular moment of knowing that their gender assigned at birth did not conform to their actual identity, but that they rather slowly

acquired the right words to name their experience while getting to know other trans individuals (Gleeson 2021). Socially reproductive work is particularly important in a context where transness is continually pathologized and cisgender identity is treated as the normative standard. Despite this, psychologists rarely positively acknowledge trans communities, the reproductive work they do, and the collective efforts and time invested in these supportive networks (Maier 2024).

Literature shows that, during psychological evaluations, different concepts of 'time' are assigned varying degrees of significance. For example, the hours spent researching how to present oneself to the psychologist, and the anxieties associated with 'getting it right', are frequently excluded from the official timeframes used to assess trans experiences (Dickinson 2021). However, as many trans individuals' accounts demonstrate, the process of researching how to present as a 'proper' trans subject constitutes a significant aspect of their experience.

Psychologists' chrononormative expectations delegitimize the diverse and multifaceted experiences of trans subjects. During assessments, REO thus oppresses accounts of gender identity formation that expand on its social nature, rather productively emphasizing its individualistic aspects. As a consequence, the socially reproductive work that members of trans communities perform for one another is not acknowledged. This perniciously renders these epistemic communities, and the important work they do, invisible.

3. Clocks that Stop Working: The Significance of Fleeting Moments

Why is there so little variation in how trajectories of trans individuals are discussed, and why are concepts of linearity and 'right' timing dominant? Much of contemporary transphobic accounts claim that 'being trans' is something new. One reason is that records of trans people's existence have been violently erased, and their histories suppressed. For instance, one of the first actions of the Nazi regime was to destroy, in 1933, the German Hirschfeld Archives, where Magnus Hirschfeld had amassed over 10,000 volumes documenting trans life. Hitler reportedly considered Hirschfeld the "most dangerous man in Germany" due to his extensive evidence that transness is far less exceptional than the right wing claims (Feinberg 1996).

However, new accounts of trans life are shaped by societal preconceptions about what it means to be trans. They are a result of early sexology's attempts to manage and pathologize trans persons (Gill-Peterson 2018). In psychological settings, narratives aligning with the diagnostic criteria of Gender Identity Disorder are required in order to access hormonal or surgical treatments (Prosser 1998: 104). Autobiographical narratives expected in such contexts “retrospectively bestow an illusion of teleological progression upon the aleatory chaos of life experience” (Amin 2014: 219). Given the linear conception of biography, many aspects of a person's experience are simplified or omitted when recounting their trajectory toward a 'completed' trans identity. As Kadji Amin argues, this process occurs “at the expense of episodes, or even fleeting moments, that would fracture or exceed it” (Amin 2014: 220). Many elements of queer narratives are omitted from the conventional narratological frameworks that trans individuals are expected to present, both in psychological settings and in everyday life.

In *Queer Phenomenology* (2006), Sara Ahmed makes similar observations regarding lesbian life: “It is given that the straight world is already in place and that queer moments, where things come out of line, are fleeting. Our response need not be to search for permanence [...] but to listen to the sound of ‘the what’ that fleets” (Ahmed 2006: 106). Ahmed's call is to make these moments intelligible by recording and discussing them, and by recognizing their value, even if they do not conform to normative frameworks. Yet, confrontations with REO might lead to the non-articulation of these highly formative fleeting moments.

Halberstam's concept of “queer temporality” (2005) offers valuable insights into queer and trans persons' experiences, and helps me to explore further negative effects of REO. Halberstam highlights that “time-warping” experiences are frequent in queer life, noting that being “out of sync” with normative life schedules is a common experience (2005). Drawing on Ahmed's queer phenomenology, I will now examine phenomenological implications of queer temporality.

Ahmed argues that queer subjects often occupy “a queer slant”, a diagonal life trajectory that intersects the conventional vertical and horizontal lines of genealogical norms. It is the “everyday work of dealing with the perceptions of

others, with the 'straightening devices' and the violence that might follow when such perceptions congeal into social forms" (Ahmed 2006: 107) that shapes how the chrononormativity of cis-heteronormativity impacts queer subjects. Confrontation with transphobia or homophobia not only shapes individuals' lives, but also their orientation in the world. It impacts time through the processes that queer subjects "have to work through." This includes evading violence through code-switching, processing micro-aggressions, ignoring the scrutinizing glances directed at one's body when presenting in ways that defy normative expectations (e.g., wearing a dress while being perceived as male), and addressing the trauma resulting from non-conformity. Trans individuals' testimonies reveal how transgressions of gender norms structure the way reality is experienced, shaping one's existence in time. These experiences also shape the decisions trans persons take during transition. Yet, these experiences do not fit within the chrononormative frameworks of identity that are linked to normative ideas of straightness and linear progress. By virtue of REO's oppressive aspect, they are often not thematized in psychological assessments, while teleological accounts are reproduced.

Amin highlights that omitting these "fleeting moments" of development, or struggles with a queer- and transphobic world, when presenting one's narrative can lead to fragmentation and invalidation. Some individuals may leave a psychologist's office or a conversation with a cisgender inquirer emotionally scarred by having to present a narrative that excludes aspects of their experience that "do not fit" within normative frameworks. If the self is, as Susan Brison observes, "an embodied, socially constructed narrative" (2017: 226), the suppression of formative incidents and 'fleeting moments' during psychological assessments can harm the self as a whole. Attending to the fleeting, exceeding moments often omitted from trans individuals' narratives can remedy some of the harms of REO. By rendering fleeting moments inarticulable, and producing grand, linear narratives, REO flattens the individuality of trans experience, undermining individuals' epistemic agencies and potentially disrupting people's relationship to their gender and conception of self.

4. Collective Continuance: A possible remedy for REO?

“Collective continuance” (CC), a term from Anishinaabe philosophies, proposes another non-normative conception of time. Could REO’s harms be remedied by introducing CC into the hermeneutical apparatus that psychologists use to evaluate trans persons? Could it even make them regret their role in perpetuating oppression through gatekeeping, and in upholding a settler colonial, cis normative epistemic system?

As Kyle Whyte shows, “collective continuance considers existence as emanating from the relationships between humans and nonhumans that are in constant motion, embracing diversity and constituted by reciprocal responsibilities” (Whyte 2020: 53). Existence is considered to be fluid, systematically organized, and in a continuous state of action and motion. Images to conceptualize time include strands in a spider’s web, snowflakes, and shoals (Whyte 2020). Anishinaabe philosophies understand identity as always shifting due to its intrinsic connection to the place, time of year, and the network of reciprocal responsibilities within which a person finds themselves. This framework demands the valuing of ‘fleeting moments’ and non-normative instances, rather than policing them. Groups that live in CC view “migration, motion, fluidity, vibration, and expansion/contraction as experiences and temporalities that should be respected and celebrated, not avoided or denied” (Whyte 2020: 57). CC is frequently employed to critique the mechanisms of settler colonialism, which not only introduced new orders of time and space, but also redefined gender through physical genocide and a logic of elimination. Whyte highlights that settler colonialism seeks to erase CC, diminish human and non-human diversity, and impose sexist and racist norms (Whyte 2020). REO’s productive and oppressive power of making trans individuals align their narratives with the demands of chrononormativity can be understood as an effect of the violent imposition of these norms.

An appreciation of CC on part of psychological evaluators could mitigate some effects of REO in regard to how people narrate their life journey. As a productive mechanism that entails the harm of epistemic domination, REO incentivizes linear, grand narrations of gender identity development that have a clear point of origin. Further, it promotes individualistic accounts. As an epistemically oppressive mechanism, REO inhibits the articulation of the socially reproductive work trans communities perform, and of the fleeting

moments that inform people's orientation in the world and the decisions they take in their transition journey. The focus of CC on constant motion, and its appreciation of non-linear and non-fixed identities, offers a necessary corrective to the concept of a fixed identity that settler colonial chrononormative frameworks propose and enforce through REO. Further, in emphasizing the social aspect of every experience, CC anchors a demand for reciprocal responsibility to the realization that all humans are socially intertwined. Ideally, this would lead to empathy on part of the evaluators.

How could an incorporation of CC into psychologists' hermeneutical resources aid Joe, who must demonstrate that he is 'trans enough' to qualify for gender-affirming care? When REO enforces chrononormativity, Joe likely experiences a loss of epistemic agency in regard to his recollection of his life. He knows that providing a linear progress from a clear point of origin will ease his chances of getting approval for the surgery, and that he should not draw too much attention to the time that he spent with his community in figuring out his transness, or the times he was unsure about what exactly he wanted. An important part of Joe's self-understanding and life is in danger of not only being obscured, but even undone in light of the loss of agency he is experiencing. The structures he encounters on his path to his desired embodiment are not interested in his real experiences, which induces emotional distress.

Paying more attention to queer temporalities, and a commitment to collective continuance on part of the psychologists, could remedy epistemic oppression, domination, and affective injustice, and could lead to Joe receiving the gender affirming care he desires without having to distort the narrative that makes up his life. After all, the relation to the psychologist is another relation that co-constitutes Joe's self. As a person partaking in collective continuance, the psychologist should enable him to embrace his transition with all the slantwise and fleeting moments that come with it.

References

- Ahmed, Sara (2006) *Queer Phenomenology*, Durham and London: Duke University Press.
- Amin, Kadji (2014) "Temporality", *TSQ: Transgender Studies Quarterly* 1(1–2), 219–222.

Bex-Priestley, Graham (2022) "Gender as Name", *Journal of Ethics and Social Philosophy* 23(2), 189-213.

Brison, Susan J. (2017) "Personal Identity and Relational Selves", in: Ann Garry, Serene J. Khader and Alison Stone (eds.) *The Routledge Companion to Feminist Philosophy* New York: Routledge, 218-230.

Dickinson, Nathaniel (2021) "Seizing the Means: Towards a Trans Epistemology", in: Jules Joanne Gleeson and Elle O'Rourke (2021) *Transgender Marxism*, London: Pluto Press, 204-218.

Feinberg, Leslie (1996) *Transgender Warriors. Making History from Joan of Arc to Dennis Rodman*, Boston: Beacon Press.

Foucault, Michel (1976) *Überwachen und Strafen. Die Geburt des Gefängnisses*, Frankfurt am Main: Suhrkamp.

Freeman, Elizabeth (2006) *Time Binds. Queer Temporalities, Queer Histories*, Durham and London: Duke University Press.

Gill-Petersen, Jules (2018) *Histories of the Transgender Child*, Minneapolis: University of Minnesota Press.

Gleeson, Jules (2021) "How Do Gender Transitions Happen?", in: Jules Joanne Gleeson and Elle O'Rourke (2021) *Transgender Marxism*, London: Pluto Press, 70-84.

Goetzke, Louka Maju (2025) "Transitioning Out of Time: Why Every Gender Transition is Always Already Late" *Time & Society* 0(0), 1-22.

Halberstam, Jack (2005) *In a Queer Time and Place. Transgender Bodies, Subcultural Lives*, New York and London: New York University Press.

Hayward, Eva (2017) "Spiderwomen", in: Reina Gossett, Erica A. Stanley, and Johanna Burton (2017) *Trap Door. Trans Cultural Production and the Politics of Visibility*, Cambridge: The MIT Press, 255-280.

Maier, Ra (2024) "'Sind Sie sicher, dass Sie nicht einfach nur schwul sind?' Die Erfahrungen von trans Personen im Rahmen der Begutachtungspraxis der staatlich-medizinischen Regulation von Trans in Österreich" MA thesis, University of Vienna, online: <https://theses.univie.ac.at/detail/70414#>.

Pearce, Ruth (2018) *Understanding Trans Health: discourse, power, and possibility*. Bristol: Polity Press.

Preciado, Paul B. (2021) *Can the Monster Speak?*, London: Fitzcarraldo Editions.

Prosser, Jay (1998) *Second Skins: The Body Narratives of Transsexuality*, New York: Columbia University Press.

Srinivasan, Amia (2018) "The Aptness of Anger." *The Journal of Political Philosophy* 26(2), 123-144.

Spade, Dean (2006) "Mutilating Gender", in: Susan Stryker and Stephen Whittle, *The Transgender Studies Reader*, New York: Routledge, 315-332.

Stockton, Kathryn (2009) *The Queer Child, or Growing Sideways in the Twentieth Century*, Durham and London: Duke University Press.

Whyte, Kyle (2020) "Collective Continuance", in: Gail Waiss, Ann V. Murphy and Gayle Salamon (eds.) *50 Concepts For a Critical Phenomenology*, Evanston: Northwestern University Press, 53-60.

Wolfe, Patrick (2006) "Settler Colonialism and the Elimination of the Native" *Journal of Genocide Research* 8(4), 387-409.

Acknowledgement

I want to thank the participants of the Workshop "Critical Temporalities", hosted by the New School of Social Research and the University of Vienna in March 2023, and especially its organizer Leyla Sophie Gleissner, for the incentive to write this paper. Further, I want to thank the participants of the Trans and Queer Theory Colloquium, Department of English and Comparative Literature, and of the Graduate Colloquium, Institute for the Study of Sexuality and Gender, both Columbia University, for their feedback. Lastly, I owe gratitude for their insightful comments to my readers Ding, Dr. Val Meneau, and Dr. Sophie Veigl.

Inhabiting the Dilemma: Wittgenstein, Feminism, and the Question of the Subject

Saori Makino (Chiba, Japan)

Abstract

This paper explores how Ludwig Wittgenstein's later philosophy can be read as a form of feminist practice, particularly in relation to a dilemma surrounding subjectivity. Drawing on Susan Hekman's analysis of the Cartesian subject, I revisit the tension between the constructed nature of subjectivity and the imperative to assert feminist agency. Rather than proposing a new theoretical resolution, I rearticulate this philosophical dilemma as an everyday one that emerges within the ambiguity of lived experience. Through a close reading of the *Last Writings on the Philosophy of Psychology, Volume II*, I focus on Wittgenstein's interest in the dichotomy of "the inner and the outer", and on how this distinction shapes our understanding of other minds. Wittgenstein does not dissolve this dichotomy; he lives with the difficulty it poses, tracing its influence within our forms of life. This practice of staying with a philosophical difficulty offers a model for feminist self-clarification – an approach that neither resolves nor escapes contradiction, but inhabits it. Ultimately, I suggest that reading Wittgenstein in this way reveals a feminist practice, rooted in attentiveness and the struggle to speak meaningfully from within an embodied life.

1. The Dilemma of the Feminist Subject

Around 1990, postmodern approaches became prominent in feminist epistemology. A postmodern approach can be broadly defined as an intellectual attitude that challenges such concepts as universality and necessity, questioning the assumption that they exist independently of any particular culture, history, or perspective. Susan Hekman, a leading postmodern feminist, articulates the relationship between feminism (implicitly, feminist epistemology) and postmodernism as follows:

If all the "feminisms" have anything in common it is a challenge to the masculine/ feminine dichotomy as it is defined in western thought. It is this dichotomy that informs all the dichotomies the postmoderns are attacking, even though they do not always make this explicit. Similarly, challenging the priority of "man" in the modern episteme must be fundamental to any feminist program. On these key issues, as well as on many others, feminism has much to gain from an alliance with postmodernism. (Hekman 1990: 2)

Hekman does not merely argue that feminism shares common ground with postmodernism and should "ally" (ibid.: 2) itself with it. She suggests that

feminism may offer a more radical challenge to modern epistemology than postmodernism does. Postmodernism attacks “the homocentricity of Enlightenment knowledge and even the status of ‘man’ himself” (ibid.: 2), which is based on dichotomies such as “rational/irrational” and “subject/object”. Despite this, she argues, postmodernism has not been as explicit as feminism in addressing the “masculine/feminine” dichotomy that lies at the very foundation of such humanism. Feminists have been the ones persistently challenging this dichotomy.

If Hekman is correct, feminist epistemology is capable of carrying out the postmodern project more radically than postmodernism alone, and should energetically do so in order to advance feminism. In fact, she frames the Cartesian subject/object dichotomy as a central issue in feminist epistemology. In *Gender and Knowledge: Elements of a Postmodern Feminism* (1990), she devotes the third of five chapters to this issue.

The rationalism at the root of the epistemology of modernity is a rationalism of the subject. As many critics of modernity have noted, the modern episteme is defined by the Cartesian dichotomy between subject and object. The driving force of the modern age is the search for certainty, the effort to use reason to establish absolute and universal truth. Since Descartes that search for certainty has been firmly grounded in the rationality of the knowing subject. Descartes’ *ego cogito ergo sum* placed the certainty that is the goal of the modern episteme firmly within man himself. (Ibid.: 62)

According to Hekman, the modern Cartesian subject, founded on such dualisms as reason/irrationality and subject/object, and on privileging one aspect of each dichotomy over the other, is not only an adversary of postmodernism but also a challenge to feminism. She states that feminism can address the Cartesian subject in three ways: “retaining”, “restructuring”, and “rejecting” (cf. ibid.: 65). The first approach, “retaining”, seeks to preserve humanism while challenging its inherent sexism. The second, “restructuring”, involves reconceptualizing the “subject” by incorporating an active role and position for “women”. The third approach, “rejecting”, entails completely abandoning any attempt to reconstruct the Cartesian subject or anything based on it (cf. ibid.: 65).

Hekman herself appears to support the second approach; she presents a favourable characterization of Foucault's challenge to the Cartesian subject. According to Hekman, Foucault portrays the "constituted subject that is wholly determined by social forces" (ibid.: 72) as the "subject that resists" (ibid.: 73). Foucault's concept of the subject brings to light how feminist discourse "is born out of resistance to the modes of discourse that, historically, have constituted the feminine subject", without relying on "a transcendental, constituting subject or an essentially female" one (ibid.: 73). Hekman then demonstrates that the ideas of Irigaray, Cixous, and Kristeva can be assessed in a similar manner. She does not advocate the complete rejection of the Cartesian subject, but supports using it as a means to conceptualize the feminist subject.

My aim in this paper is not to devise a new concept to replace the "modern subject", nor to evaluate the theoretical solutions that other philosophers and feminists have offered in response to critiques of it. In particular, Hekman has critically examined the Cartesian subject not to call for its outright rejection, but to explore how it might be reinterpreted through Foucault's notion of the "resisting subject", and thus be reappropriated for feminist purposes. Drawing on this framework, Hekman shows how feminist discourse has emerged not from a transcendental or essentially feminine subject, but from resistance to the discursive formations that have historically constituted women as subjects. This offers a way to reconceptualize feminist agency while remaining attentive to the historically situated nature of subjectivity.

While Hekman does not explicitly present her argument in terms of a dilemma, my reading of her work reveals a tension between two imperatives: to acknowledge the socially and historically constructed nature of subjectivity, and to affirm the agency and authority of those who speak as women. Hekman seeks to resolve this tension by theorizing a feminist subject modelled on Foucault's "resisting subject" – a subject shaped by power, yet capable of resistance. In contrast, my aim is not to resolve this tension, but to remain with it – to reflect on how such tensions are lived, felt, and negotiated in everyday life. The difficulty of determining whether our actions stem from genuine desires or from internalized norms, and whether our voices are truly ours or shaped by dominant scripts, is not merely theoretical. It is a condition of our ordinary experience.

My approach, therefore, is to take what I perceive as Hekman's postmodern dilemma and to rearticulate it as an everyday dilemma: one that manifests not in abstract philosophical debate but in the struggles of navigating social life. I aim to shift feminist epistemology from theoretical mastery to a practice of listening attentively to the complexities of situated voices. This resonates deeply with the later philosophy of Ludwig Wittgenstein, who treats philosophical problems not as puzzles to be solved but as difficulties in our way of living and going on. His work offers a way to remain with the tensions that arise within our forms of life – and to seek clarity, not through resolution, but through self-clarification.

2. Staying with the Dichotomy: Reading “The Inner and the Outer”

This section examines how Wittgenstein engages with the distinction between the “inner” and “outer”, not as an abstract dichotomy but as a picture, embedded in our form of life, that subtly distorts how we understand others and respond to them. This picture is not a theory imposed from outside, but a way of seeing in which we participate, often without recognizing how it influences us.

Instead of stepping outside the tensions this picture generates, Wittgenstein investigates the forms they take in our everyday language. He does not attempt to resolve these tensions by framing them as universally intelligible philosophical problems. Instead, his work models a form of critical attentiveness – an effort to remain responsive to the gestures, responses, and ways of behaving through which others appear to us. This evokes feminist concerns about how we listen, how we respond, and how we live with the conceptual frameworks that are deeply ingrained in our ways of seeing.

This section focuses on the *Last Writings on the Philosophy of Psychology, Volume II* (LW 1992, abbreviated in running text as *LW2*), one of Wittgenstein's posthumously published manuscripts. It consists of a series of notes (MS 169–176, excluding 172 and 175) from the final two years of his life (1949–1951). According to the editors, Georg Henrik von Wright and Heikki Nyman, these notes fall into three thematic groups: “the concepts of certainty, knowing, doubting, and other topics in epistemology”; “the philosophy of colour concepts”; and notes dealing with “psychological concepts and in particular with the problem of the relationship between ‘the inner’ and ‘the outer’” (ibid.:

x). Reflecting its emphasis on the third group, the volume bears the subtitle *The Inner and the Outer*.

Since *LW2* is a collection of manuscript notes, views on how it should be treated may vary. In my view, however, it is one of the culminating points of Wittgenstein's later philosophy. There are two reasons for this. First, *LW2* deepens the reflections Wittgenstein had already developed in a series of typescripts and manuscripts written between 1946 and 1949, including RPP 1980a, RPP 1980b, and *LW* 1982. Second, as the editors note, "the excellent quality of these writings [MS 173, 174, and 176] should be obvious to every reader" (cf. *ibid.*: xi).

In these typescripts and manuscripts, Wittgenstein repeatedly addresses the problem of other minds and the uncertainty to which it exposes us, either as knowers or as interlocutors. In *LW2*, this theme appears as a persistent philosophical scepticism, generated by the dichotomy between the "inner" and the "outer". If we adopt a therapeutic reading of *LW2*, we might describe its narrative as follows: the inner/outer dichotomy is a "picture" (*LW2*: 69; cf. *PI*: §115) – a way of seeing – that fixes our perspective. While the uncertainty surrounding other minds may indeed arise from everyday experience, it becomes pathological when this dichotomy intensifies and reinforces it.

Logical Impossibility lies in the lack of exact rules of evidence. (Therefore we sometimes express ourselves in this way: "We may always be wrong; we can never be certain; what we observe can still be pretence." Although pretence is only one of many possible causes of a false judgment.) (*LW2*: 94)

If Wittgenstein had treated the uncertainty surrounding other minds as a philosophical problem that could be formulated and resolved independently of any particular perspective, then the claim "Logical impossibility lies in the lack of exact rules of evidence" would amount to a surrender. For if the impossibility of knowing "what goes on in someone else" (*LW2*: 94) could be explained by "the lack of exact rules of evidence", then it would be permanently irresolvable.

I reject this interpretation, and instead wish to attend to the tone of *LW2* as a whole. What then becomes apparent is Wittgenstein's entanglement in

uncertainty about other minds. One should hesitate to attribute this merely to his personal disposition – his scepticism or his tendency toward isolation. More importantly, his tendency to become entangled in this uncertainty cannot be dismissed as a sign of irrationality. Rather, what the tone of *LW2* suggests is that we live, and engage with others, through language – in its broadest sense, including facial expressions and behaviour. Wittgenstein’s entanglement in the uncertainty of other minds is not the result of a lack of rational reflection. This entanglement arises because we live within language, and thus others’ expressions, behaviour, and displays of pain are not external to or separate from ourselves.

Let us return to the quotation about rules of evidence, and look at the passage following *therefore*. Here, Wittgenstein treats what may appear philosophically threatening – “the lack of exact rules of evidence” – as a condition we actually accept and live with. The sentence “We may always be wrong; we can never be certain; what we observe can still be pretence” is not something only a pathological sceptic would say. Even if we accept that “pretence is only one of many possible causes of a false judgment”, the former sentence nonetheless articulates a familiar aspect of how we relate to others.

Nothing more powerfully expresses Wittgenstein’s stance than his remark that “the ‘inner’ is a delusion... like a painted curtain drawn in front of the scene of the actual word use” (*LW2*: 84). The assertion that “the inner is a delusion” strikes us with considerable force, coming across as a sharp, even dismissive judgment. However, the word *delusion* does not necessarily imply that the concept of the “inner” must be discarded as a philosophical confusion. What this passage suggests, when read alongside the curtain metaphor that immediately follows it, may not be the rejection of a particular concept, but the presence of a kind of fate – one in which we find ourselves profoundly separated from the place to which we are meant to return. Wittgenstein’s metaphor gestures gently at the structure that stands between us and the everyday – an image that invites us to look at it, not to tear it down. What it captures is not a path back to untroubled immediacy, but a lived sense of distance: a horizon in our shared life that both orients and eludes us.

3. Living the Dilemma: Feminist Everydayness Revisited

One of the central insights that has emerged from Wittgenstein's later writings, particularly his *Last Writings on the Philosophy of Psychology II*, is that the uncertainty surrounding other minds is not an abstract puzzle to be solved, but a lived condition – an expression of what it means to live in language. The dichotomy between the “inner” and the “outer”, far from providing clarity, is a misleading picture that separates us from what we already inhabit: our life with others. Wittgenstein does not aim to dismantle this picture, but to remain with it – to reveal its gravitational pull, even as we continue to live through gestures, expressions, and shared forms of life. As I have argued, this expresses an understated yet enduring sense of everydayness: we are already in the world we seek to return to, yet we are estranged from it by the frameworks we have inherited.

We can now revisit the dilemma within Hekman's thought. Her project was to challenge the Cartesian subject with a feminist conception of subjectivity, showing how the subject is both socially constructed and nonetheless capable of agency. By employing Foucault's notion of the resisting subject, she sought to resolve the tension between passivity and authorship within feminist epistemology. But what appears as a philosophical dilemma in Hekman's account – a tension between historical construction and self-determination – can be seen from another angle as an everyday dilemma: one that arises not in the realm of theory, but in lived experience. For example, in a heteronormative society, can I be certain that my love for a man is my own? In a culture that assigns care work to women, can I truly care for another without compulsion? These are not questions we ask abstractly, but ones that we live through – often ambivalently, and often without any resolution.

To reinterpret a philosophical dilemma as an everyday one is not to reduce it to the personal, but to relocate it: to see how agency is enacted even under conditions of constraint, and how one's voice emerges, not despite lived ambiguity, but through it. Wittgenstein's way of staying with difficulty – of illuminating how a problem reveals itself, rather than how it might be resolved – provides a way to hear these voices. It offers an image of philosophy not as theoretical mastery but as a practice of attention—open to anyone who seeks to listen, especially to what struggles to be said in contexts shaped by silence and marginalization.

In this way, reading Wittgenstein becomes a feminist practice – not because it provides a new theory of subjectivity, but because it allows us to remain with the questions, the tensions, and the voices that constitute our shared and fractured forms of life.

4. Conclusion

This paper has revisited a central tension in feminist epistemology: the dilemma between the deconstruction of the modern subject and the need to speak from the position of “women”. Drawing on Hekman’s account, I have suggested that this philosophical dilemma can be reinterpreted as an everyday one: not a contradiction to be resolved, but a condition to be lived with. In Wittgenstein’s attentiveness to how language entangles us, I have found an attitude that matches this need to stay with our difficulties.

Wittgenstein’s texts call philosophical problems – once they have emerged as abstract discursive formations – back into the dimension of life that continues within language. In his writing, we find ourselves oscillating between two tendencies: on the one hand, resignation to the inevitable condition that there is no outside to language; and on the other, a contrasting impulse – the desire to strip away the pictures that obscure our ability to attend to words as they are used in specific, lived contexts. This oscillation between resignation and desire is itself a manifestation of the everyday dilemma.

Rather than proposing a new theory of subjectivity, I have approached feminist philosophy as a practice of reading – of listening to voices shaped by hesitation, ambiguity, and situated struggles. In this way, reading Wittgenstein has become not only an interpretive effort, but a feminist one.

References

- Hekman, Susan (1990) *Gender and Knowledge: Elements of a Postmodern Feminism*, Cambridge: Polity Press.
- Wittgenstein, Ludwig (2009) *Philosophical Investigations*, 4th edition, trans. G.E.M. Anscombe, P.M.S. Hacker, and Joachim Schulte, eds. P.M.S. Hacker and Joachim Schulte, Oxford: Wiley-Blackwell.
- Wittgenstein, Ludwig (1980a) *Remarks on the Philosophy of Psychology, Volume I*, eds. G.E.M. Anscombe and G.H. von Wright, trans. G.E.M. Anscombe, Oxford: Blackwell.

Wittgenstein, Ludwig (1980b) *Remarks on the Philosophy of Psychology, Volume II*, eds. G.E.M. Anscombe and G.H. von Wright, trans. G.E.M. Anscombe, Oxford: Blackwell.

Wittgenstein, Ludwig (1982) *Last Writings on the Philosophy of Psychology, Volume 1: Preliminary Studies for Part II of Philosophical Investigations*, eds. G.H. von Wright and Heikki Nyman, trans. C.G. Luckhardt and Maximilian A.E. Aue, Oxford: Basil Blackwell.

Wittgenstein, Ludwig (1992) *Last Writings on the Philosophy of Psychology, Volume 2: The Inner and the Outer (1949–1951)*, eds. G.H. von Wright and Heikki Nyman, trans. C.G. Luckhardt and Maximilian A.E. Aue, Oxford: Basil Blackwell.

Deep Disagreements: The Importance of Examples

Ryan Manhire (Åbo, Finland)

Abstract

The notion of deep disagreement has received increasing attention in recent years following Robert J. Fogelin's seminal text 'The Logic of Deep Disagreements'. Fogelin draws on Ludwig Wittgenstein's *On Certainty* to describe forms of disagreement that appear "immune to appeals to facts", and thus "not subject to rational resolution", and that persist despite "a general agreement on moral issues", such as the sanctity of human life. Fogelin's suggestion, that the resolution of deep disagreements may require us to "fall back on persuasion", has been criticised by Duncan Pritchard and others who maintain that resolution must remain within the confines of a broadly rational framework.

In this paper, I suggest that one of the limitations in the disagreement between Fogelin and Pritchard concerns the absence of examples of language use in the midst of a potential deep disagreement. In line with Wittgenstein's observation that "[o]nly in the stream of thought and life do words have meaning", I consider a heated exchange between a couple in Kendrick Lamar and Taylour Paige's 'We Cry Together' that could be read as a case of deep disagreement along the lines suggested by Fogelin. I argue that the interpersonal dynamics present in the example challenge Pritchard's idea of the *über hinge* commitment as a pathway towards the rational resolution of deep disagreements. Rather, the example may highlight the depth of the difficulty such disagreements present us with, as well as the need to reconsider Fogelin's suggestion of non-rational persuasion as a pathway towards resolution.

I. Fogelin on Deep Disagreement: Hinge Propositions and Persuasion

In his seminal text 'The Logic of Deep Disagreement', Robert Fogelin describes forms of disagreement that "cannot be resolved through the use of argument, for they undercut the conditions essential to arguing" (Fogelin 1985: 5). Such "deep disagreements" are characterised as both "immune to appeals to facts" (Fogelin 1985: 5), and thus "not subject to rational resolution" (Fogelin 1985: 7), and as persistent despite "a general agreement on moral issues" (Fogelin 1985: 5).

One of Fogelin's paradigmatic cases is the disagreement over the morality of abortion (Fogelin 1985: 5). As Fogelin understands the issue, opposing parties may *agree* on number facts about the biological status of the fetus, as well as in moral terms, on the sanctity of human life, and yet nevertheless *disagree* on whether abortion is morally permissible. The issue, according to Fogelin, is that there exists no shared framework "within which reasons can be marshalled" (Fogelin 1985: 3) to settle the disagreement one way or the other.

Fogelin turns to some of Ludwig Wittgenstein's remarks published as *On Certainty* (OC 1974) to elaborate on the issue. What interests Fogelin is Wittgenstein's focus on a series of propositions that have the *form* of empirical propositions, yet *cannot be treated as such*, as they seem on reflection so fundamental that neither the possibility of doubt nor justification appear to make sense "if making judgements is to be possible at all" (OC 1974: 308):

... the *questions* that we raise and our *doubts* depend on the fact that some propositions are exempt from doubt, are as it were like hinges on which those turn (OC 1974: 341).

... If I want the door to turn, the hinges must stay put (OC 1974: 343).

In the first instance, Fogelin (1985) is tempted to suggest that deep disagreements arise from a clash of such "hinge propositions", or what he calls "framework propositions" (Fogelin 1985: 5). This enables him to reconceptualise the site of the so-called abortion debate to a clash between a pro-life hinge proposition ("The fetus is a person") and a pro-choice hinge-proposition ("The fetus is not a person"), neither of which, for the respective parties, make sense to be meaningfully doubted or justified. However, Fogelin goes on to suggest that

... when we inquire into the source of a deep disagreement, we do not simply find isolated propositions ("The fetus is a person."), but instead a whole system of mutually supporting propositions (and paradigms, models, styles of acting and thinking) that constitute, if I may use the phrase, a form of life (Fogelin 1985: 5-6).

One of the more controversial points in Fogelin's account is the suggestion that since deep disagreements "by their nature, are not subject to rational resolution" (Fogelin 1985: 7), all that may be left to do when confronted by one is to "fall back on persuasion" (Fogelin 1985: 7). This suggestion draws on Wittgenstein's very brief remarks in *On Certainty* on persuasion, understood as what we do "[a]t the end of reasons" (OC 1974: §612). At one point, Wittgenstein describes persuasion as an attempt to give an opponent one's "picture of the world" (OC 1974: §262; henceforth "world-picture"). What this

amounts to in practice will vary from case to case, but for one's world-picture to change suggests a fundamental shift in the "inherited background" against which one determines what is true and what is false (OC 1974: §94).

At another point, however, Wittgenstein invites us to think of persuasion as "what happens when missionaries convert natives" (OC 1974: §612), following a clash between two people that has devolved into mutual accusations of the other as a "fool" and/or a "heretic" (OC 1974: §611). It is the latter example that Fogelin sites in his discussion of persuasion, and though he does not acknowledge it, the remark seems to reveal the very real possibility of abuses of power and/or aspects of dishonesty and manipulation associated with attempts at persuasion. Nevertheless, Fogelin maintains that sooner or later we may have to bite the bullet and acknowledge the need for non-rational forms of persuasion to resolve deep disagreements.

II. Pritchard on Deep Disagreement: The Über Hinge Commitment and Side-On Rational Persuasion

Duncan Pritchard suggests contrary to Fogelin, that non-rational persuasion is "a worrying way for deep disagreements to get resolved" (Pritchard 2021: 1120). Pritchard's "worry" here is tied both to the kinds of abuses that can be read into Wittgenstein's previous remark on missionary conversions (cf. OC 1974: §612; Pritchard 2021: 1120), and to a more general philosophical concern about the threat of epistemic relativism that may follow from epistemically incommensurable disagreements (Pritchard 2021: 1118). To address these worries, Pritchard maintains that "there will always be a rational way of engaging with the other party" (Pritchard 2021: 1123), and thus, a rational pathway towards the resolution of deep disagreements.

To this end, Pritchard proposes the existence of what he calls the "über hinge commitment" (Pritchard 2021: 1120): the general idea that each of us operates with an overarching certainty "that we are not radically and fundamentally in error in our beliefs" (Pritchard 2021: 1120). Though the über hinge commitment is not explicitly mentioned by Wittgenstein in *On Certainty*, Pritchard proposes that the idea helps to make sense of interpersonal understanding despite apparent divergences in what he calls "hinge commitments" (Pritchard 2021: 1118).

According to Pritchard, though many of my hinge commitments will diverge from those I interact with, this is largely a superficial matter. For example, though the hinge commitment “My name is Ryan Manhire”, or “My native language is English” will diverge from the hinge commitments depicting the names and native languages of others, they are nevertheless manifestations of our respective *über* hinge commitments, to the extent that we do not take ourselves (or each other) to be radically mistaken about our names or the languages we speak (Pritchard 2021: 1120).

Pritchard uses this insight to argue that these superficial divergences “provide no basis for thinking that [...] we do not have shared standards of evidence, or lack shared epistemic principles” (Pritchard 2021: 1121). For Pritchard, this means that even where divergences in hinge commitments are *not* superficial, such as in cases of divergences in religious, political, or ethical hinge commitments, if the parties “are so much as intelligible to each other”, they will “inevitably have overlapping hinge commitments”, and thus, the depth of their disagreements “will only be partial” (Pritchard 2021: 1122). This enables Pritchard to argue for a “side-on” approach to persuasion that remains within a broadly rational framework shared by all parties. Over time, Pritchard suggests, appeals to common ground (shared beliefs and hinge commitments) between the parties may enable one to shift the other’s hinge commitments:

More precisely, as their wider set of beliefs changes, so too will the specific hinge commitments which manifest their *über* hinge commitment (which never changes) (Pritchard 2021: 1123).

With this broadly rational pathway towards resolution in place, Pritchard has since refined his approach to consider the kinds of topics we might deeply disagree about, and the kinds of hinge commitments that might be associated with them. This leads to a distinction between “quotidian hinge commitments” and “axiological hinge commitments” (Pritchard 2023, 2025). “Quotidian hinge commitments” refer to the kinds of “common sense claims” (Pritchard 2023: 312) found throughout *On Certainty* that most of us agree upon, “to the extent that we would struggle to understand someone who genuinely didn’t share them” (Pritchard 2023: 312; cf. OC 1974: §156). For similar reasons of intelligibility, this category also includes certain moral hinge commitments (or “moral certainties”, as they are often referred to in related philosophical

discussions) such as Fogelin's example of the sanctity of human life (Fogelin 1985: 5) as well as other familiar examples such as "Murder is wrong" (cf. Pleasants 2009, 2015; Ranalli 2021).

"Axiological hinge commitments", by contrast, concern "the fundamental values that are encoded in one's worldview" (Pritchard 2023: 312; cf. Pritchard 2021, 2025), including but not limited to moral, political, and religious hinge commitments (Pritchard 2023: 313). For Pritchard, these commitments "capture our deepest values; what we care about at a basic level" (Pritchard 2023: 302). It follows, for Pritchard, that potential sites for deep disagreements would therefore involve clashes between axiological hinge commitments in a variety of intercultural ('Eastern' versus 'Western' values), intracultural (new versus old), and intergenerational (old versus young) disagreements (Pritchard 2025: 18).

Significantly, though both categories are equally nonsensical to doubt and/or attempt to justify, axiological hinge commitments only become suitable candidates for deep disagreements once they have become "considered judgements" (Pritchard 2023: 307). As I understand Pritchard's point here, a dispute between different cultures, for example, would be merely shallow, if the axiological hinge commitments at the heart of the disagreement remained mere by-products of the encultured world-pictures the respective parties were initiated into as children (Pritchard 2023: 307). Such unreflective hinge commitments "might well be unstable in response to challenges [...] as the subject is prompted to doubt their status" (Pritchard 2023: 307). It is only when these axiological hinge commitments are later reflected upon and explicitly endorsed that "they are bound to be highly resistant to counterargument" (Pritchard 2023: 308), and hence, according to Pritchard, suitable candidates for possible deep disagreements.

III. The Importance of the Messy World of Real-Life Cases

At this point, I want to take a step back from these competing accounts of deep disagreement to consider a limitation in the discussion, and what this limitation might mean for where to go from here. This concerns the narrowness of the examples that Fogelin and Pritchard use in support of their respective accounts.

The importance of examples for Wittgenstein is reflected in his observation that “[o]nly in the stream of thought and life do words have meaning” (Z 1967: §173). Such reminders are used to combat confusions in philosophical analysis stemming from assumptions about the meanings of words or phenomena arising from their detachment from context and use (cf. PI 2009: §340). In *On Certainty*, a related sentiment appears when Wittgenstein asks himself, in the midst of his investigation into the nature of hinge propositions,

[a]m I not getting closer and closer to saying that in the end logic cannot be described? You must look at the practice of language, then you will see it (OC 1974: §501).

In this way, as Lars Hertzberg suggests, examples from familiar practices of language “are not to be regarded as data for the construction of a philosophical story, rather the examples *are* the story” (Hertzberg 2022: 2).

To be fair, Fogelin acknowledges the importance of context towards the beginning of his discussion of deep disagreement, noting that “arguing is something people do”, and “they do it for a wide variety of reasons, in an effort to achieve very different purposes” (Fogelin 1985: 2). Nevertheless, context appears on the peripheral of Fogelin’s discussion of abortion as the paradigmatic case of deep disagreements, almost as an afterthought. For instance, though he imagines how “The fetus is a person” may be a hinge proposition bound up with religious forms of life that maintain an immortal soul enters the fertilised egg soon after conception, he concedes he does not know how accurately he is representing such a person, as his “heart is not really in it” (Fogelin 1985: 5). Similarly, though Pritchard notes several possible examples of deep disagreement, constituted by clashes of axiological hinge commitments (intercultural, intracultural, and intergenerational deep disagreements), he too seems to leave the use of these examples at the peripheral of his investigation, merely gesturing towards the kinds of topics that *might* result in deep disagreements, which, importantly, *follows from* the solution to the problem of deep disagreements he has already outlined.

What I am suggesting here is that, in different ways, Fogelin and Pritchard limit themselves to using rather shallow forms of examples in the service of preconceived philosophical stories about deep disagreement, rather than

allowing potential examples *of* deep disagreement to illuminate the phenomenon. We are thus yet to see what a particular case may look like “in the stream of thought and life” (Z 1967: §173), and what kinds of concerns and meaningful resolutions may be relevant to the parties involved (cf. Levi 2000).

In order to move the discussion forwards, I will briefly turn to an example to suggest the kinds of difficulties that the notion of deep disagreements might present us with if we consider how they occur in what Victoria Lavorerio describes as “the messy world of real-life cases” (Lavorerio 2021: 428). The example is comprised of both the lyrics and music video to Kendrick Lamar and Taylour Paige’s song ‘We Cry Together’ (Lamar and Paige 2022a, 2022b), which depict an emotionally charged argument between a man (played by Lamar) and a woman (played by Paige) in an intimate, heterosexual relationship. The argument between the man and the woman shifts across a variety of topics, from personal “accusations of lacking care and attention, and the threat or accusation of infidelity” (Manhire & Kronqvist 2025: 8), to more public derived, “general characterisations of what *all men* and *all women* are [really] like, and how the other is *just like the rest of them*” (Manhire & Kronqvist 2025: 8). These dynamics suggest that the disagreement is not *about* one particular topic, but about an unspoken difficulty in the interpersonal relation between the parties themselves, one they are both hesitant to address. This reading is supported by the audio version of the song’s closing frame, following from the disagreement left unresolved: “Stop tap-dancing around the conversation” (Lamar & Paige 2022a).

Two points make this example relevant to the current discussion:

First, in the moments where particular facts or descriptions of states of affairs are brought up, they are outright rejected by the other party. This can be seen for example in moments focusing on whether certain gifts were given in good faith or ulterior motives, and whether the man supported the woman during a period of crisis, or worked to keep her in that period of crisis so that he could maintain the upper hand in their relationship. Absent a neutral perspective from which to ascertain the truth of such claims, the argument seems to follow Fogelin’s initial description of a disagreement that is “immune to appeals to

facts” (Fogelin 1985: 5), and that persists despite agreement on moral convictions, in this case the importance of love, care, and support within the confines of an intimate relationship.

Second, both parties appear to reject the other’s appeals to facts *regardless* of the other party’s particular accusation, *if* that accusation presents the accused party in a negative light. Such rejections are then followed by counter-accusations that present the accusing party themselves in a negative light. At one point, for example, the man explicitly accuses the woman of being emotionally unstable and “dumb”, in response to her accusations against him of being distant and failing to show her adequate love and care (Lamar & Paige 2022a). In response to *these* accusations, the woman retorts that she is neither “slow nor ditsy” (Lamar & Paige 2022a), and that the man is using reverse psychology in an effort to avoid acknowledging the shortcomings she had just raised.

While the first point positions the example in the realm of Fogelin’s initial description of deep disagreements, the second point may reveal a dimension of deep disagreements that appears at odds with how Pritchard conceptualises the possibility of their rational resolvability via his appeal to the über hinge commitment (Pritchard 2021: 1120). While Pritchard at one point suggests that “a general intellectual virtue of humility would demand that one should be willing to consider the possibility that one has made a mistake, and hence take seriously the opposing judgement” (Pritchard 2023: 307) in a deep disagreement, the man and the woman’s respective über hinge commitments seem to lock them in a dynamic in which both seem to take themselves as the sole party with the right to claim that “I am being rational and moral, while you are being irrational and immoral” (cf. Manhire & Kronqvist 2025: 11). Interestingly, this seems to both foreclose the possibility of each party considering the possibility of mistake (Pritchard 2023: 307), while nevertheless enabling them to (perhaps self-deceptively) maintain the “general virtue of humility” (Pritchard 2023: 307) that they are, generally speaking, open to the possibility of mistake. It is just that both parties hold that “In *this* instance, *I* cannot be making a mistake”. Each party may even claim that they are speaking from the perspective of the kind of “considered judgement” that Pritchard suggests is required for an axiological hinge commitment to be a suitable candidate for deep disagreements. However, a considered judgement

initially *derived* from the “inherited background” (OC 1974: §94) against which one *makes* considered judgements is always at risk of becoming self-fulfilling prophecy, particularly in response to accusations that one has wronged another person (cf. Backström 2023).

What these dynamics suggest is that there may be situations in which the über hinge commitment does not point to the possibility of the rational resolution of deep disagreements, but *to the very difficulty it presents us with*. A difficulty that, as Fogelin suggests, may “not be subject to rational resolution” (Fogelin 1985: 7), even indirectly. This appears to lead us back to Fogelin’s initial suggestion that “we have to fall back on [non-rational] persuasion” (Fogelin 1985: 7). But what of the very real concerns about abuse and deception here? To this I suggest that though persuasion can be used in tandem with tactics of coercion, bribery, or deception, this does not necessitate that such tactics are inherent to persuasion itself. Indeed, what the dynamics of the example I have used today may also show is that whether a certain case of persuasion is coercive or deceptive may itself depend on the existential cost to oneself in accepting the other’s picture of the world, and what it says about oneself.

References

Backström, Joel (2023) “Unbearable Certainties”, in: Cecilie Eriksen, Julia Hermann, Neil O’Hara and Nigel Pleasants (eds.) *Philosophical Perspectives on Moral Certainty*, New York: Routledge, 76-97.

Fogelin, Robert John (1985) “The Logic of Deep Disagreement”, *Informal Logic* 7, 1-8.

Godden, David (2019) “On the Rational Resolvability of Deep Disagreement Through Meta-argumentation: A Resource Audit”, *Topoi* 38, 725-750.

Godden, David M.; Brenner, William H. (2010) “Wittgenstein and the Logic of Deep Disagreement”, *COGENCY* 2, 41-80.

Hertzberg, Lars (2022) *Wittgenstein and the Life We Live with Language*, USA: Anthem Press.

Lamar, Kendrick; Paige, Taylour (2022a) “We Cry Together” [lyrics], *Mr. Morale & the Big Steppers*. <https://genius.com/Kendrick-lamar-and-taylour-paige-we-cry-together-lyrics> (Accessed: March 12 2025).

Lamar, Kendrick; Paige, Taylour (2022b) “We Cry Together” [music video]. <https://youtu.be/wUGyZM9rcnY?si=tFUANYBqMaqlpYdG> (Accessed: March 12 2025).

Lavorerio, Victoria (2021) "The fundamental model of deep disagreements", *Metaphilosophy*, 52, 416-431.

Levi, Don (2000), *In Defence of Formal Logic*, Dordrecht: Springer.

Manhire, Ryan & Krongvist, Camilla (2025) "Tap-Dancing Around the Conversation: Difficulties in an Intimate Deep Moral Disagreement", *Synthese* 205, pp. 1-24.

Pleasants, Nigel (2015) "If Killing Isn't Wrong, then Nothing Is: A Naturalistic Defence of Basic Moral Certainty", *Ethical Perspectives* 22, 197-215.

Pleasants, Nigel (2009) "Wittgenstein and Basic Moral Certainty", *Philosophia*, 37, 669-679.

Pleasants, Nigel (2008) "Wittgenstein, Ethics and Basic Moral Certainty", *Inquiry*, 51, 241-267.

Pritchard, Duncan (2025) "Axiological Hinge Commitments", *Synthese* 205, 1-22.

Pritchard, Duncan (2023) "Understanding Deep Disagreement", *Journal of Philosophical Studies* 31, 301-317.

Pritchard, Duncan (2021) "Wittgensteinian Hinge Epistemology and Deep Disagreement", *Topoi* 40, 1117-1125.

Ranalli, Chris (2021) "Moral Hinges and Steadfastness", *Metaphilosophy* 52, 379–401.

Wittgenstein, Ludwig (2009) *Philosophical Investigations*, G. E. M. Anscombe, P. M. S. Hacker and J Schulte (trs.) 4th edn, West Sussex: Wiley-Blackwell.

Wittgenstein, Ludwig (1975), *On Certainty*, G. E. M. Anscombe and G. H. von Wright (eds.) D. Paul and G. E. M. Anscombe (trs.) Victoria: Blackwell Publishing.

Wittgenstein, Ludwig (1967) *Zettel*, G. E. M. Anscombe and G. H. von Wright (eds.) G. E. M. Anscombe (trs.) Oxford: Basil Blackwell.

Acknowledgement

Work on this paper was supported by funding from Ella Georg and Ehrnrooth Foundation for the project *Deep Moral Disagreements on Sex and Gender* (DMD-SG) and Kone Foundation (Grant 202206718) for the project *Ethical Explorations: Rethinking Our Sexual Relations* (EEROS).

Simone Weil and Wittgenstein on Character

Nimrod Matan (Tel Aviv, Israel)

Abstract

Simone Weil's discussion of the concept of character, as that which conditions our reactions, divorces it from the domain of choice and will, presenting it as belonging to the domain of *world* rather than as a psychological concept. Stemming from Kant's discussion of character, character can be conceived, following Weil, as tantamount to the world of the person as a whole rather than to the person's disposition to react. Character is unknowable: although we cannot but attribute character to people, we cannot point at something that could determinately justify the attribution of character.

My claim is that Wittgenstein's differentiation between the world of the happy and the world of the unhappy (in *TLP* §6.43) reflects a similar conception of character as *worldly* rather than psychological.

According to this view, character conditions reactions in that it determines the world in which a person acts and reacts – its limits, possibilities, even the objects that populate it, and the time of their existence. I elaborate on this conception of character through a reading of an episode in book 24 of the *Iliad*, in which Achilles exhibits his magnanimous character when he concedes to release Hector's body to King Priam.

1.

In her 1941 *Notes on the Concept of Character*, Simone Weil identifies two challenges that a definition of the concept of character should meet, given that character is an “invariant common to the possible reactions of a single person and that limits his possible reactions” (Weil 2015: 98) and that reactions are conceived of widely enough as to include acts, gestures, attitudes, and words. The first challenge is to account for how character may be modified by circumstances: how it may change in light of external conditions, despite its invariability. Posed as a moral challenge, this can be put as the individual's endeavor to prevent the “empire of circumstances” from breaking his or her character. The second challenge is to account for cases when someone acts out of character, in which we nonetheless wish to continue to attribute to them their supposed character. A variant of the second challenge involves cases where a break in character appears not to be momentary but seems rather the result of a willful effort on the part of the person in question, who endeavors to change his or her character and may even succeed in doing so.

Weil nods in Kant's direction when mentioning his idea that character is "intuited from the outside", adding: "but not by the same givens" (Weil 2015: 98). The idea expressed here is that the sequence of reactions a person produces is indefinitely unlimited and character is always attributed to a person based on an unspecified subset of this sequence. Character thus becomes something of a mystery, which is a key problem for ethics, since "we cannot pose a moral problem without putting the concept of character at its center", and moreover, for philosophy in general, since "one could not even think about a human being if one did not think that he already had a character that is the same through all his reactions". The mystery lies in the unknowability of character: although we cannot but attribute character to people, "this character cannot be known by anybody", in the sense of pointing at something that could determinately justify the attribution of character. Although we require that an individual possess a definite character in order to conceive them as a person, what we strive to define is "something defined in principle but not defined in fact" (Weil 2015: 98).

Weil's passing mention of Kant provides sufficient justification for dwelling longer on Kant's discussion of character in the *Critique of Pure Reason*, where he articulates the philosophical problem of character in terms of the need to reconcile intelligible character, an a priori notion of character, pertaining to the domain of reason, and empirical character, which is character as it is manifested in a person's actions (and we may extend Kant's *actions* to include *reactions* in general, i.e., thoughts, words, etc., following Weil) in reality. While intelligible character determines all actions a priori, "reason ... declares ... certain actions to be necessary which nevertheless *have not taken place*, and which perhaps never will take place" (Kant 2003: 308). This can therefore be taken to be a third challenge to defining character as a concept: how to account for the fact that while character a priori determines the possibilities of reaction, the scope of reaction may be limited a posteriori, so that certain reactions deemed necessary by character might simply, contingently, not happen to occur.

Let us use Achilles' character trait of magnanimity as an example. Achilles' magnanimity is famously manifested in book 24 of the *Iliad*, where he grants King Priam's plea to release the body of his son Hector so that it can be brought to burial. If Achilles is magnanimous, then he should continue to be

magnanimous even if Priam had failed to get through the Greek guard and reach Achilles' tent, or if Achilles had been killed by Paris before and not after Priam's visit, in which cases his magnanimity would not have manifested itself. And yet there is nothing but his gesture of magnanimity towards Priam to hold to when aiming to justify the attribution of magnanimity to Achilles. His magnanimity was lying in wait, as it were, throughout his life until it manifested itself in the encounter with Priam.

According to Kant, the estimation of an action is based on the intelligible character. In relation to the empirical character we may explain an action as caused also by miseducation, disposition, bad company, circumstances, etc. (Kant 2003: 311-312). Thus, in a world where Achilles did live long enough to manifest his magnanimity, as recounted in the *Iliad*, we base our estimation of the release of Hector's body as an act of magnanimity (rather than, perhaps, a token of adherence to the norms of war, and thus a manifestation of conformity), on his being magnanimous a priori, and yet the only token of his magnanimity is his gesture of releasing the body.

The concept we seek to define is therefore one which, on the one hand, provides a priori conditions for action (and thought), and on the other hand is manifest only via a series of contingent occurrences. When aiming to describe a person's character we have nothing more than his or her biography to base our description on, and yet we see character as a priori conditioning action, with a rigidity which is more decisive than disposition. The magnanimity of Achilles made it necessary that he release the body of Hector, not merely likely. He was compelled and not simply inclined to do so. Character's reach into the world is of a different order than that of disposition or inclination and other psychological concepts. The weight of the particular (biographical) case in relation to the a priori condition is categorically more substantial here than in the case of psychological concepts. Although each sort of case is contingent, there is an air of necessity to the particular case that cannot be explained by the disposition-particular action scheme.

If character, then, is not a psychological concept, which domain does it belong to?

Weil outlines the components of an attempted definition of character thus: (1) “a system of obstacles that hinder or make possible actions or thoughts in such and such a direction”, (2) the “order of things [that] exist or does not exist” for the person, and (3) the “way of being a certain sort of person in relation to time” (Weil 2015: 98-99).

The obstacles or conditions for action and thoughts are described here in terms not of psychological dispositions or capacities but rather of *things* and of *time*. Weil’s outline of a definition places character in the domain of *the world* rather than in the realm of the person. A psychological framework would have character consist in the psychical conditions for a person’s possibilities of reaction to the world, or in the generation of a series of contingent occurrences which constitute his or her biography. Under the worldly framework (*worldly* in the sense of belonging to the category of *world*), on the other hand, character is the world of the person: the objects which exist for him or her, and the time during which he or she could be described as having a world, or as living.

It is therefore clear why Weil should make an effort to divorce character first and foremost from the domain of choice or will. Action and thought are determined only by circumstances and character, not by choice, and hence in the particular case, the individual has no freedom of choice to react as he wills (Weil 2015: 99). A change in character is due to external circumstances only. Such a change can be momentary, which makes it a case of acting out of character, or enduring, in which instance it represents genuine character change. To strive to willingly change one’s character means, therefore, not to directly act (or think) but to willingly, freely place oneself “in such and such circumstances that in the course if things will act on us from outside”. The act of the will orients one’s attention “in the moments ... that do not engage us, where circumstances do not solicit any choice from us” (Weil 2015: 99-100). The role of the will lies not in orienting action directly, i.e. in making a choice in the face of a particular case, but in shaping the world so as to allow a certain character to manifest itself. Wishing to become magnanimous does not consist in insisting on acting magnanimously in a particular case but in orienting oneself through “moments that appear to be the most insignificant, the most indifferent moments of our lives” (Weil 2015: 100), so as to let one’s world be shaped in such a way that would prompt magnanimity.

2.

In the sections discussing the impossibility of ethical propositions in the *Tractatus*, Wittgenstein famously claims that the difference between the good will and the bad will does not consist in any factual difference. It is, rather, a difference in the world as a whole: “In short the effect must be that it become an altogether different world. It must, so to speak, wax or wane as a whole”. In the second part of the same section he moves from the good/bad distinction to the distinction between the happy and the unhappy, namely, between two opposed characters. He concludes the section: “The world of the happy man is a different one than that of the unhappy man” (TLP 1963: §6.43).

The claim that the difference between the happy and the unhappy consists in the *worlds* of the two being different does not only mean that it is a difference that cannot be expressed in language, since it's not about any fact *in the world* but rather about the world as a whole. I take it also to mean that being happy and unhappy are *worldly* rather than psychological conditions. To conceive of character this way means to think of it as simple: character corresponds to a world as a whole. The opposed conception of character would treat it as composite, analysable to psychological or behavioural concepts. Such analysis would be psychology. Wittgenstein objects to the very idea that a soul could be conceived of as composite when he writes that “a composite soul would not be a soul any longer” (TLP 1963: §5.5421). Thinking of the soul as simple rather than as composite means not to think of it as describable in factual propositions. To describe a soul, its character, in a non-factual, non-psychological way would mean to view the person as a whole, as if from outside all happenings and, being so, or *sub specie aeterni*, in analogy to how art presents objects (CV 1980: 7). If the world of the happy is another than the world of the unhappy, then viewing a person *sub specie aeterni* amounts to the same thing as viewing the world *sub specie aeterni* (TLP 1963: §6.45).

It is noteworthy that in Kant, too, we find a trace of the understanding that character is worldly rather than psychological, that character determines the world of the person rather than his dispositions or inclinations to act in it, when he claims that Intelligible character (and reason in general) is the cause not of action or anything in the “world of phenomena” directly, but rather of empirical character itself, which in turn is the causal source of action. If empirical character is the disposition to act (or react) in certain ways, then the

intelligible character is the cause of this disposition. It conditions not any particular action but rather the space of action as a whole, or the world of the person.

In his 1916 Notebooks Wittgenstein discusses the idea that character (or spirit) can be inferred from physiognomy and reflects on whether the relation between the latter and the former is causal. He writes:

Now is it true (following the psycho-physical conception) that my character is expressed only in the build of my body or brain and not equally in the build of the whole of the rest of the world? This contains a salient point. This parallelism, then, really exists between my spirit, i.e. spirit, and the world. (NB 2004: 85)

The parallelism between character and world is expressed in §6.43 in terms of unity of character and world. Character is therefore exhibited not only in action but is also extended to such things as the body, a person's name, and his surroundings, which no doubt belong to the "empire of circumstances" and not to the domain of the will:

Suppose we ask ourselves "what proper name would suit the character of this man"--and portray it in sound? The method of projection we use for the portrayal is something which as it were stands firm. (A writer might ask himself what name he wants to give to a person.) But sometimes we project the character into the name that has been given. Thus it appears to us that the great masters have names which uniquely fit the character of their works. (PG 1984: 179)

The encompassing, worldly nature of character can be made clearer by analogy with how we conceive of literary characters. The character of the literary character reaches out to the "reality" of the world it inhibits and determines the details of this world as it is depicted in the work of literature. When examining Bernard Williams' practice of philosophically studying literary characters as a means to arrive at ethical insights, Stephen Mulhall writes that the self-possessing character is "constituted by categorical desires, commitments or ground projects", which implies that character, as for Wittgenstein, is simple rather than composite. The key examples Mulhall examines in this context are Gauguin's "project" of "becoming a great painter"

and Anna Karenina's "project" of "finding romantic fulfilment" (Mulhall 2007: 369). The description of both Gauguin and Anna Karenina as possessed by a project is tantamount to describing them as, say, a painter and a lover, which – like the happy and unhappy – are readily understood as defining a world rather than a set of psychological traits. Being a painter or a lover means to live in a world, where a certain order of things exists and another does not, to use Weil's wording; a world that would be portrayed as a certain world (a painter's or a lover's world) in literature.

We can now return to the Achilles-Priam example. The correct way to describe Achilles' magnanimity in relation to his granting of Priam's plea is not to say that because he is magnanimous he chose to release the body of Hector, but rather that being magnanimous means that his world is so shaped that he *had* to release Hector's body. The fact that Priam infiltrated the guarded Greek camp and reached Achilles is no less an expression of Achilles' magnanimity than Achilles' reaction to Priam's plea. Thus, when Achilles asks Priam, "How could you dare to come alone to the ships of the Achaians and before my eyes, when I am one who have killed in such numbers such brave sons of yours?" (Lattimore 1967: 489), and invites the Trojan king to sit down beside him, and when he says to Priam, "it does not escape me that some god led you to the running ships of the Achaians. For no mortal would dare come to our encampment, not even one strong in youth. He could not get by the pickets, he could not lightly unbar the bolt that secures our gateway" (Lattimore 1967: 490), he does not merely describe Priam's determination, courage, or luck, he also describes the world in which his own magnanimity is manifested, and by that, in a way, he describes his magnanimity. It seems that the Homeric text makes an effort to divorce the release of the body of Hector by Achilles from the domain of Achilles' will or choice, and to present it as a facet of his biography, of his life or his fate. Although Achilles felt pity towards Priam (Lattimore 1967: 488), it is not in the moment when this feeling emerged and when they cried together that he said he would release Hector's body. Only after becoming aggravated with Priam does Homer have Achilles announce his decision, so as to detach it from the psychological context of pity and emphasize that it has already been decided a priori that he should do so, regardless of the emotions the encountered stirred in him: "No longer stir me up, old sir. I myself am minded to give Hektor back to you. A messenger came

to me from Zeus, my mother, she who bore me, the daughter of the sea's ancient" (Lattimore 1967: 490). The world of the magnanimous Achilles is such that Priam had to succeed in getting past the Greek guard, and Paris' fatal arrow had to wait until after this encounter has taken place, and it is his world that can be described as conditioning his reaction to the plea of Priam.

References

Kant, Immanuel (2003) *Critique of Pure Reason*, Translated by: Meiklejohn, J.M.D., Mineola, New York: Dover Publishing.

Lattimore, Richmond (trans.) (1967) *The Iliad of Homer*, Chicago and London: University of Chicago Press.

Mulhall, Stephen (2007) "The Morality of the Soul: Bernard Williams's Characters", in: Alice Crary (ed.) *Wittgenstein and the Moral Life*, Cambridge and London: MIT Press. 355-379.

Weil, Simone (2015) "Notes on the Concept of Character", in her: *Late Philosophical Writings*, Translated by Springsted, Eric. O., and Schmidt, Lawrence. E., Notre Dame, Indiana: University of Notre Dame Press. 97-102.

Wittgenstein, Ludwig (1963) *Tractatus Logico-Philosophicus*, Translated by Pears, D. F. and McGuinness, B. F. London: Routledge and Kegan Paul.

Wittgenstein, Ludwig (1980) *Culture and Value*, Translated by Winch, Peter. Chicago: University of Chicago Press.

Wittgenstein, Ludwig (1984) *Philosophical Grammar*, Translated by Rhees, Rush, Berkley and Los Angeles: University of California Press.

Wittgenstein, Ludwig (2004) *Notebooks 1914-1916, 2nd Edition*, Translated by Anscombe, G.E.M., Malden, Oxford and Carlton: Blackwell Publishing.

Wittgenstein vs. the Conceptual Engineers

Jacob McDowell (Chicago, United States)

Abstract

While other philosophers have claimed to have found a conservative attitude to conceptual change in the later Wittgenstein's writings, Tamara Dobler has recently brought Wittgenstein into service for the project of conceptual engineering. I disagree that Wittgenstein's philosophy is inherently conservative, but I also believe it is importantly at odds with the project of conceptual engineering. I believe, the project of conceptual engineering is founded upon certain pictures of 'meaning' and 'concepts' whose rejection was one of the most fundamental parts of the later Wittgenstein's philosophy.

I begin by outlining the project of conceptual engineering and how Dobler proposes to draw on Wittgenstein to aid this project. I then examine how Wittgenstein uses the term 'concept' in his later writings. I argue that there is an important use of the term there that aligns with 'picture' or 'dominant model'. There I argue, concepts are not wholly determinate, inexorable guides to use. Rather, uses of terms is fundamental and concepts may fit that use more or less well. On Wittgenstein's therapeutic approach, we call to mind uses of a particular term to dispel or revise faulty concepts or pictures of that term.

While other philosophers have claimed to have found a conservative attitude to conceptual change in the later Wittgenstein's writings, Tamara Dobler has recently brought Wittgenstein into service for the project of conceptual engineering. Here is a formulation of that project from Herman Cappelen that Dobler quotes approvingly:

“Motivated by removing a certain perceived deficiency in a concept, then main goal of a conceptual engineer is ‘to ask what our concepts ought to be or what extensions our terms ought to have’ (14).”

Dobler seeks to use Wittgenstein's work to respond to two challenges that have been raised for conceptual engineers: “the problems of trivialisation and implementation” 225)

The first challenge is “how can a conceptual engineer successfully *implement* her stipulative revisions?” (227) That is, it seems easy for a philosopher to stipulate that a given word will express a new concept – but much more difficult to actually get speakers implement the new concept – to reform their use of the word in line with the new concept, in favour of the old. Here is David Chalmers acknowledging just this challenge for conceptual engineering:

“... once you break down conceptual engineering into design, implementation, and evaluation, it becomes clear that concept design and concept evaluation are relatively straightforward, or at least don’t face the same sort of social obstacles. They suffice for many theoretical purposes. Concept implementation requires changing use, which is a difficult social project, but it’s possible.” (16)

The second challenge supposes the first to have been solved, but then goes on to ask, “how can we refrain from *trivialising* philosophical problem solving?” (228) What underlies the trivialization worry is that if philosophical problems are solved by conceptual engineering, then they are solved by replacing one concept for another – and that that’s not to solve the original problem but to change the subject. That is only a solution in a trivial sense.

I’ll start by outlining how Dobler addresses the implementation challenge. According to Dobler, the implementation challenge must be overstated because we constantly implement just the same sorts of revisions the conceptual engineer proposes to in everyday conversation. We do so because, according to Dobler, semantic or standing meaning is radically underspecified and needs to be pragmatically enriched or engineered in different contexts. As Dobler puts it,

“... the semantic (standing) meaning is underspecified, and it undergoes different modulations and enrichments in different contexts” (225)

Standing meaning is “underspecified” for Dobler in that words and other lexical items encode “families” of concepts that require revision or engineering to arrive at an appropriate speaker meaning.

“...a word typically encodes a whole *family* of related concepts ... communicating ... we routinely perform ‘revisions’ of standing meaning, highlighting only those concepts from the encoded family that are contextually relevant.” (230)

For these reasons, revising and engineering concepts is something that we know can be implemented because we already do it all the time:

... in the course of everyday communication we regularly ‘engineer’ underspecified word meanings in order to convey only the information

that is contextually relevant. ‘Revising’ word meanings is thus not something out of the ordinary or ... incredibly hard to implement” (234)

The idea that the standing or semantic meaning of terms is an underspecified family of concepts from which speakers select an appropriate concept is what Dobler terms ‘pluralist semantics’. Dobler’s use of the term family is meant as an echo of Wittgenstein and Dobler, in part rightly, sees both her and Wittgenstein as drawing attention to the fact that “*pluralism’s already there*” in our language – “that we use the [*same word*] in different ways for different purposes.” (244).

Dobler draws on this conception of pluralist semantics as well as Wittgenstein’s philosophical methodology to address the trivialization challenge: “By combining the idea of semantic underdetermination with pluralist semantics, it is possible to retain the notion of ‘change without discontinuity’” (225)

What Dobler takes from the idea of semantic underdetermination in this connection is the idea that engineering does not involve wholesale conceptual replacement. Instead, it involves revision or refinement of what’s already present. Dobler compares “calling this fact to mind” with the later Wittgenstein’s method of description in order to dissolve philosophical problems:

[F]ollowing Wittgenstein, my brand of pluralist conceptual engineering advocates revision by ‘calling to mind’ rather than revision by replacement. ... we remind ourselves that *pluralism’s already there*, i.e. that we use the word *know* in different ways for different purposes. (244)

Putting all these pieces together we get Dobler’s Wittgenstein inspired version of conceptual engineering, which proceeds as follows:

The first step in the engineering intervention is to recognise that a word is associated with a conceptual family: there is not one concept that a word encodes but many. A pluralist conceptual engineer does not seek to replace a ‘faulty’ concept but she brings to view the existence or the possibility of others. ... A conceptual engineer treats a philosophical problem by placing it in the pluralist perspective and *dissolves* it by revealing its normative status as a

result of accepting a certain model as dominant. We are free to champion different models for different purposes without seeking to banish individual concepts from the conceptual family. (225 – 226)

One might worry that Dobler's program will fall short of Cappelen's aims because it only allows for the selection of concepts from the pre-existing families associated with lexical items. Whereas the original hope of conceptual engineering was for the development and implementation of brand new concepts – not just the selection from some pre-set lot. Dobler still takes this to be possible though. Following David Chalmers, Dobler distinguishes *de novo* conceptual engineering from conceptual re-engineering – where *de novo* engineering is a matter of building new concepts – not just refining or selecting old ones (6). According to Dobler, *de novo* engineering is possible “insofar as it allows that the ‘model’ we currently associate with a word may shift overtime thus creating an opportunity for new candidate members to join the conceptual family.” (247)

I now turn to arguing that Wittgenstein's therapeutic approach to concepts is fundamentally different from that of Dobler and other conceptual engineers. Dobler has rightly identified a concern in Wittgenstein with revising concepts in the sense of the ‘dominant model’ or ‘picture’ we have of our use of a word. As I read the later Wittgenstein, ‘a dominant model’ or picture can lag behind the use or applications we make of an associated term – the picture we formed of our earlier uses of a concept can fail to fit the grammars of new applications of a term. Here, our use of language and our conception of that use can come apart, and attempting to fit the former to the latter begets philosophical confusions. I take this dynamic to be what Wittgenstein has in mind in *On Certainty* regarding confusions around knowledge when he wrote “at the bottom of this is a misunderstanding of the nature of our language-games.” (§599) Wittgenstein, I believe, often uses concept synonymously with ‘dominant model’ or ‘picture’ in this sense.

Consider his saying in *Remarks on the Foundations of Mathematics* that a “concept is something like a picture with which one compares objects.” (VII, §71) That is not to say that is the only use of the term ‘concept’ in Wittgenstein – as he himself says in the same section of *RFM*: “ ‘Concept’ is a vague concept.” (VII, §70).

Later, in the *Investigations*, Wittgenstein speaks of concepts forcing themselves on us: “What forces itself on one is a concept.” (§191). Wittgenstein describes himself about how certain concepts forced themselves upon him when he wrote the *Tractatus*:

... it sounds as if these languages [logical calculi] were better, more perfect, than our everyday language; and as if it took a logician to show people at last what a proper sentence looks like.

All this, however, can appear in the right light only when one has attained greater clarity about the concepts of understanding, meaning something, and thinking. For it will then also become clear what may mislead us (and did mislead me) into thinking that if anyone utters a sentence and means or understands it, he is thereby operating a calculus according to definite rules.” (1953/2009, §81)

As I read this passage, Wittgenstein is saying that his concepts of “understanding, meaning something, and thinking” failed to fit the uses he and other speakers made of those terms in everyday conversation. Attempting to squish the varieties of uses of those terms into his concept or model led Wittgenstein to the distorted dogma that what was needed was “a logician to show people at last what a proper sentence looks like.”

Wittgenstein’s therapeutic approach to revising concepts may sound broadly similar to that of conceptual engineering, but I believe that there are deep differences. Consider Chalmers’ summary of the three-stage process of conceptual engineering:

Our three stages are (a) designing a concept, such as proposing a meaning for a word, (ii) evaluating a concept, such as seeing how well this meaning can play certain roles, and (iii) implementing the concept, say ensuring the word is used to express this concept within a community. (15)

Here the goal is to change linguistic use or behaviour. In order to do that, one proposes a new concept or definition in the hope that that concept will guide linguistic behavior in the desired way. A concept here is thought of as something prior to use which determines ‘all the steps’ as it were. This is not to say that there are no cases where the application of the concept is

indeterminate – but it is supposed to be determined prior to use what the determinate and what the indeterminate cases are, and it just remains for speakers to find that out if they did not divine this from the armchair. Conceptual engineers treat the fashioning of concepts in this sense as comparatively easy: “... that concept design and concept evaluation are relatively straightforward”. (Chalmers, 15). The real difficulty is just the psychological and sociological one that Dobler labelled the implementation challenge. Wittgenstein, on the other hand, rejects the idea of that a concept is something specifiable and graspable prior to use that can serve to guide use. I take pessimistic answers to be implied in response to the following questions raised in *RFM*:

...can we not complete the construction of the concept as a receptacle for whatever application may turn up? May I not complete the construction of the ... and as it were prepare a form of language for possible employment? (V, §40)

We can of course put forward new terms and definitions for future use. But we must not fall into thinking of this as readying something “for whatever application may turn up”.

For Wittgenstein, one word (in some sense of ‘word’) may have different employments or uses: “... A meaning of a word is a kind of employment of it. For it is what we learn when the word is incorporated into our language.” (§61) Wittgenstein then encourages us to compare word meanings the functions a word has: “Compare the meaning of a word with the ‘function’ of an official. And ‘different meanings’ with ‘different functions’.” (§64) It is true that this is an invitation to comparison, rather than an injunction for identification. But in the next remark of *On Certainty*, Wittgenstein encourages us to view changes in function or employment as in turn driving changes in concepts, which in turn drive changes in the meanings of words: “When language-games change, then there is a change in concepts, and with the concepts the meanings of words change.” (§65) The point I take it here, is that use or function is primary. It is changes in that which in turn change concepts and the meanings of words.

The direction of change suggested by conceptual engineering seems to me to be exactly the opposite. Someone identifies a defective or non-ideal concept, one then develops a new improved concept. Adopting the new and improved concept, one's language use functions differently and, hopefully, better. Although there might an 'implementation challenge' – this challenge grants the conceptual engineer's overall picture, and then asks about the engineer's abilities to persuade people to adopt new concepts, or the vigilance of individuals in consistently employing the new concept in favour of the old. The Wittgensteinian challenge is instead directed at the order of priority supposed by the conceptual engineer.

Consider Wittgenstein's description of the extension of the concept of 'number' in the *Investigations*:

Why do we call something a “number”? Well, perhaps because it has a direct affinity with several things that have hitherto been called “number”; and this can be said to give it an indirect affinity with other things that we also call “numbers”. And we extend our concept of number, as in spinning a thread we twist fibre on fibre. ... (§67)

Here, as I read this remark, it is because we notice an affinity between some 'things' and other things we have already called 'numbers' that we come to call the former 'numbers'. Because our classificatory practice extends, so to does our concept of number. It is not that the concept is extended or engineered prior to these new classifications. The revised concept does not guide the new classifications. Rather, it is through the making of the new classifications, and revising our dominant model to reflect this, that our concept is extended.

Even if concepts are not to be thought of as guides to use specifiable prior to and independently of use, neither are they mere representations of that use that may accurate or inaccurate. Instead, there is an interactive or iterative relationship between use and concepts. There is something like a feedback effect, I believe, for Wittgenstein, where our understanding of our linguistic behaviour restricts how that behaviour might be extended or changed. I take this 'feedback' like aspect of concepts to be evident in Wittgenstein's remark that: “Concepts lead us to make investigations. They are the expression of our interest and direct our interest.” (PI 1953/2009, §570)

Being in the grip of a concept or picture can cause aspects of our linguistic behaviour to ossify and become particularly recalcitrant to change. Elsewhere a given concept might create apparent tensions in our uses of a word. Replacing those concepts may be an important part of change – but given Wittgenstein’s view of concepts we should not expect it to fit into the three neat, sequential and discreet stages of ‘design’, ‘evaluation’, ‘implementation’. From Wittgenstein’s use-based perspective, evaluation should be expected to first require some implementation which we evaluate and then feed back into our design. Whereas the process is, at least in theory, a sequential one for conceptual engineers, for Wittgenstein conceptual ‘design’, ‘evaluation’ and ‘implementation’ are related iteratively *both in theory and in practice*. The processes feed into one another and then back again.

In a sense, this makes for a more active possibility of adopting new concepts than that allowed for by Dobler. Recall that Dobler allowed for *de novo* conceptual engineering in that “the ‘model’ we currently associate with a word may shift overtime”. This sounds like a wholly passive process. On the Wittgensteinian approach, we actively make new uses and applications of terms all the time. Mere revisions or extensions of existing concepts may not prove adequate to capturing the evolved use. Realizing that, we may shift or change our associated concept – but that is not something that merely happens to us. Rather, it occurs in part because of our active use of the term in question.

I think that shows that Wittgenstein’s picture of linguistic change is fundamentally different from that presupposed by conceptual engineers. I now want to close by considering a possible response inspired by Dobler’s work. One might think that what’s required to bring the later Wittgenstein and conceptual engineering together is simply the acknowledge of the underdetermination of standing meaning and the consequent pluralism that Dobler argues for. I believe that there are important differences between Dobler and Wittgenstein though.

First, I do believe that Dobler is right to sense in Wittgenstein a view of linguistic change is a kind of ‘change without discontinuity’, and that the same word might have a plurality of uses. However, I believe that Dobler’s description of both phenomena is importantly different from Wittgenstein. Take first Dobler’s idea that standing meaning is underspecified. According to

this idea, words encode families of meanings from which speakers and hearers select the appropriate meaning in a given context. Wittgenstein's description of words having different senses or employments is interestingly different, however. In discussing different senses of 'understanding', such as 'understanding a sentence' and 'understanding a poem', Wittgenstein writes:

... has "understanding" two different meanings here? I would rather say that these kinds of use of "understanding" make up its meaning, make up my concept of understanding. For I want to apply the word "understanding" to all this." (1953/2009, §532)

Dobler, would presumably describe this case by saying that the word 'understanding' encodes at least two concepts and that we just select one of these in a given circumstance, thereby performing one of our regular 'revisions' of standing meaning" (230). Dobler, following Charles Travis, takes the idea of semantic underspecification to be a way of articulating Wittgenstein's criticism of the idea of meanings as bodies or shadows lying behind words. Here is Wittgenstein describing the idea to be rejected: "Meaning was something invisible behind the word which made it impossible to put it together [with other words] in a certain way – it, as it were, embodied the rules." (2017, 22).

Dobler's idea of semantic underspecification seems not to reject the idea of meaning as something lying behind words, but merely to reject the idea of unique meanings lying behind word types. Instead, for Dobler, families of meanings or concepts lie behind word types or lexical items. Recognizing something as a family resemblance concept is not to see it as encoding a family of different concepts – but to see the family resemblance structure exhibited in its use as characterizing the very concept it is.

I therefore do not believe that Dobler's idea of semantic underspecification is to be found within Wittgenstein's writings. Neither then will acknowledgement of that idea help to bridge Wittgenstein's therapeutic approach and the conceptual engineering approach. Instead, I believe, the project of conceptual engineering is founded upon certain pictures of 'meaning' and 'concepts' whose rejection was one of the most fundamental parts of the later Wittgenstein's philosophy.

References

- Cappelen, Herman (2018). *Fixing Language: An Essay on Conceptual Engineering*. Oxford: Oxford University Press.
- Chalmers, David J (2020). "What Is Conceptual Engineering and What Should It Be?" *Inquiry*, September, 1–18.
- Dobler, Tamara (2022). "Pluralist Conceptual Engineering." *Inquiry* 68 (2): 224–50.
- Wittgenstein, Ludwig (1969/1991). *On Certainty*, G.E.M. Anscombe and G.H. von Wright (eds.), Denis Paul and G.E.M. Anscombe (trans.). Oxford: Blackwell Publishing.
- Wittgenstein, Ludwig (1956/1983). *Remarks on the Foundations of Mathematics*, Revised Edition, G.H. von Wright, Rush Rhees and G.E.M. Anscombe (eds.), G.E.M. Anscombe (trans.), Cambridge, MA: MIT Press.
- Wittgenstein, Ludwig (1953/2009). *Philosophical Investigations*. Revised 4th Edition, P.M.S. Hacker and Joachim Schulte (eds.), G.E.M. Anscombe, P.M.S. Hacker and Joachim Schulte (trans.), Oxford: Wiley-Blackwell.
- Ludwig Wittgenstein (2017). *Wittgenstein's Whewell's Court Lectures, Cambridge, 1938–1941, From the Notes by Yorick Smythies*, Volker Munoz and Bernhard Ritter (eds.). Oxford: Blackwell.

Family resemblances, semantic drift, and polysemy

Henrique Mendes (Vienna, Austria)

Abstract

This article explores how polysemy—words with multiple, related meanings—silently underpins some of philosophy’s most persistent confusions. Drawing on Wittgenstein’s later philosophy, it argues that many conceptual puzzles arise not from the complexity of reality but from unnoticed shifts in the grammar of our words. Wittgenstein’s diagnosis of philosophical perplexity includes analogical projection, the craving for generality, and the formation of misleading conceptual pictures. The article adds that these tendencies are often sustained by the polysemous structure of key philosophical terms, which conceal semantic diversity under a single form. Building on Wittgenstein’s remarks, as well as insights from Rosch’s prototype theory and Lakoff’s notion of radial categories, the paper shows how family-resemblance concepts and polysemy create networks of overlapping meanings. These structures invite inferential slippages that fuel pseudo-problems and category mistakes. Two case studies illustrate the point: Wittgenstein’s analysis of the question “Where does thinking take place?” reveals how the spatial preposition “in” creates a misleading picture of the mind as a container; Augustine’s treatment of “measuring time” shows how the projection of spatial measurement onto temporal experience generates a paradox. In both cases, the grammar of familiar words masks deep structural differences. By foregrounding polysemy as a structural enabler of philosophical confusion, the article extends Wittgenstein’s therapeutic project: philosophy should not seek definitive answers but clarity about the ways language quietly misleads us.

1. The roots of philosophical confusion, family resemblances, and polysemy

In his later philosophy, Wittgenstein offers not a single diagnosis of the causes of philosophical confusion, but rather a repertoire of overlapping tendencies by which language leads thought astray. This plurality is reflected in two influential secondary sources: the first volume of G. P. Baker and P. M. S. Hacker’s *Analytical Commentary on the Philosophical Investigations* (1980/2005: 277–283) and Hans-Johann Glock’s *Wittgenstein Dictionary* (1996: 280). Each attempts to systematize Wittgenstein’s scattered remarks into a typology of philosophical error. According to Baker & Hacker, the sources of philosophical perplexity include i) misleading features of our grammar, ii) the projection of features between language games, iii) the operation of innumerable “pictures,” iv) an obsession with particular language forms, v) a tendency to model philosophical inquiry on scientific questioning, and cravings for vi) generality, vii) unity, viii) definitions, ix) explanations, x) necessities, and xi) even myths. Glock’s taxonomy overlaps significantly, acknowledging the craving for generality, projection between language games, philosophical pictures, and the

emulation of science as important factors, but also adding xii) analogies, xiii) the seduction of phenomenology, and xiv) the metaphysical quest for the unconditioned.

While the sources of confusion mapped by Glock and Baker & Hacker may appear varied, I believe they are usually underpinned by a shared structural feature: the architecture of family resemblance conceptual categories. As Wittgenstein famously argued, many of the concepts we employ in ordinary language do not conform to a model of strict definition but are instead held together by a network of overlapping similarities—no single feature is common to all, but each case bears a resemblance to others in the group. A canonical example is Wittgenstein's analysis of the word game: board games, card games, and ball games differ widely in structure and purpose, yet are grouped together not through any essence, but via overlapping similarities—a paradigm case of a family-resemblance conceptual category. This structure is clearly grounded in analogy: we group entities and phenomena not by identifying an essence or a set of necessary and sufficient features for their characterization, but by tracing chains of partial similarity.

Such analogical groupings, however, sit uneasily with the philosopher's cravings for theoretical clarity—most notably, the vi) craving for generality, for vii) unity, and for viii) definitions in the classical sense. Faced with a loosely organized concept like game, the philosophical impulse is often to seek a single feature or essence that all uses have in common—a tendency that stems from the desire to impose systematic order upon what is, in fact, an open-textured network of similarities. This craving for generality compels philosophers to ignore the local peculiarities of different uses, treating them as if they fell under a unified logical form. The craving for unity reinforces this tendency: where language offers us multiplicity and variation, the philosophical mind searches for cohesion and structural harmony.

But because family-resemblance concepts do not yield to strict definitions, this search is often frustrated—and redirected through analogies. A philosopher may fix on a single, clear-cut instance of a term and then ii) project its features onto other, less salient cases. Over time, these projections accumulate, forming what Wittgenstein calls iii) philosophical pictures (*Bilder*): seemingly coherent conceptual models in which multiple language games are illegitimately

collapsed into a single explanatory schema. These are not merely isolated errors, but ensembles of mistaken analogical projections layered over one another until they present the illusion of conceptual unity. A paradigmatic example is Wittgenstein's critique of the Augustinian picture of language in the beginning of the *Philosophical Investigations* (PI: §1). There, he identifies the assumption that every word denotes something as a projection onto all language from certain prototypical uses of words, such as pointing to objects and saying their names. This projection neglects the variety of grammatical functions words can have, such as giving orders, asking questions, making jokes, or expressing emotions (PI: §23). The resulting picture, while initially intuitive, becomes a source of confusion when used as a general model of linguistic meaning. It is worth noticing that, as this passage of the *Investigations* show, the phenomenon extends to metalinguistic terms themselves: *meaning*, *proposition*, *word*, *concept*, *sense*—; these too are most often family-resemblance terms (MWL: 273–274).

My central claim is that one specific linguistic phenomenon—*polysemy*—plays a particularly important and underacknowledged role in the emergence of philosophical pictures and the confusions they sustain. Whereas family resemblance allows us to group disparate cases under a single concept despite the absence of a strict definition—game being the canonical example—polysemy introduces semantic compartmentalization within a single word. It arises when semantic drift, often carried out through metaphor or analogy, leads to the stabilization of multiple, related but distinct senses under the same linguistic form. As an illustration, let us take a paradigmatic example of polysemy: the word “head.” In reference to the human or animal body, it denotes what is physically at the *top*, essential for life and *higher cognition*. In organizational contexts, “head” metaphorically designates the locus of *decision-making* and *authority*, the *top* of a hierarchical pyramid. On a table, the “head” is the end where an *authoritative* figure typically sits, someone responsible for making crucial *decisions* for the group. These uses share enough metaphorical or structural overlap to seem unified, yet they diverge significantly in meaning and inferential role.

Seen through Wittgensteinian lenses, polysemous terms are prone to generating grammatical illusions: their surface uniformity tempts us to treat their various senses as governed by the same logical grammar. In

philosophical discourse, this is especially dangerous. Because precision and conceptual clarity are at a premium, the embedded semantic diversity of polysemous words often goes unnoticed. As these meanings accumulate historically—sometimes feeding back and reshaping the conceptual space of older uses—the boundaries between them blur further. The result is a fertile ground for pseudo-problems and category mistakes, where arguments slip unnoticed between distinct senses of the same term. Polysemy, in this way, functions as a quiet enabler of confusion, giving the illusion of semantic unity where none exists. Words like “representation,” “object,” and “identity” are especially prone to this kind of slippage: each has both everyday and technical or philosophical senses that differ markedly in their inferential roles. A representation might refer to a physical depiction, a mental image, or a symbolic expression in a formal system—yet philosophers may treat it as if it were a unified ontological category. An object can be a tangible thing, an intentional correlate, or a logical constant, but these senses are often conflated in metaphysical debates. Identity, in turn, may signify numerical sameness, qualitative similarity, or personal selfhood—distinctions that are frequently blurred in philosophical arguments about personal identity or logical necessity. (A more extensive list of similar examples can be found in Baker & Hacker 1980/2005: 277).

In order to understand the role played by polysemy in philosophical perplexity, we must examine its structural development. As Glock (1996: 123) notes, an intriguing—even if underdeveloped—notion appears in Wittgenstein’s unpublished manuscript *Eine Philosophische Betrachtung*: the idea of “centers of variation” (*Zentren der Variation*) (MS-115: 221). There, Wittgenstein suggests that certain uses of a concept may serve as more central or paradigmatic, while others cluster around them in looser, less typical configurations. Although he never systematically develops this idea in later writings, it anticipates the key insight of what would later be called prototype theory in cognitive psychology and semantics. First articulated by Eleanor Rosch and her collaborators (Rosch & Mervis 1975), prototype theory—a framework whose Wittgensteinian inspiration is explicitly acknowledged in the title of its foundational paper—holds that many natural language categories are not defined by necessary and sufficient conditions, but are instead organized around highly typical instantiations. For example, in the

category “bird,” a robin or sparrow is typically judged more representative than a penguin or ostrich. These central cases anchor our understanding of the concept, while peripheral members are included by virtue of their resemblance to the prototype rather than by meeting a strict definitional threshold.

In *Women, Fire, and Dangerous Things*, George Lakoff deepens the prototype-based account of category structure by introducing the notion of *radiality*—networks of related but distinct senses organized around a central prototype (Lakoff 1987: 91–114). One paradigmatic example is the concept of “mother,” which includes: i) the genetic mother, who is the biological source of DNA; ii) the birth mother, the woman who gives birth; iii) the nurturing mother who raises the child; and iv) the legal mother, who is recognized as such by law, among others. While all of these senses are connected to a prototypical notion of “mother” via overlapping features pertaining to social function, biological role, or legal abstraction, they can diverge significantly in meaning and application. A woman may, for example, be a child’s mother in the legal sense but have no biological or emotional connection to the child. Lakoff argues that such polysemous structures are not arbitrary but arise through metaphorical and metonymic extension from a central prototype that concentrates the highest number of category-defining features—in this case, the nurturing, biological, legally recognized mother (Lakoff 1987: 74–76).

The radial model, I believe, helps explain how polysemy can emerge from categories structured by family resemblances. As peripheral uses are extended through analogy, domain shift, or institutional elaboration, they become semantically compartmentalized yet remain linked to the core concept. A radial structure is thus not a formless cloud but a network of senses anchored in prototypical usage, where each extension retains some traceable conceptual pathway back to the center. At the edges of these networks, it is often unclear whether we are dealing with a single family-resemblance concept or a polysemous term. This indeterminacy has important consequences for philosophy: it allows unnoticed shifts between inferential roles to take place, sustaining argumentative slippages under the illusion of semantic unity. Radiality, then, marks the threshold where analogical drift hardens into grammatical plurality—a fertile ground for philosophical confusion. In what follows, I examine how such structures manifest in the case of a few

philosophically relevant terms, showing how their polysemous layering can silently structure and sustain deeply entrenched conceptual illusions. The case of “thinking” and its supposed location, discussed by Wittgenstein in the *Blue Book*, offers a particularly instructive example, as does his discussion about the notions of “time” and “measurement” in Saint Augustine’s philosophy.

2. The “location” of thought and the “measurement” of time

Among the many cases in which polysemy quietly supports philosophical confusion, few are as emblematic as the case of thinking and its supposed location. In the *Blue Book*, Wittgenstein critically examines the deeply intuitive, yet misleading tendency to ask where thinking occurs (BBB: 7). When pressed, most people are inclined to reply that it happens “in the head,” “in the brain,” or “in the mind.” But this apparently innocent answers conceal a host of grammatical assumptions. What exactly is meant by “in”? Is this preposition used in the same sense across different domains? And in what sense is the mind something that could contain anything? Wittgenstein’s insight, as I interpret him, is that this inclination is not incidental. Rather, it arises from the very grammar of our ordinary language, in which spatial prepositions such as “in” co-occur with nominalized mental activities like “thinking” or “perception.”

This grammatical pairing leads us to treat thought as a kind of entity or event located within a bounded space—an image reinforced by the persistent dichotomy between “inner” and “outer” experiences. Yet, as Wittgenstein remarks, our tendency to localize thought “in the head” has also another source: the grammatical parallel between “thinking” and bodily activities such as “writing” or “speaking” tempts us to postulate an analogous inner activity that corresponds to the word “thinking” (BBB: 7). The mere existence of such a noun—grammatically formed like “speaking” or “walking”—suggests that we should look for a referent: a mental activity that takes place somewhere. This grammatical illusion leads us to construct a pseudo-picture of the mind as a container or enclosed space, with thought as its content, and the head as its location. This spatial model of the mental—underpinned by the polysemous layering of “in” and the grammatical form of mental predicates—demonstrates how ordinary expressions can solidify into philosophical pictures.

Another particularly telling example of how polysemy sustains philosophical perplexity is found in Augustine's reflections on the concepts of "time" and "measurement," a passage of his *Confessions* to which Wittgenstein refers when discussing how philosophical confusion arises from the grammar of such words. Wittgenstein invites us to reconsider what is really being asked when we pose questions of the form "What is time?"—noting that they often express not a demand for information but a deeper kind of grammatical unease (BBB: 26–27). This diagnosis finds a clear illustration in Augustine's own treatment of time and its measurement.

Augustine begins from the familiar assumption that time is something we can measure—but quickly encounters a paradox. The past, he notes, no longer exists, and the future is yet to come; the present, meanwhile, has no extension. So how can we measure something absent, something that lacks extension altogether? For Wittgenstein, this perplexity arises not from the metaphysics of time, but from the grammar of the word "measure," and the conceptual picture it invites. In everyday contexts, to measure is to determine the length or size of a physical object—something extended in space, visible, and usually at rest. Time, however, is not such a thing. When we say we measure time, we are not locating the edges of a temporal object with a ruler, but performing a different kind of linguistic and practical operation altogether—one often involving memory, clocks, and conventions of reckoning.

Augustine's confusion stems from the unnoticed slippage between these distinct senses of a polysemous word. Because both are called "measurement," we are tempted to treat them as grammatically continuous. But, as Wittgenstein observes, this continuity is only apparent: the similarity between the two uses exerts a kind of conceptual magnetism that disguises the difference in their grammatical roles. The question "How can we measure time if it lacks extension?" becomes puzzling only if we assume that to measure time must function like to measure a length. It is this projection of features across language games—sustained by the polysemy of the term "measure"—that gives rise to the illusion of contradiction. Here, as elsewhere, Wittgenstein's method is not to solve the riddle with a new theory, but to untangle the linguistic knots that generated it in the first place.

It is important to note that cases like Augustine's are not merely historical curiosities; they exemplify a recurring pattern in philosophical perplexity. Because of their perceptual salience and practical familiarity, concrete things and actions often serve as source domains for metaphorical and analogical extension. Words grounded in bodily experience—like “measure,” “in,” or “object”—are especially prone to this drift. As their usage migrates into more abstract or theoretical domains, their meanings may diversify without becoming visibly distinct. This gives rise to polysemy: a silent sedimentation of divergent inferential roles beneath a common lexical form. In the case of “measure,” the grammar appropriate to spatial objects is subtly imposed upon temporal experience. Such unnoticed shifts often go unchallenged precisely because they masquerade as semantic continuity. It is at these fault lines—where grammatical projection, craving for conceptual unity, and polysemous drift converge—that some of philosophy's deepest confusions take hold. Recognizing and defusing them, as Wittgenstein suggests, is not a matter of resolving ancient paradoxes, but of learning to see how language quietly leads us astray.

3. Conclusion

My aim here was to foreground an underexplored source of philosophical confusion in Wittgenstein's later work: polysemy. While much attention has been given to the projections between language games, analogies, and philosophical pictures, I have argued that these are often sustained by a deeper semantic mechanism—namely, the multiple, divergent uses embedded within a single linguistic form. Polysemy gives rise to the illusion of conceptual unity while concealing subtle shifts in meaning and inferential role. In philosophical discourse, where precision is expected, such shifts can go unnoticed, paving the way for pseudo-problems and category mistakes. The cases of thinking's “location” and time's “measurement,” as examined in the *Blue Book*, illustrates how ordinary terms can silently carry the residue of metaphor and grammatical drift, sustaining longstanding philosophical perplexities. By drawing attention to polysemy as a structural enabler of these confusions, we add a new dimension to the therapeutic ambitions of Wittgenstein's later philosophy. Several other central philosophical terms—such as “object,” “truth,” or “identity”—may be similarly affected, and warrant further grammatical attention.

References

- Baker, G. P.; Hacker, P. M. S (1980/2005) *Wittgenstein: Understanding and Meaning, Volume 1 of An Analytical Commentary on the Philosophical Investigations. Part I: Essays*, 2nd ed., Oxford: Wiley-Blackwell.
- Glock, Hans-Johann (1996) *A Wittgenstein Dictionary*, Cambridge, MA: Wiley-Blackwell.
- Lakoff, George (1987) *Women, Fire, and Dangerous Things: What Categories Reveal about the Mind*, Chicago: University of Chicago Press.
- Rosch, Eleanor; Mervis, Carolyn (1975) "Family resemblances: Studies in the internal structure of categories", *Cognitive Psychology* 7 (4), 573–605.
- Wittgenstein, Ludwig (1953) *Philosophical Investigations*, G. E. M. Anscombe (ed.), New York: Wiley-Blackwell. [PI]
- Wittgenstein, Ludwig (1958) *The Blue and Brown Books: Preliminary Studies for the 'Philosophical Investigations'*, Rush Rhees (ed.), Oxford: Harper & Row. [BBB]
- Wittgenstein, Ludwig (2000) *Wittgenstein's Nachlass. The Bergen Electronic Edition*. Edited by the Wittgenstein Archives at the University of Bergen under the direction of Claus Huitfeldt. Oxford: Oxford University Press. [TS/MS]
- Wittgenstein, Ludwig (2016) *Wittgenstein: Lectures, Cambridge 1930–1933: From the Notes of G. E. Moore*, David Stern, Brian Rogers, and Gabriel Citron (eds.), Cambridge: Cambridge University Press. [MWL]

White Fraternities: A Tale of Two Smotherings

William Moix (Binghamton, NY, United States)

Abstract

I argue that distinctively raced institutions - *white fraternities* are the sites of dangerous self-censorship that are enabled by and further perpetuate pernicious ignorance which harms the subjects of their ignorance (i.e. women and persons of color). I conceptualize this phenomenon of self-censorship through Kristie Dotson's *testimonial smothering*. It will follow that due to an internal instance of this phenomenon – what I call *localized testimonial smothering* – possible insider/outsider agents of change will be unable to express dissent, effectively curtailing socially just practices in these institutions. As members of these fraternities have a high potentiality to assume powerful social roles post-graduation, they bring with them the same pernicious ignorance into those socially salient roles. This results in women and persons of color being incompetently understood in their appeals to those who have the power to make notable changes or to cast judgment in an impactful way. Thus, these marginalized social identities self-censor to avoid their misunderstood testimonies being used to their detriment. Drawing from Dotson again, this latter phenomenon curtails just social practices on a broader scale than that of its localized counterpart – Thus, I call *societal testimonial smothering*. By conceptualizing the interaction of these two distinct instances of self-censorship – *localized testimonial smothering* and *societal testimonial smothering* – I argue that the former results in the latter, and vice versa, creating a feedback loop. These problematic epistemic practices, left unchecked, will perpetuate a pernicious, systematic ignorance of marginalized experience.

Fraternities have long been prominent sites of collegiate social life and have offered professional opportunities via networking post-graduation. Boasting a pedigree of esteemed alumni who have gone on to hold salient positions of power, fraternities have often been cited as social institutions that remain largely resistant to integration and engage in heinous acts against groups outside of their own (Clowney 2022, 1). Particularly heinous are instances of sexual assault against women and overtly racist attitudes held against persons of color (Black, Belknap, and Ginsburg 2012, 400, 415; Armstrong, Hamilton, and Sweeney 2006, 492). In citing these two particular acts, a specific brand of fraternity may begin to materialize, the “white fraternity.” I posit that these particular brands of fraternity, per their white male social composition, are the sites of dangerous self-censorship that are enabled by and further perpetuate a pernicious ignorance that harms the subjects of their ignorance (i.e. women and persons of color). Members of fraternities have a high potentiality to assume powerful social roles post-graduation, bringing with them the same pernicious ignorance into those socially salient roles. This

results in women and persons of color being incompetently understood in their appeals to those who have the power to make notable changes or to cast judgment in an impactful way. Thus, these marginalized social identities self-censor to avoid their misunderstood testimonies being used to their detriment.

It will be my goal in this paper to identify this progression from what I call a *localized testimonial smothering* to a *societal testimonial smothering*. More clearly, this localized testimonial smothering is located in the *localized* social space shared between fraternity members; and, the societal testimonial smothering is located within the broader social polity shared by those now-graduated members and those women and persons of color whom they marginalize. In section 1, I will outline the theoretical framework in which I am situating my discussion, Kristie Dotson's account of testimonial smothering. In section 2, I will rely on Jan Willem Wieland's account of willful ignorance to explain the brand of pernicious ignorance that plagues the fraternities in question. In section 3, I will quickly sketch the brand of fraternity – "white fraternity" – that will be the focus of my discussion. With this established context, I will argue, in section 4, for a two-fold progressive account of testimonial smothering that begins at the localized level, which then *progresses* and influences testimonial smothering at a broader, societal level that I discuss in section 5. If my argument holds, I will have shown that women and persons of color self-censor in part because they find their testimonies fall on incompetent ears; and, these incompetent ears result from fraternal brotherhoods that nurture and instill a pernicious ignorance.

Section 1: Testimonial Smothering

Testimonial smothering occurs when the speaker of an utterance "perceives their immediate audience as unwilling or unable" to appropriately uptake their testimony (Dotson 2011, 244). As a result, the speaker abbreviates their testimony to only offer information that the hearer competently understands. It will become apparent later on that, especially in my discussion of localized testimonial smothering, this abbreviation can result in a complete lack of communication between a speaker and interlocutor.

Dotson offers three routinely existing circumstances that characterize instances of testimonial smothering to provide a schematic for identifying them:

(1) The content of the testimony must be unsafe and risky

Unsafe and risky content in testimony is anything that could run the risk of being misunderstood, due to unintelligibility on the part of the hearer, and forming false beliefs to the detriment of the speaker. These false beliefs of concern here are those that could cause “social, political, and/or material harm” (Dotson 2011, 244).

(2) The audience must demonstrate incompetence with respect to the content of the testimony of the speaker

The audience demonstrates competence when, in regard to the speaker’s testimony, they find it accurately intelligible, meaning that they can “clearly comprehend the testimony and [...] would be able to detect possible inaccuracies in their comprehension” (Dotson 2011, 245). Inversely, incompetence is the deprivation of this comprehension and the inability to detect inaccuracies in one’s faulty comprehension.

(3) Testimonial incompetence must follow, or appear to follow, from pernicious ignorance (Dotson 2011, 244).

Pernicious ignorance is any sort of ignorance, in a given context, that is consistent with or follows from a predictable epistemic gap in cognitive resources (Dotson 2011, 238).

Schematizing testimonial smothering with these three common circumstances will prove vital later when mapping out localized and societal accounts of testimonial smothering. To review, testimonial smothering occurs when the speaker deems their testimony as unsafe and risky. This perception of risk comes from an audience demonstrating an inability to accurately comprehend their testimony, presenting an opportunity for false interpretations of testimony. This inability to correctly comprehend testimony comes from a consistent epistemic gap in cognitive resources.

It would seem a logical next step, then, to describe what sort of ignorance is pervasive amongst our fraternal audience to explain their incompetence towards speakers’ testimony inside their ranks and beyond. I believe Jan Willem Wieland’s account of willful ignorance to be especially apt for this conversation. As will become apparent in later sections, fraternities, and their

members, have a vested interest rooted in loyalty and personal convenience which breed potent instances of willful ignorance.

Section 2: Willful Ignorance

Willful ignorance is commonly seen to come from an individual's will to be so ignorant. Rather than being kept from knowledge, it seems to be the case that one keeps themselves from it because it is more convenient for them. In other words, willful ignorance seems to posit that one keeps themselves ignorant because doing otherwise would make *them* worse off, despite how easy it may be. Wieland points out that, while this simple account is generally indicative of willful ignorance, it sheds no real light on what this convenience entails and the serious ramifications retaining convenience has. By this simple account, the fact that it would be more convenient for me to not know how tall Mount Everest is, despite how easily I could google it, would be a case of willful ignorance. Clearly, this instance of willful ignorance has no serious ramifications, other than in a sudden-death game of geographical trivia.

In light of accounting for the force of convenience, Wieland builds from Moody-Adams' earlier conditions of willful ignorance to present us with a schematized account which we will later apply for our purposes:

S's ignorance of p is willful if

- (i) p implies that A, an action of S or another agent, is wrong;
- (ii) S should have considered p;
- (iii) S could have considered p;
- (iv) But S does not consider p;
- (v) Because this is inconvenient for S (Wieland 2017, 111).

Using one of Wieland's examples, borrowed from Moody-Adams, we can apply this framework to the case of a university administrator.

Consider a university administrator who is confronted with allegations of harassment against one of their colleagues. In this scenario, the administrator refuses to investigate, becoming willfully ignorant of the fact that their colleague has harassed someone. Applied to Wieland's framework, (i) the

proposition “there should be no harassment in schools” implies that the harassment by their colleague is wrong and, further, that doing nothing to stop it is wrong. As such, (ii) the university admin should have known about the harassment and that this situation should be rectified. Further, (iii) the admin could have known about the harassment. The point is that it was well within the admin’s ability to investigate the harassment and be aware of the facts. But, (iv) the university admin does not investigate the harassment. This is because (v) it is inconvenient to investigate a colleague and find them guilty of harassment (Wieland 2017, 109–10).

Condition (v) requires a further mechanism to explain why it is inconvenient for the admin to investigate their colleague: *Backward-looking self-interest and other-interest*. Roughly, backward-looking self-interest is concerned with an agent wanting to stay ignorant about past behavior. Other-interest refers to an agent caring for another agent and, as such, does not want to know that they would do wrongful things. Applied to our case here, it seems then that it could be in either sense. The university admin could care about the school’s image, an extension of them, and demonstrate a backward-looking self-interest. Alternatively, their image could not matter much to them but, rather, the loyalty between colleagues could be a motivating force. Either way, investigating a colleague is rendered inconvenient and, thus, the university admin remains willfully ignorant (Wieland 2017, 110–11).

Now that we with our conceptual tools in hand – Weiland providing willful ignorance as the driving force of Dotson’s testimonial smothering – it may seem time to analyze the powerful social institutions that are fraternities. Yet, it would be important to take one last pit stop to outline the brand of fraternity to which we will be applying our tools, white fraternities.

Section 3: White Fraternities

Fraternities are commonly known as existing on college or university campuses within a larger Greek system which is split by gender, sororities standing their opposite. Fraternities and sororities can take on different racial, ethnic, and, as of recently, gender/sexual-oriented flavors. Despite this apparent diversity among Greek organizations, there stands the paradigmatic fraternities and sororities at the top, those dominated by white “brothers” and “sisters” (Clowney 2022, 18–19).

Clowney, in the first wide-scale compositional examination of American Greek fraternities, notes that colleges such as the University of Arkansas – which remained entirely segregated until 2004 – and the University of Alabama – which did not fully integrate until 2014 – are only a small piece of the Greek system which remains a “site of fierce resistance to integration” (Clowney 2022, 1–2). An important point of note in why segregation amongst Greek organizations matters is the social power that they have on campuses. Further, these same fraternities that remain largely segregated are important points of access to the broader job market beyond college. Thus, those who are allowed access to these socially privileged groups on campus are allowed a fast track to socially salient positions of power post-college.

The social power wielded on campus by fraternities is no laughing matter. Clowney cites the words of a Cornell student who states that “If you don’t plan on being in a frat or sorority [...] it can affect your ability to make friends. You will be excluded from a lot of events and things” (Clowney 2022, 10). As college is an unfamiliar space with possibly an overwhelming amount of people, any limitations or barriers presented in the way of making friends and finding familiar social spaces could be disastrous for an incoming freshman. These fraternal spaces while offering undergraduate comfort also offer postgraduate comfort. As Clowney found, “joining a fraternity produces large gains in social capital and increases in expected future earnings by approximately 36%” (Clowney 2022, 15). It can be assumed that along with these higher earnings come jobs with equal parts prestige, or more social weight. Thus, those who are allowed access to these groups are allowed eased access into influential positions post-graduation.

The stark numbers that Clowney presents show that prominent schools’ fraternities and sororities around the country are dominated by white members. Auburn, Ole Miss, Nebraska, Alabama, and Clemson are made up of above ninety percent white members while five other top institutions are above the eight-fifth percentile (Clowney 2022, 21). In contrast, black students make up less than 1% of those members in sororities and 1.6% of men in those fraternities despite making up 12% of all undergraduates in the United States (Clowney 2022, 24).

The purpose of this sketching of the Greek system in America is to paint a picture of those who are allowed the most access to the social rewards afforded by fraternities and sororities while in college and afterward. While not described in detail here, I take for granted these white fraternities provide positions of social superiority and pull relative to white sororities. An example is when white sorority sisters in some houses only gain access to live within their house once they have accumulated a number of points that are gained by attending fraternity house parties. This example fits well into the assumption that - even while in a privileged position qua sorority - those sorority members act in response to their relatively privileged correlates, fraternities. It is for this reason that I am focusing my discussion on fraternities due to their socially privileged statuses relative to those of women, both in sororities and beyond, and people of color, who are left excluded from these groups.

Section 4: Localized Testimonial Smothering

Now that I have outlined the social institution that I will analyze – white fraternities – and provided the tools of analysis – Dotson’s testimonial smothering and Wieland’s willful ignorance – I will argue for my two-fold progressive account of testimonial smothering that begins among those white fraternity social spaces and then influences a societal testimonial smothering when those members assume positions of power in society-at-large.

Applying Dotson’s account of testimonial smothering, individual fraternity members may find the content of such utterances as “I think that Smith sexually assaulted someone” or “Hey, Jones has been saying a lot of racist comments to in the group chat” to be unsafe and risky to their credibility in their fraternities. The perceived risk finds purchase in the earlier note about how white fraternities come with a powerful amount of social status within and outside of the collegiate space. If one were ostracized or not given a recommendation from their fraternity because of a loss of social status, this could be seen as a palpable risk qua fraternity member. Per Dotson’s accounts, this could be enough reason for one to self-censor the content of these testimonies.

The perception of risk and a lack of safety comes from the testimonial incompetence demonstrated by those that they attempt to report to. This testimonial incompetence is demonstrated, for instance, by other fraternity

members dismissing the story or offering less damaging conceptions of it. Notably, they do not fully comprehend the force of the testimony. This testimonial incompetence appears to follow from a pernicious ignorance, namely willful ignorance.

This willful ignorance, afforded by Wieland, explains that other loyal fraternity members have a backward-looking self-interest and an other-interest which labels the acknowledgment of sexual misconduct and racist transgressions in the fraternity as inconvenient. The former interest manifests from the members not wishing for their fraternity, and their credibility as a part of it, to be discredited by charges of sexual misconduct and racism. The latter interests manifest from loyalty to their fellow members, whom Smith and Jones are a part of, wishing to protect them. In both senses, those confronted by such testimonies and demonstrate testimonial incompetence do so because of the inconvenience of acknowledging the facts. It is not because they could not have confirmed such allegations or that they should not have but rather that such information would present an inconvenience that could threaten the fraternity if acknowledged.

Thus, to retrace this logical line, fraternity members who wish to testify against racial transgressions or sexual misconduct against their fellow members may typically find that their testimonies are incompetently understood by those in the fraternity to whom they try to report it. I argue that this incompetence comes from a willful ignorance that is vested in the interests of the fraternity. As a consequence, those harmed by the fraternity – cited by the testimonies – are ignored and left open for repeated harm. Because of the testimonial incompetence shown by fellow fraternity members, those who originally wished to report transgressions self-censor to retain the favor of the fraternity they are a part of.

Holding that this localized testimonial account is plausible, I will now argue, quickly, that the professional launchpad that is a fraternity's networking structure does not just bring fraternity members into powerful social positions. Further, they bring along situated knowledge and willful ignorance, aided by earlier localized testimonial smothering, into a larger social polity that influences the further testimonial smothering of those subjects of earlier testimony – women and persons of color.

Section 5: Societal Testimonial Smothering

The term societal testimonial smothering comes integrally tied to the localized testimonial smothering I just outlined. What I attempt to explain by this two-fold relation is that the considerations/experiences of women and persons of color, previously subjected to ignorance in the fraternal space, are directly affected in broader social context post-graduation because now-graduated fraternity members assume positions of great power in the large social polity qua fraternal networking. Societal testimonial smothering follows as such.

Women and persons of color regularly find that they are unrepresented, underpaid, and subjected to adjacent structural injustices in the broader social landscape. Per the influence of fraternal networking, the persons who hear these testimonies or have an impactful say in making change are men raised through white fraternal structures. As a consequence of localized testimonial smothering, the claims of women and persons of color fall on incompetent ears that were insulated to filter out the claims of such individuals due to faulty epistemic practices found in fraternities. This broader testimonial incompetence appears to follow from the same willful ignorance that ignored the testimonies about the harm done to those marginalized social identities. Thus, a consistent epistemic gap regarding the trials faced by those persons in a broader social sense becomes manifest. As a result, women and persons of color self-censor their testimonies addressing similar harms because they perceive such testimony as risky and unsafe. This perceived lack of safety lies in the consistent formation of false beliefs that occur from misunderstood testimony, and the potential social, political, and/or material harm that could come from it. Consequently, this self-censorship on a broader scale creates a feedback loop where men nestled in these brotherhoods continue to lack the proper acknowledgment of the considerations/experiences of women and person of color and are continually thrust into socially powerful positions beyond college per there fraternal affiliations; thus continuing the cycle.

References

- Armstrong, Elizabeth A., Laura Hamilton, and Brian Sweeney. 2006. "Sexual Assault on Campus: A Multilevel, Integrative Approach to Party Rape." *Social Problems* 53 (4): 485–92. <https://doi.org/10.1525/sp.2006.53.4.483>.
- Black, Tyra, Joanne Belknap, and Jennifer Ginsburg. 2012. "Racism, Sexism, and Aggression." In *African American Fraternities and Sororities*, edited by Tamara L. Brown, Gregory S. Parks,

and Clarenda M. Phillips, 2nd ed., 400–418. *The Legacy and the Vision*. University Press of Kentucky. <http://www.jstor.org/stable/j.ctt2jcq14.22>.

Clowney, Stephen. 2022. “Do Fraternities Violate the Fair Housing Act? An Empirical Study of Segregation in the Greek Organizations.” *SSRN Electronic Journal*, 1–28. <https://doi.org/10.2139/ssrn.4065056>.

Dotson, Kristie. 2011. “Tracking Epistemic Violence, Tracking Practices of Silencing.” *Hypatia* 26 (2): 238–45. <https://doi.org/10.1111/j.1527-2001.2011.01177.x>.

“Greek Political Leaders - NIC | North American Interfraternity Conference.” n.d. Org. North American Interfraternity Council. Accessed December 6, 2022. <https://nicfraternity.org/greek-political-leaders/>.

Wieland, Jan Willem. 2017. “Willful Ignorance.” *Ethical Theory and Moral Practice* 20 (1): 109–11.

What If Nothing Were the Case? On Empty Space in Wittgenstein's *Tractatus*

Filippo Mosca (Rome, Italy)

Abstract

In section 2.013 of the *Tractatus Logico-Philosophicus*, Wittgenstein claims that “every thing is, as it were, in a space of possible states-of-things” and that “this space” can be thought of “as empty.” This echoes Kant’s assertion in the *Critique of Pure Reason* that one can “very well think that there are no objects to be encountered” in space. Despite the resemblance, Wittgenstein’s notion of empty space differs sharply from Kant’s. This paper develops along three lines. First, it examines the status of empty space within the Tractarian framework in contrast to Kant: unlike Kant’s physical-geometric space without objects, Wittgenstein’s empty space is the configuration of logical space in which nothing is the case. Second, it addresses the apparent paradox of thinking of space as empty (the very thought that nothing is the case constitutes something that is the case) by appealing to the distinction between factual and logical incompatibility. Third, it traces a plausible source for this distinction in Wittgenstein’s wartime *Notebooks*, where he challenges Russell’s view on fully generalised propositions, and moves from treating them as tautologies to recognising their contingent, non-logical character.

1.

In 2.013 of the *Tractatus logico-philosophicus*, Ludwig Wittgenstein states:

Every thing is, as it were, in a space of possible states-of-things. I can think of this space as empty, but not of the thing without the space. (2.013)

Wittgenstein’s remark seems to closely echo the following passage from Immanuel Kant’s *Critique of Pure Reason*:

One can never represent that there is no space, though one can very well think that there are no objects to be encountered in it. (A 24)

The affinity is striking (Several scholars have pointed it out. See, for instance: Black, 1964: 50, Engel, 1970: 490, Morris, 2008: 37, Klaassen, 2015: 14, Zambito 2019: 49). However, as we shall see, when Kant and Wittgenstein invoke the idea of empty space, they have a different notion of emptiness in mind. This paper develops along three lines: (a) it considers, in contrast to Kant, how space can be (thought as) empty according to the Tractarian framework [2–3]; (b) it explores the apparent paradox involved in *thinking* of space as *empty* [4–5]; (c) it shows that Wittgenstein had already laid the groundwork for dissolving this apparent paradox in his wartime *Notebooks*.

2.

In the passage from the *Critique of Pure Reason* to which Wittgenstein seems to be alluding, Kant clarifies his account of space from the *Transcendental Aesthetic*. Space is not an absolute reality independent of us, but a subjective representation. However, *subjective* does not mean *arbitrary*. On the contrary, although space is represented by the subject, this representation is *necessary*. Why? Because it is an *a priori* representation – that is, it necessarily precedes any subject's experience.

Space is a necessary representation, a priori, that is the ground of all outer intuitions. (A 24)

We do not derive space from experience, as if objects merely happened to be spatial and we inferred their spatiality from encountering multiple instances of them. Rather, space is presupposed a priori by any experience of objects.

As a thought experiment, we can try to imagine space as empty – that is, without objects. But we cannot do without space. To have representations without space would mean abandoning (one of) the very form(s) that makes representation possible, and thus, in effect, ceasing to have any representation. But there is no point in talking about objects in abstraction from our capacity to represent them (B XXVI). Consequently, to cease to have any representation would mean to cease to have any *object* at all.

3.

Wittgenstein's remark in 2.013 closely resembles Kant's wording, but their points diverge significantly. Two key differences are worth highlighting.

The first key difference concerns the kind of space at stake. The space in 2.013 is not Kant's physical-geometric space. It is "space" only figuratively (*gleichsam*): it consists not of locations, but of possibilities – of any kind. (Physical-geometric space is not entirely absent from the *Tractatus*. See, for instance, 2.0121, 2.0131, 2.01251. However, when Wittgenstein refers to it, space appears merely as one example among others – typically alongside time and colour. Spatial possibilities are thus just one subset within the broader space evoked in 2.013.) Such space is what Wittgenstein elsewhere calls *logical space* (1.13). This space is *logical* in virtue of its connection with thinking and possibility: it "contains" everything that is possible to think (a thought is a

logical picture: 3), and therefore (see Section 4) everything that is possible (3.02, 5.61).

The second key difference concerns how space can be empty according to Wittgenstein.

Kant implies that while we cannot imagine objects without space, *we can imagine space without objects*. Wittgenstein, however, would likely reject this as a genuine possibility. His “objects” are quite different from ordinary objects. They are the indivisible elements of “the substance of the world” (2.021). While the world “is everything that is the case” (1), substance “subsists independently of what is the case” (2.024) – hence, objects exist *necessarily*. Indeed, their necessary existence is an indispensable condition for the meaningfulness of our propositions (2.021, 2.0211 and 2.0212).

It is obvious that an imagined world, however different from the real world it may be, must have something – a form – in common with the real world. (2.022)

This fixed form consists of the objects (2.023)

Possible worlds differ in so far as they consist of different *configurations* of objects (2.0271 and 2.0272). Objects, however, necessarily exist in every possible world we can imagine. So, we cannot imagine anything that does not involve objects. (I use “imagine” and “think” interchangeably – both because “imagine” is often used to translate *denken* in English versions of the *Tractatus*, and because the *Tractatus* makes no room for *imagining* as distinct from *depicting*, and thus from *thinking*. See Section 4.)

Wittgenstein’s notion of emptiness thus differs markedly from Kant’s. When Wittgenstein states, “I can think of this space as empty,” he does not mean ‘empty of objects’, as Kant might suggest. Rather, Wittgenstein means, ‘empty of states-of-things.’

States-of-things are the atomic combinations of objects that constitute complex facts (2), which in turn make up the world (1.1). Unlike facts, states-of-things are *immediate* combinations of objects (2.01). (“What is the difference between *Tatsache* [fact] and *Sachverhalt* [state-of-things]? *Sachverhalt* is, what corresponds to an *Elementarsatz* if it is true. *Tatsache* is, what corresponds to

the logical product of elementary prop[osition]s when this product is true [...]" – LW to BR, 19.8.19.) Like facts, they are *contingent* combinations. Insofar as it is "[w]hat is the case" (2), a fact is also, *eo ipso*, what might not have been the case. But the same applies to states-of-things: they may obtain or fail to obtain (2.05, 2.06, 2.062).

The non-obtaining of states-of-things is perfectly conceivable. One can even conceive of the non-obtaining of *all* states-of-things. In that case, the "space of possible states-of-things" would be empty – none of the possible states-of-things would transition from being merely possible to obtaining. Yet its conceivability would remain unaffected. However empty, this space could still be imagined.

4.

Following Page (1997), we can think of *logical space* as the total set of elementary propositions – each depicting a possible state-of-things, along with all possible truth-value assignments over them. This complete "list plus truth table" schema yields all the ways the world might be: that is, logical space.

Following Cerezo (2005), I take the idea that *logical space is empty* to be captured by the truth-function φ , which yields F for all rows in the truth-table of all elementary propositions except for the all-F row.

p₁	p₂	p₃	.	p_n	φ
T	T	T	.	T	F
T	T	T	.	F	F
.	.	.	.	.	.
.	.	.	.	.	.
F	F	F	.	T	F
F	F	F	.	F	T

φ represents the possible world in which none of the objects that constitute the substance of the world stand in any relation to one another. In short, it represents logical space as empty: it is true that nothing is the case.

As we have seen, 2.013 implies that φ is thinkable. But what does it mean, properly speaking, to think that φ ? To answer this, we must first consider what a *thought* is in the *Tractatus*.

A logical picture of facts is a thought. (3)

A thought is a *logical picture*. This, in turn, implies that a thought is a *picture*. In fact, as a logical picture, a thought may be regarded as the paradigm of pictorial representation. As Wittgenstein states:

Every picture is *also* a logical picture. (On the other hand, not every picture is a spatial picture, for example) (2.182).

But if a thought is a picture, then a thought is also a *fact*. Indeed, we make ourselves pictures of facts (2.1) by means of those peculiar *facts* which are pictures.

A picture is a fact (2.141)

A thought, then, is a fact – that is, something that is the case (2). But if a thought is something that is the case, then the proposition “ φ ” expressing φ – i.e. “(it is true that) nothing is the case” – *can never be true*. For a proposition is simply the perceptible manifestation of a thought (3.1, 3.11, 3.12), and the thought that nothing is the case *never happens to be true*. Indeed, in thinking that nothing is the case, there would at least be something that *is* the case. That something is *the very thought* that nothing is the case.

One might object that the reference to “thinking” is metaphorical. Perhaps what Wittgenstein means is simply that ‘*it is possible* for logical space to be empty’, and not that we can literally *think* it. In that case, the emptiness of logical space would not depend on the occurrence of a thought thinking it, and “ φ ” might indeed be true.

But this option is ruled out within the Tractarian framework. “What is thinkable is also possible” (3.02). Conversely, what is possible is also thinkable. There is no such thing as a *possible but unthinkable* situation. Hence, to say that

‘it is possible for logical space to be empty’ is just to say that ‘we can think of logical space as empty.’ In both cases, the proposition expressing that emptiness never actually turns out to be true.

5.

How should we understand 2.013, if “ φ ” can never be true? How can we be said to think the situation in which nothing is the case, if the very act of thinking it entails that something is the case?

The answer lies in the nature of φ ’s impossibility of being true. This impossibility is not logical, but *factual*. That is, φ is not a contradiction. To paraphrase Wittgenstein’s remark in 6.1232, the general invalidity of φ expresses only an *accidental*, not an essential, general falsity.

Like φ , a contradiction cannot be true (4.64). However, there are two crucial differences between the impossibility of φ and that of a contradiction.

First, the impossibility of a contradiction being true is revealed in the formal structure of contradictory propositions (4.1211). The truth-conditions of, say, “ $p \cdot \sim p$ ” are mutually exclusive (4.46). The specific content of the proposition is irrelevant – only form matters. By contrast, the impossibility of φ ’s being true stems not from logical form but from the incompatibility of two specific situations: the situation in which nothing is the case, and the situation in which a thought occurs.

Second, a contradiction fails to represent anything at all. I have just said that in contradictions the content expressed is irrelevant. But it would be more accurate to say that no content is expressed at all. Though part of the symbolism (and thus not nonsensical, but merely senseless) (4.461 and 4.4611), contradictions *are not pictures of reality* (4.462). Like tautologies, they say nothing. (4.461). By contrast, the proposition expressing φ is a picture of reality. Though never true, it still says something. It has sense because it could, *in principle*, be true.

To understand a proposition is to know what is the case if it is true. [...] (4.024)

We understand “ φ ” because we grasp what would be the case *if it were true*, even if the situation it describes can never *de facto* obtain.

6.

The notion of empty space in 2.013 highlights a fundamental distinction between a logical incompatibility and an incompatibility among facts, however *general*. The former is unthinkable: “[w]e cannot think anything illogical [...]” (3.03). The latter, though *always* false, remains thinkable.

In these final sections, I aim to identify one plausible source for this distinction in a set of entries from the *Notebooks*, where Wittgenstein implicitly questions Russell’s account of fully generalised propositions.

In *Theory of Knowledge* (1984) Russell famously introduces *forms*. According to Russell, a form is

the fact that there are entities that make up complexes having the form in question. [...] For example, the form of all subject-predicate complexes will be the fact “something has some predicate”; the form of all dual complexes will be “something has some relation to something”. (Russell, 1984: 114)

Forms are facts like those expressed by “ $(\exists x) F(x)$ ” or “ $(\exists x, y) xRy$ ”, and so on. These facts are unusual in two respects: they are extremely general, since every instance of judgment presupposes one; the propositions expressing them are always true, as there is always something that has some predicate, or something that has some relation to something, and so on.

Crucially, Russell seems to argue that since they are always true, they are *logical* truths.

The dualism of true and false, with all its attendant distinctions, presupposes propositions, and does not arise so long as we confine ourselves to acquaintance, except, possibly, in the case of abstract logical forms; and even here there is no proper dualism, since falsehood is logically impossible in these cases. (1984: 141)

“ $(\exists x, \varphi) \varphi x$ ” or “ $(\exists x, y) xRy$ ” count as *logical* (sic) forms because, in these cases, falsity is *logically* (sic) excluded. If the facts they express did not obtain, it would be impossible to make judgments exhibiting the corresponding logical structure. They provide the “foundations of the factual discourse” (Pears, 1987: 129).

7.

Wittgenstein's *Notebooks* show that he, too, was initially inclined to assign fully general propositions a special status. He describes such propositions as, in a sense, *isolated from the world*. In entries from 14–15 October 1914, he writes:

It is clear that we can form all the completely general propositions that are possible at all as soon as we are merely given *a language*. And that is why it is scarcely credible that such connexions of signs should really say anything about the world. [...] (Ms-101,43r[2])

The proposition is supposed to give a logical model of a situation. It can surely only do this, however, because objects have been arbitrarily correlated with its elements. Now if this is not the case in the quite general proposition, then it is difficult to see how it should represent anything outside itself. (Ms-101,44r[4]et45r[1])

Here I regarded the relations of the elements of the proposition to their meanings as feelers, so to say, by means of which the proposition is in contact with the outer world; and the generalization of a proposition is in that case like the drawing in of feelers; until finally the completely general proposition is quite isolated. [...] (Ms-101,46r[2])

Fully generalised propositions do not require external verification. They are true *in isolation from* what is the case. Yet the entry of 15 October ends with a doubt:

[...] But is this picture right? (Do I really draw a feeler in when I say $(\exists x).\varphi x$ instead of φa ?) (Ms-101,46r[2])

This hesitation deepens the next day:

Now, however, it looks as if exactly the same grounds as those I produced to shew that " $(\exists x, \varphi) \varphi x$ " could not be false would be an argument shewing that " $\sim(\exists x, \varphi) \varphi x$ " could not be false; and here a fundamental mistake makes its appearance. For it is quite impossible to see why just the first proposition and not the second is supposed to be a tautology. [...] (Ms-101,46r[3]et47r[1])

The “fundamental mistake” lay in treating fully generalised propositions as tautologies – that is, as *logically* undeniable. Wittgenstein not only comes to see the possibility of the falsity of, say, “ $(\exists x, \varphi) \varphi x$ ”, but also recognises that, logically, he has no more reason to believe in the impossibility of its truth than he has to believe in the impossibility of its falsity – that the *negation* of “something exists and there is a predicate, and there is something that has a predicate” is a tautology.

Fully generalised propositions are not isolated from the world. Their complete generality does not render them non-material:

The generality of the completely general proposition is accidental generality. It deals with all the things that there chance to be. And that is why it is a material proposition. (Ms-101,58r[3])

Similarly, the proposition “something is the case” – which seems to follow naturally from the claim that the world is everything that is the case (1) – should also be considered *material*.

After all, the difference between “something is the case” and fully generalised propositions is a matter of degree. “Something is the case” can be understood as a higher-order generalisation of quantified existential claims such as “ $(\exists x) \varphi x$ ” or “ $(\exists x, y) xRy$ ”, and so on – just as “Such and such is the case” in 4.5 generalises over the logical prototypes in 3.315 (Cf. Bronzo, 2015: 247). Given an exhaustive list of Russellian logical forms, “something is the case” would represent the most abstract formulation derivable from them.

So what holds for “ $(\exists x, \varphi) \varphi x$ ” also holds for our “ φ ”. It is quite impossible to see why “something is the case” – and not “nothing is the case” – should count as a tautology. In fact, “ φ ” is not a tautology. It asserts something *about the world*. As such, its negation, however factually unobtainable, must be logically possible – that is, thinkable.

I can think of a possible world in which everything, including my own thinking, is “absorbed” into an empty space where nothing is the case. The world is everything that is the case (1). Yet, it is perfectly possible (to think) that *nothing is the case* – that there is no world at all.

References

- Black, M. (1964). *A companion to Wittgenstein's Tractatus*. CUP Archive.
- Bronzo, S. (2014). *Words, Sentences, and Speech Acts: An Interpretation and Defense of the Context Principle* (Doctoral dissertation, Doctoral Dissertation to be submitted to the Department of Philosophy of the University of Chicago in).
- Cerezo, M. (2012). "Possibility and logical space in the *Tractatus*". *International Journal of Philosophical Studies*, 20(5), 645-659.
- Engel, S. M. (1970). "Wittgenstein and Kant". *Philosophy and Phenomenological Research*, 30(4), 483-513
- Kant, I. (1998). *Critique of pure reason* (P. Guyer & A. W. Wood, Trans.). Cambridge University Press. (Original work published 1781/1787)
- Klaassen, T. (2015). "Wittgenstein as a Kantian Philosopher". URL: <https://philpapers.org/archive/KLAWAA.pdf>.
- Halimi, B. (2016). "The Concept of "Essential" General Validity in Wittgenstein's *Tractatus*". *Early Analytic Philosophy-New Perspectives on the Tradition*, 283-300.
- Morris, M. (2008). *Routledge philosophy guidebook to Wittgenstein and the Tractatus*. Routledge.
- Page, J. (1997). "Unconfigured Tractarian Objects". *Philosophical Investigations*, 20(1), 39-50.
- Pears, D. (1987). *The False Prison: A Study of the Development of Wittgenstein's Philosophy*. Volume 1. Clarendon Press.
- Wittgenstein, L. (1921), *Logisch-philosophische Abhandlung*. In: "Annalen der Naturphilosophie", 14, 185–262. Second Ed.: *Tractatus logico-philosophicus*. London, 1922.
- Wittgenstein, L. (2000), *Wittgenstein's Nachlass: The Bergen Electronic Edition*, ed. Claus Huitfeldt, Oxford, Oxford University Press.
- Zambito, P. F. (2019). "Logic is a geometry of thinking". *Space and Spatial Frameworks in Wittgenstein's Writings* (Doctoral dissertation).

Acknowledgement

Thanks to Nicola Gianola for the nice and painful discussions.

Epistemic and Affective Advantages of Being Oppressed

Carmen Mossner (Osnabrück, Germany)

Abstract

Standpoint theorists claim that being oppressed, although characterised by disadvantages, can also offer a position of epistemic privilege, where, for instance, experiences of oppression can provide insights into the dynamics of oppressive structures. However, the literature has so far only focused on the epistemic dimension, treating affective experiences as mere contributors to knowledge about oppression. I challenge this narrow view by arguing that affective experiences of oppression harbour *affective advantages* that are *not reducible* to epistemic knowledge. In particular, I suggest that oppressed individuals experience a greater variety of emotions linked to oppression, while also experiencing them with a greater intensity. This can, for instance, foster a heightened sensitivity to structural changes, facilitate the development of tools for emotion regulation, and expand their affective repertoire in ways that enhances their empathic abilities. I further illustrate that these affective advantages are not merely a contributing factor to epistemic insights but hold an intrinsic value in themselves that are hardly accessible to non-oppressed groups. Crucially, recognising the value of affective experiences of oppressed individuals can aid in implementing a more just society.

1 Introduction

Oppression is commonly associated with structural disadvantages for certain individuals or groups. Women are oppressed as women through norms and social structures that downgrade feminine appearances, abilities, and character traits in ways that make women systematically more prone to experiencing physical harm, to being disadvantaged on the job market or to being sexualised and objectified. Along similar lines, people of colour are disadvantaged, so are disabled people, homosexual people, trans people, and so on. Most of the debates on oppression focus on the various ways in which oppressed individuals are disadvantaged. Standpoint theorists, on the other hand, have proposed that oppressed individuals may as well experience advantages that can, on a societal level, aid in developing a more just society. They argue that oppressed people have an *epistemic advantage* (e.g., Dror 2022) or *epistemic privilege* (e.g., Fricker 1999; Toole 2024) due to their social standpoint if they become aware of and reflect upon their social position as a marginalised individual. These advantages are said to concern knowledge about power dynamics and social and economic structures that contribute to oppression, like sexism, classism, or racism (Friesen and Goldstein 2022), and maybe even moral knowledge in general (Dular 2024).

In the literature so far, quite a bit has been said on the kind of epistemic advantage that is at issue. Yet, while recent discussions on *affective injustice* have already acknowledged that the unjust treatment of individual's affective lives cannot be reduced to mere epistemic injustice (Gallegos 2021), debates on advantages due to oppression have so far failed to recognise this. Affectivity is discussed as experiences *contributing* to the epistemic advantage oppressed individuals can have, but is not regarded as an *advantage in itself*, independent of epistemic knowledge derived from experiences of oppression. However, I argue, being oppressed does not only offer a standpoint of epistemic advantage, but also one of *affective advantage*, where the affective experiences of oppressed individuals can grant them advantages, such as a sensitivity for recognising structural changes, developing tools for emotion regulation that are also of help in non-oppressive circumstances, and a greater affective repertoire that can be conducive to their empathic skills. Importantly, these advantages can benefit society as a whole, if recognised and taken as a means to identify unjust or oppressive societal structures.

Section 2 provides a brief introduction into the debate on epistemic advantage. Based on this, section 3 argues that oppression does not only offer epistemic but can also lead to affective advantages. I close in section 4 with concluding remarks.

2 Epistemic Advantages

The concept of epistemic advantage is rooted in feminist standpoint theory and the broader tradition of socially situated knowledge. It refers to the idea that individuals from marginalised or oppressed groups can, under certain conditions, develop an epistemic superiority in specific domains of knowledge concerned with social structures. This is sometimes referred to as the “inversion thesis” (e.g., Dror 2022; Wu 2023) because it challenges conventional assumptions that those in privileged positions hold superior knowledge. While privileged groups may have access to power and resources, they often lack direct insight into the structures that maintain their privilege. Oppressed individuals, by contrast, must navigate these structures daily, allowing them to attain a particular standpoint with deeper critical insight. More generally, epistemic advantages are said to comprise “identifying dubious assumptions, developing hypotheses, or promoting objectivity” (Friesen and Goldstein 2022, p. 663). These advantages primarily

concern oppressed individuals or groups, as, for example, they have been forced to develop tools to navigate their oppressive environment, potentially making them more adaptive and resilient to change. Yet, these advantages are just as important from a societal perspective. After all, if we grant that those who are oppressed have a greater sensitivity to oppressive social structures and power dynamics, they are arguably the most valuable experts in order to identify injustices and develop a more just society (ibid.).

While there is a wide consensus that being oppressed *can* offer an epistemic advantage in the social domain, many have argued for a weak version of the epistemic advantage hypothesis, claiming that an epistemic advantage does not follow automatically from oppression and is not reducible to an individual's social identity (Dular 2024). Yet, how exactly this advantage comes about, has been a matter of debate. One approach ascribes a “double vision” (Jeong 2024, p. 108) or “double consciousness” (Medina 2013, p. 44) to oppressed individuals. Often times, they are forced to not only deal with their own perspective, but also to comprehend the oppressive perspective that is imposed on them by society. The ability to compare these two visions and detect distortions or discrepancies between the two, allows them to view the world from two (or multiple) angles and identify oppressive dynamics and inequalities (ibid.).

Another explanation adverts to the need for critical reflection. Being marginalised is *per se* not enough to attain an epistemic advantage. In principle, oppressed individuals have a far greater incentive to change the structures that oppress them than individuals in power or those that benefit from the oppression of others have (Friesen and Goldstein 2022). Beyond this initial incentive, oppressed individuals must also engage in critical reflection that questions their own position in society and its potential for knowledge about oppressive structures (ibid.). This can be cultivated, among other things, through education, activism and resistance.

Yet another explanation stresses that being oppressed can foster certain epistemic virtues (Medina 2013), such as “*epistemic humility*” (p. 43) as the awareness of one's own limitations, “*intellectual curiosity*” (p. 43) as a greater desire to seek alternative explanations, and “*open-mindedness*” (p. 44) as the willingness to consider multiple perspectives. Ashton (2019) has broken down

how experiences of privilege or oppression and the corresponding development of certain vices and virtues has epistemic effects on, for instance, how likely a person is to notice or engage in friction, i.e., the contact between different perspectives. These virtues in oppressed individuals can contribute to a more rigorous and critical approach to knowledge.

Although many proponents of the epistemic advantage hypothesis recognise that the experiences of oppressed individuals are not merely of an epistemic character or simply a matter of unjust distribution of material and social goods but are instead affective phenomena as well, affectivity is always downplayed as a kind of experience that merely feeds into the epistemic knowledge of oppressed individuals. The following argues that advantages do not only concern what oppressed individuals can know in virtue of their characteristic lived experiences and what their knowledge can teach us about our society, but also what they can feel and what their feelings can teach us about our society on their own, irrespective of their merely mediating role in the process of knowledge generation.

3 Affective Advantages

Debates on oppression developed a strong focus on *epistemic* standpoints, injustices and advantages. More recent debates have applied the concept of *epistemic injustice*, i.e., the claim that oppressed individuals are sometimes discredited in their ability as *knowers* (see e.g., Fricker 2007), to the affective domain, claiming that in some cases, oppressed individuals do not (only) face an epistemic injustice but also an *affective injustice* that cannot be reduced to epistemic injustice. That is to say, they are not (only) discredited in their capacity as knowers, but (also) as emoters, for example by being denied certain emotions. Lorde (1984), for instance, describes how a white woman responded to her after she fiercely criticised the injustices that women of colour face: “Tell me how you feel but don’t say it too harshly or I cannot hear you” (p. 125). It is expected that injustices are expressed solely through emotion-deprived descriptions. Research on affective injustices points out that the significance of affectivity in oppression has not been taken serious so far. This very same gap, I argue, can be found in the literature on epistemic advantages as well. Oppressed individuals may not only inhabit a social position that provides them with an epistemic advantage, it may also offer them an *affective advantage*. The following illustrates how deeply entrenched

affective oppression is and how it, in turn, might entail affective advantages that go beyond mere epistemic ones.

Imagine scientists have developed an Empathy Transfer Device that should help reduce the cleft between oppressed and non-oppressed individuals. It allows people to temporarily experience the emotions and sensations of other people as if they were their own. Chris and Paul are the first to try the new device. While Paul has lived a privileged life, free from discrimination, Chris has experienced systemic oppression, poverty and social exclusion. The program requires Paul to live with Chris' emotions for one month. Although he retains his own memories and identity, he now feels the fear of walking through certain neighbourhoods, the anxiety of being judged based on stereotypes, and the frustration of having his intelligence questioned unfairly. At first, Paul tries to rationalize these emotions, believing they may be exaggerated. But as the month progresses, he begins to recognize the depth and diversity of the emotional burden that Chris carries, sometimes daily. Not only is he experiencing some emotions with an intensity that he has not experienced before, he is also confronted with emotions that he has never felt before. His privileged life has spared him from situations that cause feelings of existential threat, such as financial worries or threats of physical harm. He realises that all those emotions, fear, exhaustion, distrust, are deeply ingrained affective experiences that shape various aspects of life, such as behaviour and self perception. After the month in which the transfer device transferred Chris' emotions has passed, Paul is asked to describe what he has learned. He admits that no amount of theoretical knowledge, about different ways of living, about financial differences, about discrimination in general, could have prepared him for the weight of these emotions. His understanding of injustice has shifted – not because he has gained more theoretical or practical knowledge about oppression, but because he has lived, even if only temporarily, the affective reality of oppression.

The above thought experiment taps into the intuition that affective experiences can hardly be conceptualised without losing their affective weight and significance. Descriptions of oppressive structures will not be able to capture the complexity of oppression if they do not incorporate the affective experiences of those that suffer from oppressive structures. And even if descriptions try to include the affective experiences of oppressed individuals,

we may need to realise that mere descriptions cannot fully capture the range, intensity and significance of those experiences. In other words, we need to allow affectivity itself into the discourse and acknowledge, that affectivity is not merely an interesting bystander in epistemological matters but a source for advantages (and injustices) in itself. As exemplified by the thought experiment, affective experiences of oppression provide insights that cannot be captured by descriptions of or knowledge about oppression. Precisely such insights have been reported on by people like Julia Serano (2007) after transitioning from a biological male body to a female body. Remaining questions are what leads to these advantages and what they consist in.

Affective advantages stem, in part, from an *increased variety* of emotions that goes along with experiences of oppression. There is no doubt that oppression forces individuals into situations that non-oppressed individuals may never encounter. The average male worker may never encounter the financial issues of a single mother that struggles to pay for her monthly bills. He has no reason to feel the financial worry or pressure which the mother may feel daily. The heterosexual couple may never encounter the stares of disgust that homosexual couples see when holding hands in public. They have no reason to feel unwanted and excluded, or to feel the social pressure to be different and to question their own feelings that homosexual couples often face. A teenage son usually does not encounter the worried looks of his mother when leaving the house at night. He has no reason to worry much about being physically harmed or sexually harassed while his sister is frequently told to watch out at night, stay in crowded areas or better not leave the house at all after it gets dark. Overall, oppression forces people into more situations that trigger strongly negative affective responses that amount to a greater variety in felt (negative) emotions.

Not only are oppressed individuals more likely to experience a greater variety of emotions, they are also more likely to experience some emotions, those inflicted through oppression, with *greater intensity*. Although, of course, non-oppressed individuals may experience the same kind of emotion in a non-oppressive circumstance, a crucial difference is that they likely do not experience it frequently, or even daily. It is likely not an emotion that regularly accompanies them such that it significantly shapes their way of thinking, behaving, and emoting. A white woman's worry of not finding an

apartment on the market may only reflect a fraction of the worry a refugee of colour has due to constantly being rejected based on stereotypes. This does not make the suffering of non-oppressed individuals (with respect to the kind of oppression in question) any less important, but it serves to illustrate that the suffering of oppressed individuals cannot and should not be downplayed based on the fact that, in other circumstances, non-oppressed individuals experience similar emotions.

This increased variety and felt intensity of emotions may account for a few advantages. For instance, it can contribute to what Mossner and Walter (2024) called an “affective repertoire”. They argue that in order to empathise with another person, i.e., accurately *feel into* the other person’s affective state, one must have experienced a remotely similar emotion before. An affective repertoire is thus the collection of affective experiences a person has made up until now. In the context of digital communication, Mossner and Walter argue that spending large parts of one’s life in rather bleak and repetitive online environments can negatively affect the variety of one’s affective repertoire. One of its downsides is its impeding effects on empathy, as feeling into a person that experiences an emotion one has never felt before is difficult, if not impossible. Crucially, it also hinders the recognition that the other person is experiencing a potentially negative emotion in the first place. In the case of oppressed individuals, over the course of their oppression, they have become familiar with various emotions and emotional expressions in response to their oppression. They will, most likely, recognise *that* another person is feeling oppressed much faster than someone who has never been oppressed, and they will, most likely, also be able to understand and feel into the other person much better, providing a sense of understanding and community (the same phenomenon has been observed in a different context by Osler and Krueger (2022)). In other words, being oppressed can positively affect one of the pillars of empathy, i.e., the emotions one is familiar with and that one can recognise and feel into in others. This is in line with findings suggesting that oppressed individuals exhibit greater empathy than non-oppressed individuals, at least towards other marginalised people (Turkoglu 2023).

Further, it can increase an oppressed person’s ability to “read the room”. In order to navigate the world while being oppressed, it is often required to attend to nuances in one’s environment—who else is present, how do they

behave, what mood are they in, how do they react to my own behaviour, what is their social status, etc. Changes in tension or the general atmosphere may thus be more accessible to individuals who are primed to look out for such nuances already. This may not only aid them in situations they are used to, but it can also help them notice, i.e., develop a feeling for, when others are being oppressed. Similarly, oppressed individuals may develop tools for emotion regulation that are necessary to navigate the world, such as actively maintaining a supportive social environment, which can better prepare them for new situations that privileged people may not be prepared for. Further, it can make them more adaptive to new situations, in which others would be overwhelmed, or more sensitive to changes in society that bring about new oppressive structures. Crucially, as stressed before, all of this does not simply contribute to an epistemic standpoint but provides a way of being connected to others and the world surrounding oneself, which cannot be brought about through knowledge.

4 Conclusion

This paper has suggested that, while standpoint theory nicely captures the epistemic advantages that can go along with being oppressed, it fails to account for the possibility of affective advantages. Despite incorporating affective experiences merely as a means to gather epistemic knowledge, affectivity is not regarded as a source for advantages that are independent of theoretical and practical knowledge about oppression. In contrast, I have argued that some of the advantages that oppressed individuals can have are not reducible to an epistemic standpoint but are an irreducible characteristic of the affective experiences they make. Mere descriptions of or factual knowledge about oppression cannot replace having affectively lived through the experiences of oppressed individuals. I have shown that advantages comprise an increased sensitivity to structural injustices and social changes, a broader affective repertoire, an enhanced ability to "read the room", and advantages in the development of tools for emotion regulation. Recognising these advantages is not to romanticise oppression but rather to acknowledge the critical insights that arise from it and the potential to implement social change towards a more just society.

This paper also serves as an appeal to debates on epistemology. It seems that we have been paying too much attention to what we can know about the world

while downplaying what our affectivity can tell us about the world that cannot be captured by epistemic knowledge. Debates on affective injustice and the presented suggestion of affective advantages lead the way but there are still quite a few steps to take in order to drag affectivity out of its shadow.

References

- Ashton, N.A. (2019). Relativising epistemic advantage. In M. Kusch (Ed.), *The Routledge Handbook of Philosophy of Relativism* (pp. 329–338). Routledge.
- Dror, L. (2022). Is there an epistemic advantage to being oppressed?. *Noûs*, 57(3), 618–640.
- Dular, N. (2024). Standpoint moral epistemology: the epistemic advantage thesis. *Philosophical Studies*, 181(8), 1813–1835.
- Fricker, M. (1999). Epistemic oppression and epistemic privilege. *Canadian Journal of Philosophy Supplementary Volume*, 25, 191–210.
- Fricker, M. (2007). *Epistemic injustice: Power and the ethics of knowing*. Oxford: Oxford University Press.
- Friesen, P., & Goldstein, J. (2022). Standpoint theory and the psy sciences: can marginalization and critical engagement lead to an epistemic advantage?. *Hypatia*, 37(4), 659–687.
- Gallegos, F. (2021). Affective injustice and fundamental affective goods. *Journal of Social Philosophy*, 53(2), 185–201.
- Jeong, T. (2024). Epistemic Diversity and Epistemic Advantage: A Comparison of Two Causal Theories in Feminist Epistemology. *Hypatia*, 39(1), 97–117.
- Lorde, A. (1984). The uses of anger: Women responding to racism. In *Sister Outsider* (pp. 124–133). Ten Speed Press.
- Medina, J. (2013). *The epistemology of resistance: Gender and racial oppression, epistemic injustice, and the social imagination*. Oxford University Press.
- Mossner, C., & Walter, S. (2024). Shaping Social Media Minds: Scaffolding Empathy in Digitally Mediated Interactions?. *Topoi*, 43(3), 645–658.
- Osler, L., & Krueger, J. (2022). ProAna worlds: affectivity and echo chambers online. *Topoi*, 41(5), 883–893.
- Serano, J. (2007). *Whipping girl: A transsexual woman on sexism and the scapegoating of femininity*. Berkley: Seal Press.
- Toole, B. (2024). Standpoint epistemology and epistemic peerhood: A defense of epistemic privilege. *Journal of the American Philosophical Association*, 10(3), 409–426.

Turkoglu, O., Canavan, M., & Icduygu, A. (2023). The economic limits of empathy? Analysing variation in the attitudes of oppressed minority groups towards refugees. *Journal of Ethnic and Migration Studies*, 49(19), 4850-4869.

Wu, J. (2023). Epistemic advantage on the margin: A network standpoint epistemology. *Philosophy and Phenomenological Research*, 106(3), 755–777.

On Context in Wittgenstein's Philosophical Investigations

Gonzalo Nuñez (Talca, Chile)

Abstract

Drawing on the contrast between the *Tractatus* and the *Investigations*, the paper explores how the notion of context evolves from supporting logical determinacy to grounding an anti-essentialist account of meaning in language games. While contextualism suggests that meaning depends on specific circumstances of use, Wittgenstein's later view resists equating context with meaning, emphasizing instead the normative role of language games. The study further engages with recent debates on contextualism and supervenience, particularly concerning vagueness and the stability of truth conditions. It argues that meaning is not fully determined by syntactic or semantic properties alone, and that vague terms challenge the assumption that truth conditions rigidly supervene on sub-sentential components. Finally, the distinction between context and language game is clarified: while context refers to contingent situations, language games provide the normative background that enables meaningful use. Wittgenstein thus offers a form of contextualism grounded in practice, without reducing meaning to occasion-specific use.

1. Introduction

Frege's ideas have undeniably influenced Wittgenstein's philosophy of language. His work has been essential not only for understanding the critical positions that Wittgenstein adopted in his both early and later philosophy, but also for analyzing and for interpreting the central aspects of that transition from the *Tractatus* to the *Philosophical Investigations*. However, one of the Fregean perspectives on meaning that Wittgenstein inherited transversally in his philosophical thought is the so-called *Principle of Context* (PC). Mainly, this principle contends that a word has no meaning other than its contribution to the general meaning of the proposition in which it occurs.

This study examines the PC's evolution in both the *Tractatus* and the *Investigations*: whereas in the former, the notion of context supports the logical determinacy of meaning within the pictorial theory of language, in the latter it underpins an anti-essentialist view of meaning grounded in the concept of language games. From a pragmatic standpoint, contextualism holds that the meaning of words primarily depends on concrete situations of use. A possible consequence of this view is the destabilization of truth-values insofar as they supervene on context-dependent uses of linguistic components. This

work critiques the supervenience thesis by appealing to vagueness arguing that words may lack context yet remain meaningful through their normative integration within a language game.

2. The Principle of Context

Frege introduces PC in the introduction of *The Foundations of Arithmetic* (also see FA, §62) by saying that “never to ask for the meaning of a word in isolation, but only in the context of a proposition” (1973: p. xxii). His philosophical project aimed to formally demonstrate that the premises and procedures of mathematical proofs are entirely derivable from the general laws of logic. This objective constitutes the core of Frege’s *Begriffsschrift* wherein he develops an ideography or formal notation designed to capture the basic inferences of thought. The system expresses the logic of a priori conceptual content that underlies all valid reasoning. As Frege notes, “the most reliable way of carrying out a proof, obviously, is to follow pure logic, a way that, disregarding the particular characteristics of objects, depends solely on those laws upon which all knowledge rests” (BS: 5). What only matters to logic is the pure nature of thoughts.

Frege’s principal concern is the logical precision of concepts, which demands a strict separation between the psychological and the logical, as well as between the subjective and the objective. Thoughts constitute the conceptual content of judgments, and judgments are what bear reference to truth or falsity. Accordingly, thoughts are neither mental representations nor linguistic constructions; rather, they are abstract and objective entities independent of the empirical content of private mental states and the grammatical structures of natural language. Logic, therefore, is conceived as a system of laws of thought that determines the conceptual content of judgments—laws which underlie linguistic meaning and the normativity inherent in any form of reasoning that aims at truth.

PC follows directly from Frege’s effort to separate the logical from the psychological. If a word could determine its meaning independently of a proposition, it might reduce to subjective states or contingent features of natural language, undermining the objectivity of logical form. Frege also asserts that “the reference of a sentence may always be sought, whenever the reference of its components is involved; and that this is the case when and

only when we are inquiring after the truth value” (Frege, 1892, p. 62). This reflects his compositional principle: a proposition’s meaning is determined by the meanings of its parts and their logical structure. Since truth guides inquiry, the truth-value of a proposition depends on both the reference of its terms and their articulation. Thus, PC holds that a word has no meaning outside a proposition, just as a proposition has no truth-value if any of its parts lack reference.

(PC) For any term x and any proposition p , necessarily, the meaning of x depends on p and its components; (ii) x means nothing without p ; (iii) p ’s value depends on x ’ having a reference.

It is crucial to take into account Frege’s distinction between sense (*Sinn*) and reference (*Bedeutung*). In this regard, although a word cannot fully determine its reference independently of the propositional context, it can nonetheless express a sense by itself. Therefore, the claim ‘ x means nothing without p ’ should be understood exclusively in relation to x ’s capacity to establish a reference rather than its capacity to meaningfully convey a sense.

2. The Principle of Context in Wittgenstein’s *Tractatus*

Wittgenstein embraces Frege’s PC in different ways throughout his philosophy. The general idea of context itself seems to be significant with respect to the ontological commitments to the notion of object assumed by the author in the *Tractatus*. The Austrian author states at the very beginning of his book the strong ontological thesis the “the world is the totality of facts, not of things” (TLP 2001: 1.1). The world as a whole is not composed of objects (*Gegenstände*), but of facts (*Tatsachen*). This assertion does not entail the absence of objects in the world, but rather indicates that objects for themselves do not constitute the world. Facts, understood as states of affairs (*Sachverhalte*) involving combinations of objects (TLP 2001: 2.01), are what language can genuinely depict to constitute the world as a totality. Put it differently, the world is not merely a collection or aggregation of objects; it encompasses the potential arrangement of these objects into states of affairs, which are capable of rendering propositions true or false.

Thus, being a possible constituent of some state of affairs is a necessary condition for the existence of an object (TLP 2001: 2.011). This is what Wittgenstein means when he states that an object has already inscribed in its

internal nature the possibility of occurrence in a certain state of affairs (TLP 2001: 2.012). This ontology also involves an epistemological constrain; namely, to know an object is to know all its combinatorial possibilities at once, so that no new further ones are knowable (TLP 2001: 2.0123). In this regard, Wittgenstein seems to embrace a certain ontological contextualism for objects (simples) according to the following terms:

Given an object x , necessarily, (i) x is a metaphysically indivisible entity, (ii) x exists in the context of a state of affairs as a possibility of logical combination with other objects, and (iii) knowing x implies determining *a priori* all possible contexts of its logical combinations.

The notion of context in this ontology entails that the existence of an object cannot be taken in isolation from its possible combination with other objects in the context of a state of affairs. Given the Tractarian pictorial theory, language mirrors the world, so that an object requires the context of a state of affairs to exist as much as a name requires the context of a proposition to have meaning. Wittgenstein claims:

Only propositions have sense; only in the nexus of a proposition does a name have meaning (TLP 2001: 3.3). An expression has meaning only in a proposition. (TLP 2001: 3.314)

Unlike Frege's view of sense as an abstract realm of thoughts, Wittgenstein conceives it as the logical possibility of the existence or non-existence of states of affairs representable by propositions (TLP 2001: 4.031). A proposition shows its sense through its internal logical form, which it shares with the world; thus, understanding a proposition is to grasp its sense, with no further explanation needed (TLP 2001: 4.02–4.023). Only propositions have sense, as only they can be true or false. Names, by contrast, do not describe but refer directly to objects. Propositions depict facts and thus possess sense, whereas names have meaning by indicating objects as their referents—propositions are like arrows, names like points (TLP 2001: 3.144). While names are the basic units of meaning within propositions, propositions are the basic units of sense in the relation between language and world. As objects exist only within states of

affairs, names gain meaning only within propositions. A name refers meaningfully to an object only when it functions as part of a proposition capable of depicting a possible fact.

3. Context and Use: An Anti-essentialist Thesis of Meaning.

Although the *Tractatus* and the *Philosophical Investigations* represent distinct philosophical paradigms concerning language, they should not be seen as marking a radical rupture between Wittgenstein's early and later thought. Several of his central theses from both periods may appear incompatible or even contradictory. Nonetheless, many reveal elements of continuity or reflect critical reformulations of ideas he maintained throughout his philosophical work—one of which is his adoption and reinterpretation of PC.

As in the *Tractatus*, Wittgenstein remains influenced in the *Investigations*—albeit for different reasons—by the idea that words are not meaningful in isolation. While his early work holds that words require the context of a proposition to determine meaning logically, his later philosophy argues that words acquire meaning through use within language games, thus freeing them from logical fixation. The *Tractatus* asserts that the sense of propositions must be fully determined, and that the possibility of simple signs (or names) is a necessary condition for such determinacy (TLP 2001: 3.23). By contrast, in the *Philosophical Investigations*, Wittgenstein critiques his earlier view of logic as the deep structure of language and essence of all things, whose universal significance awaits discovery (PI 2002: §§89–92). The sublimation of the proposition—described by Wittgenstein as a “queer thing”—involves the tendency to treat logic as a transcendental intermediary between language and the world (PI 2002: §94). Consequently, logical analysis and the determinacy of sense become an unquestioned ideal, akin to a pair of glasses through which everything is viewed, such that “it never occurs to us to take them off” (PI 2002: §103).

In the *Investigations*, the notion of context challenges any form of essentialism—that is, the view that a word's meaning is fixed by an underlying universal form. Wittgenstein applies PC not only to words within propositions but also to their function in concrete circumstances of use. Thus, meaning is not

something an expression possesses independently and later brings into a context (Conant 1998, p. 233). This revised conception is illustrated in the following passage:

You say to me: “You understand this expression, don’t you? Well then—I am using it in the sense you are familiar with.”—As if the sense were an atmosphere accompanying the word, which it carried with it into every kind of application. If, for example, someone says that the sentence “This is here” (saying which he points to an object in front of him) makes sense to him, then he should ask himself in what special circumstances this sentence is actually used. There it does make sense. (PI 2002: §117)

Wittgenstein denies that sense is “an atmosphere accompanying the word” wherever it is used. In the *Tractatus*, sense is the logical possibility of a proposition’s truth or falsity, based on strict pictorial correspondence with reality. In contrast, the *Investigations* present sense as arising from concrete contexts of use within language games, not from a fixed, a priori structure. Sense is not a hidden essence within the word. For instance, the expression “This is here” seems vague—its referents are unclear—but it is only meaningless when isolated from context. Understanding its meaning requires not an essential correspondence with the world, but sensitivity to the specific circumstances in which it is used.

4. The Meaning of a Word is its Use: A Commitment to Contextualism?

Contextualism may be defined as the view that the content of a linguistic expression (type) depends on the specific circumstances of its utterance (token). According to this view, a word’s meaning is not determined by its internal syntactic structure or by abstract semantics alone. Rather, meaning arises from the particular contexts in which the word is used—contexts that include concrete situations, communicative intentions, presuppositions, human practices, and socio-cultural environments—what Wittgenstein called *Lebensform*. For sentences, Bridges (2010) characterizes contextualism in semantic terms: on a given occasion of use, ε is a particular expression of a sentence S , and ε is contextual if the semantic properties of the expressions composing S are insufficient to determine ε ’s content. Language, therefore, is not an autonomous system governed solely by internal semantic rules. Beyond the compositional dependence of sentences on their parts, meaningfulness also

presupposes non-semantic features rooted in the human practices through which language is enacted.

How should contextualism be understood in the *Philosophical Investigations*? Does Wittgenstein's view of meaning as use imply a form of contextualism? Conant (1998) interprets the Context Principle in the *Investigations* through the lens of sense and nonsense: nonsense arises when words are "out of circulation" (PI 2002: §500) or when language is "on holiday" (PI 2002: §38), that is, when expressions lack a context of use. What Wittgenstein retains from Frege is the rejection of psychological reductions in meaning. Sense depends on the alignment between an expression and its context within a language game. Thus, meaning is not determined solely by propositional function or formal rules, but by how expressions operate across concrete uses. Sentences, as Conant notes (1998: p. 242), are tools meant to function in varied contexts.

Contextualism may imply either the elimination (Dobler, 2011) or weakening (Voltolini: 2010) of meaning, as it holds that truth conditions depend entirely on context. Although Wittgenstein introduces language games, there is no clear evidence he supports a full contextualism of truth conditions, where truth values vary with each use. Instead, truth conditions may trivially supervene on the meanings of expressions (Voltolini: 2010). For example, "The south is very rainy" is trivially true if its terms refer to specific contextual conditions. Thus, once the meanings of sub-sentential components are given, so are the truth conditions of the sentence.

5. Vagueness and Context

The compositional thesis assumes that the truth conditions of a sentence supervene on the application conditions of its sub-sentential parts. However, it is possible to think of a sentence in which its sub-sentential parts are semantically vague, but its truth conditions can be clearly determined. Consider the following sentence:

(S) The singer is ready.

Although the truth conditions of *S* may be determinate depending on the context (e.g., being ready to perform, to board a plane, to begin an interview, etc.), its sub-sentential component 'ready' expresses an indeterminate meaning. If the truth conditions of *S* supervene on the relevant properties of

‘ready’, then no change in the former can occur without a corresponding change in the latter. That is, if vagueness is a relevant semantic property of ‘ready’, and ‘ready’ is vague when *S* is true in some circumstance, then ‘ready’ must be precise (or non-vague) when *S* is false. If the vagueness of ‘ready’ is essential to *S*’s being true, and in a different context *S* turns out to be false, it is reasonable to infer that ‘ready’ can no longer retain vagueness as one of its relevant properties.

The critical issue with this supervenience account lies in the fact that any supervenience thesis presupposes a relation of determination between the properties of what supervenes and those upon which it supervenes. The truth conditions of a sentence are said to supervene on the relevant properties of its constituent terms. If vagueness is among those relevant properties—introducing borderline cases in which the application of a term is not clearly determined—then there ought to be a rigid (or at least stable) correspondence between the vagueness of the terms and the truth conditions of the sentences they compose. However, in the practice of ordinary language, the same degree of vagueness in a term can give rise to different truth conditions across various contexts of use, thereby creating tension with the supervenience thesis. For a Wittgensteinian view, the contextualist supervenience thesis seems to be too rigid if we consider vague concepts.

It is important to distinguish between the notions of context and language game, as they are not synonymous. Wittgenstein uses ‘language game’ in various ways: to describe linguistic practices (PI 2002: §23), specific domains of language use, such as construction (PI 2002: §2) or market transactions (PI 2002: §1). There is no single, exhaustive definition of a language game. In contrast, ‘context’ plays a less prominent role in the *Investigations*, where Wittgenstein refers to concrete circumstances of utterance—such as intentions, purposes, expectations, and presuppositions. The concept of a language game, however, carries a broader philosophical significance. It emphasizes that meaning is not merely context-dependent but embedded in a wider network of rules and human practices, both linguistic and non-linguistic. In short, while ‘context’ refers to contingent, particular situations of use, a ‘language’ *game* denotes a more stable, normative framework that grounds the linguistic and social institutions in which those contexts occur.

This distinction does not imply that contexts—contingently circumstances of uses—might not be relevant to understand the meaning of words. Rather, it does imply that words become senseless when no rules of a language game—habitually norms of human practices—govern their use within a community of competent speakers. In this regard, although the meaning of words might occur without a specific context, in no case does it independently of a language game. Someone might unknown the specific context of an expression and yet understands how it could be used according to some rules. This can be interpreted from Wittgenstein's own thoughts:

After he had said this, he left her as he did the day before.” —Do I understand this sentence? Do I understand it just as I should if I heard it in the course of a narrative? If it were set down in isolation I should say, I don't know what it's about. But all the same I should know how this sentence might perhaps be used; I could myself invent a context for it. (A multitude of familiar paths lead off from these words in every direction.) (PI 2002: § 525)

Lack of knowledge about the context of an expression does not necessarily entail misunderstanding its meaning. Even without any information about the specific context of use—such as what was said or done the day before, as in Wittgenstein's example—one may still grasp how a sentence could be used in some possible, even unpredictable, situation, insofar as one can imagine a context. However, this is only possible if one understands the relevant language game or the application rules shaped by a history of human practices and habits.

A language game thus provides the linguistic and non-linguistic framework within which a context of use may arise, though no particular or predefined context is needed to constitute a language game. Is Wittgenstein, then, a contextualist? In some respects, yes—meaning cannot be grasped in isolation, as Frege argued. But not in the strict sense: for Wittgenstein, words may still be meaningful without a specific context, so long as they belong to the normative structure of a language game.

References

Bridges, J. (2010) “Wittgenstein vs Contextualism”, in: Arif Ahmed (ed.) *Wittgenstein's Philosophical Investigations: A Critical Guide*, Cambridge: Cambridge University Press, 109-128.

Conant, J. (1998) "Wittgenstein on meaning and use", *Philosophical Investigations*, 21(3), 222-250.

Dobler, T. (2011) "Two Conceptions of Wittgenstein's Contextualism", *Lodz Papers in Pragmatics*, 7(2), 189-204.

Frege, G. (1973) *The Foundations of Arithmetic*, (trans. J. L. Austin), New York: Harper & Brothers.

Frege, G. (2002) "*Begriffsschrift*: formula language, modeled upon that of arithmetic, for pure thought", in: Jean van Heijenoort (ed.), *From Frege to Gödel: A Source Book in Mathematical Logic, 1879–1931*, Cambridge (Massachusetts): Harvard University Press, 1-82.

Voltolini, A. (2010) "Is Wittgenstein a Contextualist?", *Essays in Philosophy*, 11(2), 150-167.

Wittgenstein, L. (2002) *Philosophical Investigations*, (trans. G. E. M. Anscombe), Oxford, UK: Basil Blackwell.

Wittgenstein, L. (2001) *Tractatus Logico-Philosophicus*, (trans. D. F. Pears and B. F. McGuinness), Oxford, UK: Routledge.

Acknowledgement

This work is part of the research project "A study on conceptual vagueness based on the notion of 'fuzzy limits' in Wittgenstein's late philosophy", **FONDECYT Iniciación N° 11240226** sponsored by the National Agency for Research and Development (ANID in Spanish) depending on the Ministry of Science, Technology, Knowledge and Innovation of the State of Chile.

Two Ways of Taking the Ordinary Seriously

Eylem Özaltun (Istanbul, Turkey)

Abstract

It is reported that Anscombe and Austin disliked each other and rarely (See “Pretending” 1958) engaged directly with each other’s work. Their philosophical agreements and disagreements are not explicitly worked out in their writings. A comparative analysis of their work is left to us. In this paper, I engage in a comparison of not their views on various subjects, but rather their respective philosophical methods. Although Austin and Anscombe were interested in similar topics and lived in the same cultural atmosphere, the spirit in which they engaged in philosophy was fundamentally different. Or so I argue in this paper. I present this fundamental difference by focusing on the various ways in which they take the ordinary seriously. I show that Austin’s use of the ordinary as a philosophical method evolved considerably between “Other Minds” (1946) and “A Plea for Excuses” (1956). I argue that while “Other Minds” focuses on the ordinary as an antidote for overlooking the logical complexity, and hence is very much a Wittgensteinian text, the later version of Austin’s method we find in “A Plea for Excuses” is at odds with the use of ordinary for logical investigation we find in Wittgenstein’s and Anscombe’s work.

The way we talk is the starting point of philosophical investigations for many philosophers, especially for the analytical philosophers of the post-*Tractatus* era. Hence not surprisingly the ordinary language is philosophically significant for both Austin and Anscombe. However, as we will see, there are important differences in the way in which ordinary language is used by Austin and Anscombe. In what follows, I list some preliminary observations as a first step towards a comprehensive comparison.

Anscombe’s Complaint

Anscombe explicitly finds Austin’s engagement with the ordinary language problematic. A clear expression of this can be found in the opening lines of “Pretending”.

“Offered ‘pretending’ as a philosophical topic, I should want to distinguish between mock performances and real pretenses. The difference, so far as I have noticed, is not pointed to by any of those differences between the grammatical constructions respectively appropriate, sometimes to one nuance of sense and another, sometimes to one word and another closely related one, which are Professor Austin’s special interest. Hence he disregards it, and lumping dissimilar things together, finds that in ‘the basic case’ the one who is pretending must be giving a ‘current personal

performance' in someone's presence in order to disguise what he is really doing." (Anscombe 1958: 83)

Anscombe's criticism of Austin's views on 'pretending' is not pertinent to our current purposes. Leaving their debate on 'pretending' aside, I want to focus on the methodological complaint Anscombe voices in these lines. Here Anscombe claims that due to paying attention to wrong things in the use of language, Austin misses philosophically relevant distinctions. The philosophically significant is obscured by Austin's detailed analysis of the nuances in the senses of different words and grammatical constructions of English. In this analysis, Austin, according to Anscombe, is losing track of the phenomena under investigation.

In a passage from "The Intentionality of Sensation," we can find Anscombe making the same point by an anecdote:

"Again, things do not always look the same shape, colour, size and so on, but we commonly look at and describe them, saying, e.g., 'It's rectangular, black and about six foot in height,' without paying attention to how they look—indeed we might say that often things *look* to us, strike us, not as they look but as they are! (Conviction that *only* so is 'looks' used rightly was the cause of confusion to an over-confident ordinary-language philosopher on an occasion famous in Oxford: F. Cioffi brought in a glass vessel of water with a stick in it. 'Do you mean to say', he asked, 'that this stick does not look bent?' 'No' said the other bravely: 'It looks like a straight stick in water.' So Cioffi took it out and it *was* bent.)" (Anscombe 1965:14)

Here Anscombe does not give the name of this over-confident ordinary-language philosopher. She probably assumes that everybody knows who is, given that the bent-stick incident was well-known at the time. According to a recent biography of Austin (Rowe 2023), Brian McGuinness confirmed that this over-confident ordinary-language philosopher was Austin. I take this to be another place where Anscombe makes a methodological complaint: a certain view on how we are to use the word 'look' precludes Austin from seeing what

is in plain view. Hence this is another place where Anscombe diagnoses what is wrong with ordinary-language philosophers' way of using ordinary in philosophical investigations.

The Method of "A Plea for Excuses": Linguistic Phenomenology

In "A Plea for Excuses" Austin explains his method as "to proceed from 'ordinary language'... by examining *what we should say when, and so why and what we should mean by it.*" (Austin 1956: 7) This statement of his method might make it seem that the interest of the study is the language use itself. To avoid such misunderstanding Austin adds:

"When we examine what we should say when, what words we should use in what situations, we are looking again not merely at words (or "meanings", whatever they may be) but also at the realities we use the words to talk about: we are using a sharpened awareness of words to sharpen our perception of, though not as the final arbiter of, the phenomena." (Austin 1956: 8)

So in using such a method Austin is interested in investigating reality and not just linguistic practice. However, it is not obvious why a sharpened awareness of words would sharpen our perception of reality. That would depend on what sort of sharpening in our understanding of the way we use our words is being sought here. Anscombe's complaint in 'Pretending' and also in "Intentionality of Sensation" is that Austin's hyper-awareness of nuances is not the right sort of sharpening; if anything, it is detrimental to our seeing the phenomena aright.

We can glean more about the nature of Austin's method from what he says in explaining why he chose 'excuses' as his topic of investigation (cf. Austin 1956: 7-9) From his defense of the vocabulary of 'excuses' as a fruitful place to dig in for philosophical insight, we can extract the following three features of his method "linguistic phenomenology":

1. *Purity*. As we proceed from the ordinary language, we are to focus on *the pristine regions of ordinary language*: that is, the regions that are not contaminated by philosophical study. So we should not start with words like "reality" "knowledge" "cause" "freedom" "reason" "object", etc. These words express concepts that philosophers pondered on long enough that

even ordinary usage is not immune from the theories they put forward. Hence as far as these words are concerned our ordinary language is contaminated with theory and the use of such words cannot be the starting point of linguistic phenomenology.

2. *Ahistoricity*. Relatedly in using this method, we are to *disregard the history of philosophy altogether*. And when we focus on the pristine regions of ordinary language we can do so. This will open the possibility of studying “even spontaneousness, without remembering what Kant thought, and so progress by degrees even to discussing deliberation without for once remembering Aristotle or self-control without Plato.” (Austin 1956: 9)
3. *Subtlety*. We should study the pristine regions of ordinary language which are concerned with the most pressing practical matters. In topics that matter to the ordinary man most, to facilitate everyday communication and cooperation, a rich vocabulary that is capable of making subtle distinctions emerges organically. In studying these not philosophically but *practically generated subtleties of the ordinary language*, we are to gain a sharpened sense of reality.

The thought that underlies these features seems to be that we need a fresh start for our philosophical investigations, free from the influence of existing philosophical theories. Yet it is not at all obvious why a close examination of *practically generated subtleties of the ordinary language* would reveal philosophically interesting distinctions. Take for example color words. For a fashion designer, an interior designer, a painter, and a butterfly taxonomer, colors pose practical challenges. These challenges are addressed by a rich vocabulary of color words which enables us to talk about subtle differences of colors in several dimensions. But why should a study of the nuances of the sense of color words be significant for the philosophical study of color perception? Anscombe rejects this assumed path from subtlety to significance.

Anscombe’s focus on the ordinary language as a starting point of a logical investigation does not exhibit any one of these three features.

1. Anscombe asks the most well-known philosophical questions and in answering or dissolving them engages in traditional philosophical debates. The concepts she is interested in are central concepts of

philosophy (cause, reason, knowledge, action, substance, sensation, experience, time, reality, etc.) and the use of language that is relevant to studying these concepts cannot be found in the pristine regions of ordinary language.

2. Anscombe used history in an inspiring way in producing original work. For example, she wouldn't dream of talking about practical deliberation, as Austin hopes, without engaging, admiring, and criticizing Aristotle's work on practical syllogisms. She would not even be interested in the use of "I", and write "The First Person" 1975, were she not fascinated with the Cogito argument since 1944. She would not criticize the semantic account of "I" if she was not sensitive to the philosophical concern that animates Cogito argument. Anscombe's work is one of the best examples of how to do problem-based contemporary philosophy on the most central philosophical questions while engaging with the history of philosophy in a critical and not just purely interpretive way. Hence ordinary language is never used as a refuge from the history of philosophy for Anscombe.
3. Anscombe is not interested in the subtleties of English. Whatever she aims to uncover by focusing on how people operate in particular circumstances was never meant to be parochial in this sense. *Her investigations into the grammar of our concepts are always logical, and her focus on particular circumstances is an antidote for overlooking the logical complexities in favor of one fundamental logical form.* That is why her appeal to language use involves the broader and more general aspects of language use that tend to survive in translation.

Moreover Anscombe, like Wittgenstein, carries out the the investigation into the logical complexities by looking at the ordinary in comparison with the non-ordinary, the existing use in contrast with the fictional, imagined, constructed one. Particular language games used by Wittgenstein and Anscombe as objects of comparison to our ordinary usage are intentionally less rich, less subtle, and more primitive than our ordinary use of language. What I have in mind here is the use of language games like builders in *Philosophical Investigations* (PI 2009: §2) and A-users in "The First Person" (Anscombe 1975: 24). One important function of considering these language games and comparing them to our ordinary use is to leave aside subtleties and richness Austin praises so much in the ordinary, so that we can focus on logically salient, broader aspects of our

use. For Anscombe and Wittgenstein, the ordinary is illuminated partly by comparison to the non-ordinary, and the sophisticated by contrast to the primitive. In thinking about these made-up language games Anscombe and Wittgenstein are appealing to a shared, human understanding of what makes sense which goes beyond the native mastery of English. Hence for Anscombe, Austin's fascination with subtleties of English is a sublimation of the ordinary, is at least as pernicious to our understanding as the sublimation of logic and probably less fruitful than that.

The Method of "Other Minds": Grammatical Investigation

Austin's 1946 paper "Other Minds", on the other hand, is not an example of what he calls "linguistic phenomenology" in "A Plea for Excuses". In fact, it is a good example of the type of interest in the ordinary language which I associated above with Anscombe and Wittgenstein.

1. In "Other Minds," Austin does not focus on a pristine region of ordinary language. On the contrary, he focuses on the words that appear in the expression of most sophisticated theories as well as in ordinary practice, such as "knowledge," "real," and "certainty."
2. "Other Minds" deals with central philosophical problems such as the possibility of knowledge, whether knowledge is a subjective attitude, whether knowledge is a species of belief, whether sensations provide a basis for all our empirical knowledge, whether there is an asymmetry in our knowledge of the physical and mental.
3. "Other Minds" focuses on the particular cases in which we say "I know" and "I believe" and what else we say when we are challenged. It distinguishes how we back up our initial claims to knowledge and belief, and what we say and do when we cannot back up these claims. In this investigation, there is a great deal of emphasis on what we ordinarily say. One of the central observations of "Other Minds", for example, is that we ask, "How do you know?" and "Why do you believe?" but do not ask "Why do you know?" and "How do you believe?". In making such observations, the focus is not on the rich vocabulary with fine distinctions of sense, but rather on the general structure of different ways in which we defend various positions we take or have in the face of reality. So, what is at stake are the practices of justification, persuasion, raising doubt, reaching an agreement, etc. The features of ordinary

language that are revealing about these practices are the broader aspects of our language use which are not parochial but crosslinguistic.

What we ordinarily say as opposed to what we read in philosophical treatises is an interest only in so far as it reveals how we settle certain questions in the particular settings in which they arise and enables us to discover some important logical distinctions that were covered up during the abstraction required for the general theories. The nuances that are conducive to such a discovery must be distinguished from the logically irrelevant subtleties of a particular language. “Other Minds” studies the concept of certainty and its relation to knowledge and to do so it looks closely at our ordinary ways of dealing with uncertainty. This requires attention to the use of expressions in such dealings but only at the right level of specificity. Not every subtlety of use of the word “certain” or “sure” would be relevant. Austin expresses his sensitivity to this point as he catches himself when he is tempted to indulge in such subtleties:

“‘Being certain’ tends to indicate confidence in the current perception. Perhaps this comes out in our use of the concessives ‘to be sure’ and ‘certainly’, and in our use of such phrases as ‘certainly not’ and ‘surely not’. But it may be unwise to chivvy language beyond the coarser nuances.” (Austin 1946: 165)

Indeed this is exactly what Anscombe finds unwise in what she calls “Professor Austin’s special interest” in 1958 as we have seen above. Judging from this passage, Austin of 1946 would agree. From these preliminary observations, I conclude that there has been a shift in Austin’s use of the ordinary for philosophical illumination between “Other Minds” and “A Plea for Excuses”. While the former’s use of ordinary language is very much in line with Anscombe’s and Wittgenstein’s, the latter’s is not. The shift I suspect occurred around 1950. Hence, in these two papers, Austin exhibits two different ways of taking the ordinary seriously that I sketchily distinguish above.

References

Anscombe, G.E.M (1958) “Pretending” in (1981) *Metaphysics and Philosophy of Mind: Collected Philosophical Papers Volume II* Oxford: Basil Blackwell, 83-93.

Anscombe, G.E.M (1965) "The Intentionality of Sensation" in (1981) *Metaphysics and Philosophy of Mind: Collected Philosophical Papers Volume II* Oxford: Basil Blackwell, 3-20.

Anscombe, G.E.M (1975) "The First Person" in (1981) *Metaphysics and Philosophy of Mind: Collected Philosophical Papers Volume II* Oxford: Basil Blackwell, 21-36.

Austin, J.L. (1946) "Other Minds", *Aristotelian Society Supplementary Volume*, Volume 20, Issue 1, 148–187.

Austin, J.L.(1956) "A Plea for Excuses: The Presidential Address" *Proceedings of Aristotelian Society*, New Series, Vol 57 (1956-1957), 1-30.

Rowe, M.W.(2003) *J.L. Austin: Philosopher and D-Day Intelligence Officer* Oxford, GB: Oxford University Press

Wittgenstein, L. (2009) *Philosophical Investigations*. P. M. S. Hacker & Joachim Schulte (eds.) Wiley-Blackwell.

The default use of gender-neutral pronoun they and our epistemic (and non-epistemic) lives

Jeanne Pang (Vienna, Austria)

Abstract

In this paper, I will discuss how epistemic interventions, that aim at the flourishing of our epistemic agency, take place in our daily interpersonal exchanges, and how such interventions can plausibly address epistemic oppression. I grant that addressing epistemic oppression is also a necessary means of resisting social oppression.

An example of such an epistemic intervention will be discussed: the default use of gender-neutral pronoun they in English-speaking contexts. I will first illustrate some scenarios of its implementation and argue that it is not merely a linguistic practice, but also an epistemic practice, with the effects in our epistemic lives that it addresses the shared social imagination about gender. This intervention specifically addresses gender norms ascription that arguably sustains epistemic and gender oppression, by drawing our attention and making us more sensitive to alternative epistemic resource prompting us to a more adequate social imagination compared to the dominant ones, as well as facilitating our self-knowledge and knowledge about others. Not less significantly, I shall draw attention on its contribution in the resistance to non-epistemic oppression (gender oppression in this case).

1. Social knowledge and our epistemic lives

Before proceeding, here are my premises throughout this paper. First, I take epistemic interventions to aim at the flourishing of our epistemic agency, not merely the acquisition of true beliefs, the latter which the standard views primarily focus on (see for example Ahlstrom-Vij 2013a, 2013b; Bullock 2018; Goldman 1991).

Secondly, my view on epistemic interventions allows diverse types, not only the two types (epistemic access control or epistemic nudging) that are focused on in standard accounts. Other than assuming them to be merely enabling or disabling, or drawing agents' attention to evidence or true beliefs, this paper aims to illustrate how epistemic interventions can also be about fostering sensitivities to our epistemic environment and habits. I also aim to show how challenging it can be for epistemic agents to revise their body of knowledge, in the presence of alternative epistemic resources, or the salience of their presence.

A specific type of knowledge, social knowledge, is the concern of this paper. Social knowledge broadly includes self-knowledge and knowledge about

others. Epistemic interventions can enhance our epistemic agency by making us sensitive to our *lack of* knowledge about others, i.e. being sensitive to what we don't know, encouraging us to reconsider certain assumptions about others. As an epistemic intervention, such as the default use of gender-neutral pronoun, serves to (1) resist gender oppression that manifests as gender-norms ascription and (2) enable contribution of alternative epistemic resources by other people, including gender nonconforming people, and therefore more adequate social imaginations.

Thirdly, drawing on José Medina's (2013) interactionist view, I hold that we must exercise our epistemic agency reciprocally and interchangeably in order to flourish. That means, we don't perform our epistemic roles merely as informants for each other, but dynamically as inquirers, accessors and evaluators, in the way that we collaborate in the pursuit and use of knowledge, generating and interpreting possible meanings (7–95). As the flourishing or failure of our epistemic agency cannot be achieved by individuals alone, but occur within and across social groups, epistemic interventions aiming at the flourishing of our epistemic agency, in light of our social reality being structured by epistemic and non-epistemic forms of oppressions, are *prima facie* justified.

2. The place of gender ascription and gender pronouns in our epistemic and social lives

In line with Dembroff and Wodak (2018) and Kukla and Lance (2022), I hold the view that using gendered pronouns (*he/she*) in English, along with other acts of gender ascription, is a speech act that does more than merely describe or attribute properties to individuals. It is not about making truth claims but about imposing the dominant social imagination of gender onto individuals and perpetuating the current social arrangements and expectations about how one should act in social spaces. When a speaker attributes gendered pronouns to an individual, they are not just expressing of the referent's gender, or describing what they have assumed or perceived about that person's biological sex assigned at birth. They are also assigning a set of social norms and roles that are associated with that gender along with their perceived biological sex, signalling to others (and themselves) how this individual is positioned within

their society and what social roles they ought to fulfil. As the very use of the gendered pronoun *can* reinforce gender roles ascription, it does not require speakers' explicit intention on reinforcing gender stereotypes.

This is similar to how holding the belief of “women are caretakers” contributes to gender oppression. In both cases, upholding such a belief and ascribing gender to others through language inserts a prescriptive force into the social realm, masked as descriptive statements. For instance, consider a sport player saying, “I lost (the match) because *she* is my teammate”, in contrast to other players with their male teammates. While the player *may* be referring to a specific individual who they believe didn't perform well, they *may* also be implying that because their partner's gender, i.e. being a *woman*, suggests less capability, given stereotypes about *women's* abilities in most sports. Regardless of the speaker's intention of referring to their teammate's gender, by invoking *she* in this case, the speaker situates their teammate within the dominant social imagination of gender and likely perpetuates this. Therefore, refraining from the use of gendered pronouns avoids enacting these gender ascription practices that contribute to essentialising gender roles as natural roles and further upholding them as norms.

3. Going by singular *they* by default as an epistemic intervention of resistance

This example of an epistemic intervention is the default use of *they*, and *them*, as singular, third-person pronoun forms specifically in English-speaking contexts. By replacing singular, gendered pronouns *he* (or *him*) and *she* (or *her*) by the default use of *they* (or *them*), the speaker avoids to explicitly express or ascribe gender to the people they refer to. The main goals are, by refraining from gender ascription, resisting gender oppression, and challenging the epistemic oppression that both constitutes and sustains gender oppression.

This practice is focused on English-speaking contexts, because in many languages, gender ascription works differently. For instance, on the one hand, there is only one third-person pronoun and that is gender-neutral in Bengali (“shay”), and there is no gender-specific pronouns in Cantonese (“keoi5”) and spoken Chinese Mandarin (as all are “tā”). On the other hand, languages like Czech and German not only have gendered pronouns, but their nouns and adjectives are also inflected for gender (such as “chytrý filozof” for a smart male philosopher and “chytrá filozofka” for a smart female philosopher in

Czech; “ein Philosoph” for a male philosopher and “eine Philosophin” for a female philosopher in German). However, this does not mean that gender ascription does not exist for the former two languages, nor there is no way to address gender ascription for the latter two. It means that to address gender ascription associated with gender and epistemic oppression within those languages requires different strategies.

The epistemic intervention in question can apply to those language speakers who speak English, be English their first language or not. It is because the default use of singular *they*, as I will argue, works to draw attention from the interlocutors to the shared background assumptions we are socialized to have that associate and ascribe roles and norms with a person based on their perceived or imagined bodily features. As I illustrate in the examples below, instead of assuming a hierarchy of epistemic authority in speakers who use *they* by default as they refer to a third person, it functions as an epistemic intervention through motivating a communicative dynamic between the interlocutors, the interplay of their epistemic roles as informants, inquirers, evaluators. It does so by invoking our epistemic authority, in making our reflective choices on and taking responsibility for how we process and evaluate our epistemic resources as well as our knowledge acquisition.

Besides, going by singular pronoun *they* by default is not a sufficient, but a necessary epistemic intervention to address gender role ascription, at least in the times as I write this paper; for it can no longer become useful when it loses its salience in drawing interlocutors’ attention or raising their awareness and sensitivity.

Below are three scenarios.

Scenario 1

Jean: *They* play good rally and *their* shots have very high quality.

Matthias: Who do you mean?

Jean: (Pointing at a person) *They*.

Scenario 2

Quinn: *They* are very caring for people around.

Strauss: Is *she* joining us today?

Quinn: No, *they* are at work.

Scenario 3

Mahek: I recognise you are using *they* for everyone.

Lan: Yes, I don't want to assume a gender for people.

Mahek: What pronoun do you go by?

Lan: I go by *they* as well.

Mahek: Okay, I will try to practice that.

In scenario 1, it can be a semantic issue to begin with. Matthias may assume that there is more than one person that Jean is referring to. At the end of this conversation, Matthias may think that Jean refers to “a group of people *like this* individual player play very well” at the sport in question, but not this individual. It can also simply puzzle Matthias because of their gender perception of the person Jean refers to, either female or a male, does not match with Jean's gender-neutral reference. In either case, Matthias appears to be unfamiliar with the use of singular *they* pronoun as a linguistic practice. Moreover, it is likely that Matthias has no dispositions to gender-neutral pronouns or grammar, as well as epistemic resources, ranging from testimonies from other people to a different social imagination about how gender may look like for themselves and others; and Jean's use of singular *they* pronoun catches Matthias attention.

In scenario 2, Strauss does not know who Quinn is referring to, but assumes the person being a woman for Quinn's description of them being caring. It is not so much about whether the person Quinn talk about is a female person. Yet, Quinn's response with *they* noticeably contrasts with Strauss' use of pronouns. Strauss can interpret Quinn's response in various ways, such as Quinn was communicating the person's non-binary gender identity,

In scenario 3, Mahek appears to misinterpret Lan's default use of singular *they* pronoun as a matter of self-identification, that Lan, for example, identifies as non-binary. Lan can go on to elaborate their point of this practice, in the same conversation or another occasion. Nevertheless, even before that, it raises Mahek's awareness that their interlocuter, and possibly other people, don't understand gender in the way as Mahek does.

As ambiguous conversations can often be, epistemic enhancement through interactions is not always effective and immediate; it is often not successful by single effort. Additionally, given the dynamic feature of our epistemic interactions, even speakers such as Jean, Quinn and Lan, who use singular *they* pronoun by default in their conversations with other people, are not “immune” from the reactions of their interlocuters, in holding their confidence as knowers and their resistance to the dominant social imagination of gender. Whether they are confident specifically about their practice of the use of singular *they* pronoun, as a resistance to the dominant social imagination and the gender oppression enacted thereby, or are in general epistemically confident in their ability to contribute to social knowledge, partially but also significantly depends on the uptake of their epistemic resources by their interlocuters. This also implies that even though epistemic interventions aim to facilitate the epistemic agency of the oppressed, it also does so to that of the privileged. That is because the privileged hold the epistemic power to have epistemic resources shared in, and to effect what becomes part of the collective pool of social knowledge (Dotson 2014).

Moreover, using singular *they* can avoid misgendering (Dembroff and Wodak 2018) as there is no background assumption on a person’s gender in the interactions, and it avoids, in contrast to our current language practices of gender pronouns often do, denying a person’s gender self-ascription assaults their “core bodily integrity and autonomy” (Kukla and Lance 2022, 1144). Therefore, as it resists the dominant practice of assuming and ascribing gender to individuals, using gender-neutral pronouns by default respects individuals with their self-ascribed gender identities.

Yet, another question may arise: Does this practice backfire towards *women* and their situations, particularly to deny the oppression of women? How can we intervene with care of this? In addition to the default use of gender-neutral pronoun not excluding going by an individual’s identified pronouns and respecting a person’s self-ascribed gender, it can cooperate in contexts where specifying a person’s circumstance of either oppression or privilege, related to their gender ascription, is necessary to challenge gender norms and expectations associated with them. For example, we can refer to a female person as *she*, when addressing the harm done to them due to their perceived gender, or refer to a male person as *he*, when discussing someone

essentialising their gender privilege and perpetuating norms of female inferiority, such as female people should be caretakers, stay at home, or refrain from certain activities due to their gender.

References

- Ahlstrom-Vij, Kristoffer. (2013a). *Epistemic Paternalism: A Defence*. Palgrave Macmillan.
- Ahlstrom-Vij, Kristoffer. (2013b). "Why we cannot rely on ourselves for epistemic improvement". *Philosophical Issues*, 23, 276–296.
- Bullock, Emma. C. (2018). "Knowing and Not-Knowing For Your Own Good: The Limits of Epistemic Paternalism". *Journal of Applied Philosophy*, 35(2), 433–447. <https://doi-org.uaccess.univie.ac.at/10.1111/japp.12220>.
- Dembroff, Robin & Daniel, Wodak. (2018). "He/She/They/Ze". *Ergo*, 5(14). <https://doi.org/10.3998/ergo.12405314.0005.014>.
- Dotson, Kristie. (2014). Conceptualizing Epistemic Oppression. *Social Epistemology*, 28(2), 115–138. <https://doi-org.uaccess.univie.ac.at/10.1080/02691728.2013.782585>.
- Goldman, Alvin. (1991). Epistemic paternalism: Communication control in law and society. *The Journal of Philosophy*, 88(3), 113–131.
- Kukla, Quill & Lance, Mark. (2023). Telling Gender: The Pragmatics and Ethics of GenderAscriptions. *Ergo*, 9(42). <https://doi.org/10.3998/ergo.2911>.
- Medina, José. (2013). *The Epistemology of Resistance: Gender and Racial Oppression, Epistemic Injustice, and Resistant Imaginations*. Oxford: Oxford University Press.

Acknowledgement

My great many thanks to Paulina Sliwa, my master's supervisor, for both of their inspirations and support for my project. My gratitude to Shohei Matsumoto for taking great interest in with my project and the thoughtful discussions. Also, thanks to Matt Dougherty, who read my early drafts of my thesis, from which my paper is divided. And Lasse, as well as many friends and those who engaged in conversations with me, keep me thinking about the working of epistemic interventions in our everyday life.

Strange Props: Reflections on the Status of Hinges in and beyond *On Certainty*

Samuel Pedziwiatr (Hagen, Germany)

Abstract

This paper critically reflects on the debate over the propositionality of hinge certainties in the analytic reception of Ludwig Wittgenstein's *On Certainty* (OC) within hinge epistemology. The paper proceeds in three steps. First, I present some biographical and historical context on OC and Wittgenstein's grammatical analyses of "knowledge" and "knowing" in the *Nachlass*. Second, I outline the main philosophical conflict to which OC responds: the irresolvable tension between skeptical (*modus tollens*) arguments and G.E. Moore's (*modus ponens*) defense of Common Sense Realism. This epistemological stalemate is presented as a paradigmatic example of the kind of traditional philosophical dispute that Wittgenstein's grammatical method aims to dissolve. Third, I critically reassess several proposals from hinge epistemology that seek to overcome this impasse, with particular attention to how they account for the propositionality of hinges and the problem of epistemic closure. In this context, I advance considerations that cast doubt on a central tenet of *framework* readings of OC, according to which hinges constitute grammatical rules and are therefore not susceptible to skeptical challenges or rational evaluation. The paper concludes with an outlook on potential upshots of an anti-essentialist Wittgensteinian perspective for the development of a feminist hinge epistemology.

A metereological journal of the mind. You shall observe what occurs in your latitude, I in mine. — Henry David Thoreau, *Journal* (August 19th, 1851)

Hinge epistemology (HE) holds that knowledge rests on unquestioned and non-justified basic commitments, or *hinges*, within our common (epistemic) practices. The claim that the conditions of the possibility of knowledge and reasoning consist in groundless, arational features of our acting is inspired by Wittgenstein's hinge metaphor in *On Certainty* (OC):

"[...] the *questions* that we raise and our *doubts* depend on the fact that some propositions are exempt from doubt, are as it were like hinges on which those turn. [...]"

If I want the door to turn, the hinges must stay put." (OC 1969: §341-3)

In OC, this metaphor illustrates how scientific researchers must rely on prior theoretical knowledge and practical experience in their investigations to generate new insights. In HE, the term *hinge* is used in a more technical and general sense to refer to primitive epistemological features upon which apt beliefs are presupposed. Accordingly, *hinge proposition* refers to statements

expressing such basic certainties. Unlike traditional approaches in epistemology that view “basic beliefs” as the bedrock of knowledge and are prone to justificatory infinite regress, HE situates knowledge within a respective form of life and lived practice. HE thus offers an interesting alternative to foundationalism and coherentism, particularly for those who wish to respond to the *skeptical challenge* while accounting for social, cultural, historical, and practical aspects in knowledge-production processes. In this vein, HE has developed “Wittgenstein’s original insights into a systematic approach to issues concerning justification, knowledge, scepticism, relativism, disagreement, testimony, trust, epistemic injustice and more” (Coliva 2022: 51)

Like most metaphors, the notion of *hinge* is open to competing interpretations. This is reflected in the debate over how primitive *hinges* relate to genuine propositions, in the standard sense of bipolar truth-evaluable statement. Whereas Annalisa Coliva (2022) and Duncan Pritchard favor selective or hybrid conceptions of the propositionality of hinges, Danièle Moyal-Sharrock (2022) considers *nonpropositionality* an essential feature: hinge certainties represent reflex-like, natural (or second-nature) *animal* responses; they are nonveridical and cannot express empirical or factual content. Despite such differences, many HE proponents endorse a *framework* reading of OC, according to which hinge propositions express *rules of grammar* that represent bounds of sense, and which therefore cannot not be subjected to rational appraisal. (Coliva 2022: 11)

In the following, I present critical reflections on hinges’ propositionality and cast doubt on the claim that hinges constitute rules of grammar. Section 1 situates OC in historical and biographical context, and provides some methodological background on Wittgenstein’s grammatical analyses of “knowledge” and “knowing” in the *Nachlass*. Section 2 presents the logical structure and underlying assumptions of the basic conflict between skepticism and Moore’s Common Sense realism at the heart of OC, focusing on Moore’s *here is one hand* argument. In section 3, I critically reconsider the status of hinge certainties in the treatment of *Moore’s Shift* in OC and HE. In this context, I examine how Pritchard (2012), Coliva (2016), and Moyal-Sharrock (2022) characterize hinges’ propositionality and deal with the problem of

epistemic closure. Finally, I give an outlook on potential upshots of the defended anti-essentialist Wittgensteinian perspective for developing a feminist HE.

1. The Grammars of “Knowledge” and “Knowing”

As G.E.M. Anscombe and Georg Henrik von Wright write in the preface to OC, Wittgenstein discussed G.E. Moore’s “Proof of an External World” and “Defence of Common Sense” with Norman Malcolm during his visit to America in 1949. While both papers are essential for reading OC in its right relief, it is important to remember that these lines of investigation and engagement with Moore’s theory were by no means new to Wittgenstein. (Rhees 2003: 5)

Malcolm’s memoir recounts a philosophical altercation between Wittgenstein Moore that took place a decade earlier: In April 1939, Wittgenstein caught wind that Moore criticized his views in a talk to the Moral Science Club, to which other faculty members were not invited. Wittgenstein reportedly responded “like a war horse” (Malcolm 2009: 30) and confronted Moore during an at-home discussion on May 2nd, 1939. The notes by Malcolm document how Wittgenstein attacked Moore’s claim that “it is certain that” had the same meaning for sense-statements and thing-statements, by trying to show that the usage is grammatically deviant in the former cases. A day later, Wittgenstein reluctantly apologized for treating Moore “as if stupid”. (Wittgenstein, Moore, Malcom & Citron 2015: 78)

Questions about the nature of knowledge feature prominently throughout the entire *Nachlass*. Plato presents a further, albeit less obvious, influence on OC. In August 1931, Wittgenstein wrote the following remark (MS 111: 69) about *Theaetetus* (146c-e), which he later transferred to the *Big Typescript*:

“(Socrates asks the question what knowledge is and he isn’t content with an enumeration of instances of knowledge. But we don’t pay much attention to that general concept, and are happy if we understand shoe-making, geometry, etc.)” (BT 2005: 56e)

As I have argued elsewhere, the self-contrast with Plato provides a useful lens for understanding Wittgenstein’s grammatical method, and coincides with a shift away from key Tractarian tenets such as the logical independence of elementary propositions. (Pedziwiatr 2024) Several related remarks in the

Nachlass elaborate upon Wittgenstein's criticism against Socrates' desire to find the purported *essence* of "knowledge itself". (MS 114: 85r; TS 213: 66r; MS 157a: 60v) Where Plato seeks underlying principles, general definitions, or a foundation of knowledge, Wittgenstein invites philosophers to ask what concrete instances of actual knowledge look like. Wittgenstein eschews metaphysical *über-knowledge*.

Since the *Tractatus*, Wittgenstein consistently maintains a fallibilist and propositional conception of (factual) knowledge, connected to his conception of propositionality and verifiability. In this regard, Wittgenstein's treatment of "knowledge" in the *Nachlass* shows a high degree of continuity. The aim of Wittgenstein's grammatical analysis is not to *secure* knowledge across all domains, but to provide conceptual clarification, as a pocket notebook entry from the early 1930's illustrates: "The foundations we mean pervade rather than underly mathematics + the sciences. (siehe Augustinus et cum effunderis super nos, non tu iaces, sed irigis nos.)" (MS 153b, 2v.) The answer to the philosophical question "what is knowledge?" that he proposes is to analyze meaningful uses of "knowledge" and "knowing" in specific epistemic situations (empirical science, everyday interactions, psychological research, practical handiwork skills, mathematical knowledge, ...).

Crucially, Moore's provocation does eventually lead Wittgenstein to reconsider a key assumption of his grammatical analysis of "I know..." statements in OC, as Malcolm's conversation notes from 1949 indicate. Moore's *here is one hand* argument and list of truisms seems to provide exceptions to the grammatical rule that the ordinary use of "know" is connected to verificationism or "making sure". Moore's truisms present *experiential* propositions, but the *certainty* they are endowed with seems to afford them a special logical status. (Malcolm 2009: 73) Since Wittgenstein is clear that these hinges primarily represent experiential propositions, it appears implausible to call them rules of grammar, even though Wittgenstein *likens* them to rules of grammar. It is this logical status of *certainty*, which is not mere subjective *conviction*, that the *hinge analogy* in OC §341 drives at.

2. The Skeptical Challenge and Moore's Shift

Wittgenstein's sustained grammatical treatment of Moore's *here is one hand* argument is one of OC's main threads. Moore originally develops the argument

in “Proof of an External World” (1939) in response to Kant’s formulation of the question how the existence of “things outside of us” could be demonstrated. However, the argument applies to skeptical challenges more generally.

Using classical predicate logic, the *skeptical challenge* is usually represented as a generic modus tollens of the following form, where p is a random fact or piece of knowledge that is either true or false, and s is a skeptical illusion scenario that threatens my entire knowledge base, such as a Cartesian evil demon, the brain in the vat hypothesis, radical doubt about the continuing existence of unperceived objects, the idea that reality is just a dream, or the more recent simulation hypothesis:

(1) $I \text{ know } (p) \rightarrow I \text{ know } (\neg s)$

(Sk 2) $\neg I \text{ know } (\neg s)$

$\therefore$ (Sk 3) $\neg I \text{ know } (p)$

The skeptic grants that if I *know* that a certain fact obtains, e.g., that there are roses in my garden, it follows that I also know that I am not subject to a skeptical illusion, such as hallucinating about having a garden with roses. By fiat, a certain level of *epistemic closure* is assumed, according to which relevant knowledge can be obtained through logical entailment from other knowledge. The crucial premise is (2), according to which I can never know that I am not subjected to a skeptical scenario. The skeptic alleges that it is impossible to *prove* that what we believe to know about the outer world is not just some kind of illusion. This modus tollens argument can be applied to any piece of knowledge and therefore threatens to undermine the very possibility of knowledge.

Moore agrees with (1), but denies (Sk 2) and (Sk 3). The *Moorean Shift* consists in reversing the burden of justification, by using (1) and the negation of (Sk 3) in a modus ponendo ponens to deny (Sk 2). The relevant passage in the 1939 paper reads:

“How can I prove that there have been external objects in the past? Here is one proof. I can say: ‘I have held up two hands above this desk not very

long ago; therefore two hands existed not very long ago; therefore at least two external objects have existed at some time in the past, Q.E.D.” (Moore 2010: 148)

More formally, we can present the argument as follows, where *p* stands for certainties such as *here is one hand*:

(1) *I know (p) → I know (¬s)*

(M 2) *I know (p)*

∴ (M 3) *I know (¬s)*

The above reconstructions of the *skeptic’s challenge* and *Moore’s Shift* contain several implicit assumptions and gloss over countless historical nuances, but they provide a simple and clear picture of one of the central underlying *philosophical* conflict that OC sets out to *dissolve*.

Who is right: Moore or the skeptic? If one assumes that the logical inferences are solid – and there seem few reasons to deny this *prima facie* – then how one resolves this philosophical problem becomes a question of *plausibility*, and, ultimately, faith in the respective premises. The epistemological debate has reached a stalemate at this juncture. Philosophical “temperament” will decide whether one favors (Sk 2) or (M 2). The debate between skeptics and anti-skeptics is therefore not resolvable purely by means of logic; everything depends on which premises one is inclined to accept. Wittgenstein’s opening move in OC seems to reaffirm this: “If you do know that *here is one hand*, we’ll grant you all the rest.” (OC 1969: §1)

At this point, OC takes a step back to take a closer look at the *philosophical* quarrel presented above, and analyzes how the problem has been posed from a grammatical perspective. In order for classical predicate logic to be applicable, several conditions must be met, chiefly among them referentiality, sharpness of predicates, law of excluded middle, bivalence, consistency, and domain closure. But is “I know (*here is one hand*)” a sharp and truth-evaluable predicate – are the expressions given above well-formed? This is tied up with the question of the propositionality of the respective hinge propositions. As OC §5 indicates, the *propositionality* of any given sentence, including a hinge proposition such as (M2) is relative to a reference frame in a given language

game, to what I am willing to accept as *determinants* for the respective proposition. Various competing proposals exist in the HE literature for how to resolve the epistemological conflict, which I will briefly turn to in the following section.

3. The Propositionality of Hinges and Epistemic Closure

Where do hinges enter the philosophical debate outlined in section 2, and how should we understand their propositional status? A direct reading might treat (M 2) as *experiential hinge propositions* of the *here is one hand* variety. Statements falling under (M 3), such as “there is an external world”, are further candidates for hinges of a more general variety, commonly referred to as *über hinges*. The latter systematic conception of hinges is favored by many HE theorists. Annalisa Coliva argues that such hinges are “necessary in order for us legitimately to take perceptual experiences to bear onto beliefs about mid-size physical objects” (2016: 19), and Duncan Pritchard posits the *über hinge* commitment “that one is not radically and fundamentally mistaken in one’s beliefs”, from which other hinges are derived. (2012: 267)

One potential motivation for ascribing propositionality to *über hinges* is a desire to uphold *epistemic closure*: “How could one recognise that a certain historical event (e.g., the battle of Austerlitz) took place at such-and-such a date, and that this entails that the universe has been around for more than 5 min, and yet not adopt a positive propositional attitude [...] to the entailed proposition?” (Pritchard 2012, 266) Pritchard reacts to this puzzlement by developing an elaborate *no-belief* view, which renders our propositional attitudes towards hinges not as beliefs, but as special non-doxastic commitments. By contrast, Annalisa Coliva is willing to accept that the closure principle fails in a limited cases concerning hinges; even if we possess justification for statements of the type (M 2), there can none for (M 3) *qua hinge*. (2016: 20) Still, Coliva argues in favor of a conception of hinges as minimally truth-apt, such that hinges have semantic content and can be embedded in conditionals. (2016: 19)

These arguments in favor of propositionality are decisively rejected by Danièle Moyal-Sharrock. She grants that we are able to verbally formulate hinges for *heuristic purposes*, but emphasizes that hinges can’t be *said*. On her view, the logical relationship between hinges and rationally known propositions is that

between (grammatical) rule and propositions. (2022: 11) From this perspective, the problem of closure becomes irrelevant: since hinges are not genuine propositions, there is no real *inference* taking place in the case that Pritchard considers: “The so-called entailed proposition is not a proposition at all; it is the invisible hinge on which believing that the battle of Austerlitz took place in 1805 revolves.” (2022: 11) The radical consequence is that hinges fulfill no *justificatory* purpose. As arational preconditions of reasoning, they are completely immune to rational evaluation. For Moyal-Sharrock, the relevant questions concern the *status* and not the *content* of the hinges.

While there is much to be said in favor of Moyal-Sharrock’s arguments against treating *über hinges* as propositional and her careful attention to the *animal* in OC, I remain unconvinced that hinges express logical or grammatical rules. On my reading, grammatical rules in Wittgenstein’s sense guide meaningful use and establish certain connections between language and phenomena. They are, to use un-Wittgensteinian phrases, descriptions of processes of *meaning-making*, not *knowledge-making*. Following Coliva (2016), I believe that we should radically decouple questions of semantics and meaning from epistemological problems.

Moreover, in the context of OC, the (experiential) content of hinges does appear relevant for their privileged status in our thinking and acting. I remain skeptical that it is actually the case that a general *über hinge* such as “the external world exists and is older than 5 mins” underpins certainties such as *here is one hand*, the battle of Austerlitz took place in 1805, or “I see a tree over there?”, as all three leading HE proponents’ arguments considered above presuppose.

Rather, I’m inclined to see it the other way around: I do not have any good reason for doubting in the existence of an external world, or for formulating a philosophical tenet that the external world exists (as a *metaphysical* statement), because such general principles in no way make my *experiential hinges* more certain. The former seem, if anything, parasitic on the latter. For this very reason, *über hinge* generalizations and truisms cannot act as actual hinges in any meaningful sense. My practical knowledge is in no way affected by whether or not the skeptic is right about the world “actually” being a simulation, for example. In the context of *Moore’s Shift*, the conclusion (M 3)

takes the form of a philosophical or metaphysical claim, and Wittgenstein critically remarks throughout OC that Moore's argument does little to alleviate the skeptic's *metaphysical* doubts. (OC 1969: §19) Typically, we defer to basic certainties over abstract metaphysical claims in practical reasoning. If I were to suddenly and repeatedly experience weird and unusual sensory experiences, this might call my *animal* certainties into question.

While it is possible to systematically categorize various types of natural and acquired hinges (Moyal-Sharrock 2004: 106), it is impossible to give a comprehensive taxonomy of hinges that would cover all conceivable epistemic cases. This is a point that some of Wittgenstein's close students were adamant about: "One misunderstanding people make about the propositions which Wittgenstein says cannot be doubted is to try to classify them, or find out what they all have in common." (Rhees 2003: 78). There is no sharp demarcation between hinges and regular propositions. This anti-essentialism concerning hinges is made quite explicit in one of the final remarks published in OC:

"There are however, certain types of case in which I rightly say I cannot be making a mistake, and Moore has given a few examples of such cases. I can enumerate various typical cases, but not give any common characteristic. [...]" (OC 1969: §674)

Hinges are like rules of logic or grammar in that they seemingly stand firm. They are beyond any reasonable doubt because of what we already know about the world and our place in it. It is in this sense that hinge propositions express principles of human thought and research. But on my reading, they complement and do not necessarily constitute rules of grammar. What makes many of the propositions Wittgenstein is confronted with by G.E. Moore's list of certainties so strange from a grammatical perspective is that in certain cases they apparently seem to challenge the traditional bivalence and truth-evaluability of propositions Wittgenstein espouses in the *Tractatus*. As Wittgenstein remarks, I can pretty much use any *proposition* to derive any other propositions, but this does not necessarily amount to a proof. (OC 1969: §1) If this is the case, however, then what is the point of employing and relying on hinges in epistemology at all? Regardless of whether or not they can be

articulated propositionally, their main value lies in how well they shed light on the underlying and usually unquestioned tacit principles that structure our epistemic agency.

Outlook: Which Feminist Hinge Epistemology?

The field of HE continues to benefit the systematic reception of Wittgenstein in contemporary philosophy and stimulate a wealth of interesting work in applied epistemology. Recently, HE has been extended to new areas including political epistemology (Ranalli 2022), philosophy of technology (Hermann 2025), and feminist epistemology (Carmona & Gómez-Ledo 2024). A pluralist and open-ended understanding of hinges that moves beyond the search for *über hinges* promises a more nuanced framework for addressing epistemic problems in various domains. Recent contributions to HE by Coliva (2025) and Pritchard (2025) reflect a broader trend within HE towards more context-sensitive and piecemeal work.

In this paper, I have argued that the grammatical rule conception of hinges is untenable. My arguments against attempts to identify *über hinges* that universally underpin epistemic agency remain cursorily developed. To briefly recap: Section 1 underscored Wittgenstein's anti-essentialism and suggested that, historically, his primary focus likely lay in *experiential* hinge certainties. Section 2 presented generic logical reconstructions of the skeptical *modus tollens* challenge versus Moore's Shift towards a *modus ponens* argument. Section 3 considered the contributions of Pritchard (2012), Coliva (2016), and Moyal-Sharrock (2022), and discussed their respective treatments of hinges' propositionality and epistemic closure. Against these proposals, I stressed the significance of experiential hinges and argued that framework readings of OC inadequately capture hinges' function as principles of thought.

These reflections have not been motivated by Wittgensteinian purism, but rather by a desire to reanimate some of the historically grounded and underappreciated aspects of OC. Where Wittgenstein emphasized that knowledge is heterogenous, variegated and multifaceted, there is a temptation in systematic HE towards developing unifying theories of knowledge. While I am not opposed to such theorizing out of principle, it does not square neatly with OC's analyses into the grammar of "knowledge", "doubt", and "certainty".

A potentially liberating aspect of OC lies in its invitation to reflect on our lived epistemic practices, rather than construct grand epistemological theories. In this respect, HE would perhaps do well to pay closer attention to the epistemically less remarkable, more commonplace *experiential* hinges embedded in our everyday interactions with the world and other people. Feminist epistemology has provided valuable tools for studying how various social norms, institutions, and interactions come to bear in the construction of knowledge. From this perspective, what Wittgenstein has to say in OC appears very much in line with insights that feminist theorists have been trying to teach philosophers for a long time: What we (think we) know is significantly shaped both by contextual factors and our implicit values and actions. It is such structural features that need to be rendered visible in order for us to be able to critically discuss, challenge, and change them.

References

- Carmona, Carla; Gómez-Ledo, Ignacio (2024) "A Performative Feminist Hinge Epistemology: Making Room for Feminist Hinges", *Hypatia*, 1-22.
- Coliva, Annalisa (2016) "Which Hinge Epistemology?", in: Annalisa Coliva and Danièle Moyal-Sharrock (eds.) *Hinge Epistemology*, Leiden: Brill, 6-23.
- Coliva, Annalisa (2022) "Which Hinge Epistemology between Animal, Biscopic and Constitutivist?", in: Contantine Sandis and Danièle Moyal-Sharrock (eds.) *Extending Hinge Epistemology*, London: Anthem Press, 51-70.
- Coliva, Annalisa (2025) "Hinge trust", *Philosophy and Phenomenological Research*, 1-20.
- Malcolm, Norman (2009) *Ludwig Wittgenstein. A Memoir*, Oxford: Oxford University Press.
- Moore, G.E. (2010) "Proof of an External World", in: *Philosophical Papers*, Oxon/New York: Routledge, 127-150.
- Moyal-Sharrock, Danièle (2022) "Wittgenstein's Hinge Certainty", in: Contantine Sandis and Danièle Moyal-Sharrock (eds.) *Extending Hinge Epistemology*, London: Anthem Press, 3-26.
- Pedziwiatr, Samuel (2024) "Now You're Talking My Language. Wittgenstein Reads the *Cratylus*", in: Herbert Hrachovec & Jakub Mácha (eds.) *Platonism: Proceedings of the 43rd International Wittgenstein Symposium*. Berlin: De Gruyter, 457-474.
- Pritchard, Duncan (2012) "Wittgenstein and the Groundlessness of Our Believing", *Synthese* 189: 255-272.

Pritchard, Duncan (2025) "Axiological hinge commitments", *Synthese* 205: 49.

Ranalli, Chris (2022) "Political Hinge Epistemology", in: Contantine Sandis and Danièle Moyal-Sharrock (eds.) *Extending Hinge Epistemology*, London: Anthem Press, 127-147.

Rhees, Rush (2003) *Wittgenstein's On Certainty. There – Like Our Life*, Oxford: Blackwell.

Wittgenstein, Ludwig; Moore, G.E.; Malcolm, Norman; Citron, Gabriel (2015) "A Discussion Between Wittgenstein and Moore on 'Certainty' (1939): From the Notes of Norman Malcolm.", *Mind* 124 (493): 73–84.

Realist Intuitions in Standpoint Epistemology: Strategies for Reclaiming "Truth"

Victoria Pérez Monterroso (Madrid, Spain)

Abstract

In the following pages we will address an underexplored question within standpoint feminist epistemology: the epistemic and political viability of reclaiming the notion of truth. While standpoint theorists have reconceptualised related epistemic concepts, such as objectivity, truth has largely been stranded due to its association with value-neutrality and, therefore, political conservatism. On the contrary, we argue that a feminist reappropriation of truth is not only possible but also epistemologically and politically advantageous. Drawing on the works of authors such as Linda Martin Alcoff and Miranda Fricker, we will show that criticisms of truth within feminist epistemology primarily target the traditional notion of truth as correspondence, leaving open the possibility of alternative, non-naïve conceptions of the notion. Additionally, we retrace what we call a “minimal realist intuition” in the works of authors such as Sandra Harding and Donna Haraway, which should be rendered compatible with this rethinking of truth in feminist terms. Ultimately, we argue for a situated, non-neutral, and politically conscious understanding of truth that strengthens, rather than undermines, the core commitments of feminist standpoint epistemology, both epistemic and political.

The notion of truth has not been a focal point for standpoint feminist epistemology. Not only has it not been redefined, as has been the case with related concepts such as objectivity, but there is also a substantial body of feminist arguments against the reconceptualization of this notion. These arguments warn against such an endeavour by 1) asserting a radical and unbridgeable incompatibility between the concept of truth and certain conceptual innovations proposed by the standpoint epistemology -primarily the situated knowledge thesis-, and 2) anticipating the negative epistemological and political consequences that might result from this effort, a return to the “God's eye” perspective, to the understanding of knowledge as a the unbiased mirror of nature, that is: a revival of the problematic and outdated concept of truth as correspondence with reality, together with the underlying conservative political framework it entails. As Sandra Harding has noted: “The truth ideal in science supports tendencies toward inequality” (Harding 2006: 133).

Hence, we can point to only two notable attempts at a reconceptualization of this notion: the one undertaken by Linda Martín Alcoff (1987; 1996) and the one by Miranda Fricker (2007; 1994). As the former author claimed: “It is my

view that talk about feminism and social science needs to become (at least occasionally) talk about feminism and epistemology, and this must fundamentally become talk about truth. What do we mean by truth? What can we possibly mean?" (Alcoff 1987: 122). Building on the work of these authors, especially in that of the latter, our aim is precisely to answer this question: what would "truth" be in feminist terms? Or in other words, and more specifically, how to render a notion with such a questionable background compatible with the project of feminist standpoint epistemology?

We will argue that a reappropriation of the concept of truth in feminist terms is not only possible, but also significantly useful and beneficial for the tasks envisioned by feminist epistemology, both epistemological and political. Our goal, then, is to advocate for the viability of constructing a feminist notion of truth that does not necessarily revive outdated epistemological and political frameworks and ideas. We are confident in the plausibility of this task for two main reasons: 1) although, as we have mentioned, there are numerous arguments against such a task, these arguments target a very specific notion of truth, the classical account of truth as correspondence, and therefore do not apply to all possible notions of truth; and 2) as we will show, it is possible to trace within classic feminist standpoint epistemologists what we call a "minimal realist intuition", which suggests that the world imposes normative, negative limits on our beliefs: "nature constrains our beliefs" (Harding 1995: 338) and "[h]ow we interact with nature both enables and limits what we can know about it (Harding 2006: 141). A reconceptualization of truth through a feminist lens would enable us to justify and further substantiate this "minimal realist intuition", which we assert is present but not fully articulated (and occasionally unrecognized) in some foundational texts of standpoint epistemology. This is particularly evident in the works of Sandra Harding and, to some extent, in certain passages by Donna Haraway. By undertaking this endeavour (namely, by identifying these realist intuitions and associating them with a redefined feminist notion of truth) we could advance arguments supporting the proposition that there are "more or less truthful" theories about the world, rather than merely "less false" ones in the terms that, in alignment with Quine, Harding suggested. Our aim is to identify a conception of truth that aligns with the situated knowledge thesis, that is, one that is not defined in terms of value neutrality. This is crucial if we are to retain and

further enhance the significant role that political values play, and should continue to play, in feminist epistemology.

Our initial task then is to examine how the arguments advanced by standpoint feminist epistemology, while ostensibly trying to challenge the general notion of truth, are in fact employed with a particular focus on the concept most prevalent in the positivist philosophy of science: the notion of truth as correspondence with reality. To outline the issue, since the tradition of standpoint epistemology is broad, we will address the two main arguments put forward against the idea of restoring the notion of truth. While the first argument relates primarily to the epistemological aspects of the issue, the second argument is of a political nature.

According to the first argument, certain theoretical developments advanced by standpoint epistemology, particularly the situated knowledge thesis, render it impossible to reclaim a coherent concept of truth. Standpoint epistemology proposes that epistemic relations do not exist in a social vacuum. Rather, the social and political positions occupied by epistemic actors, entangled in power relations, affect their knowledge of the world. We are of course referring to the situated knowledge thesis, which states that knowledge, ethics, and politics are inextricably related to each other. "However, stances like Harding's go beyond asserting that knowledge is always socially situated. By proposing to "start thought from women's lives" the author states that not every social position render the same epistemic results. The fact that every epistemic position is constrained by a given social position and political commitment does not imply that any epistemic position becomes equivalent, or relative, to others. We are not condemned to epistemological relativism because more valuable accounts of the world, of nature, can be distinguished. That is, even though every epistemic position is dependent on a set of political and social values, we can still recognize less partial and perverse knowledge claims: those consciously guided by feminist political interests. Standpoint theory allows us then to articulate a project of epistemology in which political values do not appear as extrinsic elements that distort knowledge, but as key pieces that are "legitimate, and not merely unavoidable, components of theory-choice" (Alcoff 1987: 112). Standpoint epistemology unveils value-neutrality as the myth it is.

The question, then, is how does the concept of truth fit into this picture? As we have already mentioned, this kind of question has not even been proposed in the texts we are focusing on. Once it has been revealed that value neutrality is not a virtue, but an obstacle, and the concept of truth being traditionally linked to value neutrality, it seems that the only possible path is to leave it behind. If we maintain a conception of truth as correspondence, as a mirror of nature (a conception held by authors such as Harding), it is clear that we cannot conceive of political values as anything other than a distortion. As we have seen, this understanding of value as distortion would be entirely incompatible with the project outlined above, for we would be failing to adhere to the situated knowledge thesis.

Nevertheless, we will see that another approach is possible if we examine the way in which Harding herself manages to reclaim the notion of objectivity, which has also traditionally been understood in terms of value neutrality. However, we will deal first with the second type of argument put forward against the idea of restoring the notion of truth. As previously mentioned, the second argument is of a political nature and claims that making use of the notion of truth would lock us irremediably on the road to inequality: "Any truth ideal that insists on eliminating cultural elements from science, let alone on refusing to critically examine the cultural elements in one's own society's sciences, and especially those that have been identified by others, works against the nourishment of democratic social relations to which the institutions of science are supposed to be committed. Insistence on or assumption of such a truth ideal advances inequality" (Harding 2006: 143).

Yet, we ask, is this the only possible notion of truth? Can we examine and embrace that the social and political elements of our epistemic relations are not distortions, but virtues, and continue to maintain a notion of truth that moves us away from, rather than towards, inequality? We firmly believe that the answer must be affirmative for we can find passages in which even Harding seems to acknowledge that other uses of the notion of truth are possible: "For example, feminists must still struggle mightily to get recognition from dominant institutions of women's actual living conditions (...). Victims of racial discrimination continue to have to fight for similar recognitions. It can seem crucial to these groups to be able to claim that their analysis are objectively true, not just their own particular perceptions" (Harding 2006: 144).

Thus, the political argument that truth leads irremediably to inequality, proposed by Harding, seems to be countered if we consider this possible social/political use of truth, which, curiously enough, is also proposed by Harding herself. This alternative conception of truth is in accordance with that put forward by Fricker, to which we will return later.

First, we must draw some brief conclusions about the two main arguments raised against the reappropriation of truth that we were discussing. Starting with the second argument, we consider it loses its strength as soon as we manage to find a definition of truth that is not automatically aligned with inequality. And as we have seen, both Harding and Fricker put us on track that this task might be possible. As for the first argument, it is clear that defining truth in terms of value neutrality would be incompatible with the situated knowledge thesis. However, since other uses and definitions of the concept are possible, it should be clear that this argument would only be relevant if we were to maintain the traditional concept of truth as correspondence, which is not our aim. We could then seek other possible uses of truth that are compatible with conscious commitment to feminist political values.

In fact, this is precisely the task that Harding herself sets out to accomplish, not with respect to the notion of truth, but with respect to objectivity. Although the positivist framework has been persistent in defining objectivity in terms of value neutrality, Harding considers that, given its great importance for the disciplines of epistemology and philosophy of science, the re-signification of this notion is possible, "if we move beyond many of the distorting features of modernity's conceptual framework" (Harding 1995: 346). At this point, it seems unjustified to maintain this position with respect to the notion of objectivity, but not with respect to the notion of truth, of equal or even greater importance than the former. As we advanced above, if we examine the way in which Harding redefines the concept of objectivity, we can see that all the arguments proposed by the author in favour of this task (despite the questionable framework of the notion) could be applied in the same way in favour of the endeavour of reclaiming the notion of truth. In conclusion, it seems that the two arguments presented against this task, while trying to challenge the general notion of truth, are in fact only applicable to the notion of truth as correspondence.

Having dealt with this first issue, we shall turn to what we called the question of the “minimal realist intuition” that we said could be traced in some standpoint epistemology foundational texts.

If standpoint epistemologists are reluctant to reclaim the notion of truth because of the possible epistemic and political consequences that this might entail, we can assume that, in the same way, we won't find an explicit defence of realism at any point in the work of these authors. Realism in the naive sense and truth as correspondence usually go hand in hand, so if the latter is rejected outright, the former is also rejected in the same way. However, we would like to argue in favour of the idea that, although authors such as Harding and Haraway do not explicitly advocate realism, we can trace non-naive realist intuitions in their texts. In Haraway's words “The social justice movements need to have simultaneously an account of radical historical contingency for all knowledge claims and knowing subjects, a critical practice for recognizing our own ‘semiotic technologies’ for making meanings, and a no-nonsense commitment to faithful accounts of a ‘real’ world” (Haraway 1998: 579). If we turn to Harding's proposal, we can find several passages similar to the following: “There are ‘facts’ of nature and social relations against which hypotheses always must be tested through rules of good reasoning about observations, though just what are those facts and rules is always in principle and often in practice open to debate” (Harding 2001: 512). This realist intuition we are retracing does not only appear in the above-mentioned, more classical authors from standpoint epistemology. It can also be found in texts by more recent authors, such as those of Fricker and Alcoff. While Fricker states that “a (non-naive) realist account of empirical belief must be sustained” (Fricker 1994: 95), Alcoff cautions us against underestimating “the significance of no longer being able to claim correspondence for any scientific theory or claim” (Alcoff 1987:116).

However, while Fricker and Alcoff explicitly advocate for realism, and argue that it is not incompatible with the core of standpoint epistemology, this is not the case in Harding's and Haraway's texts, where terms such as “real world” and “nature” appear always in quotation marks. As a result, it can be noted that realism is present, but not fully articulated. This strategy (using quotation marks around these questionable terms but still employing them) can be problematic and is ultimately unnecessary since non-naive versions of realism,

like non-naïve versions of the notion of truth, are not, as we suggest, incompatible with the commitment to feminist standpoint epistemology. Rather, we actually need this commitment to realism if we are to defend our feminist political commitments, as Fricker noted: "We must have a realist account of empirical belief, for in such discourse the political presupposes the empirical as a normative constraint. Moreover, I have made the stronger claim that if we were to absolve ourselves of the need for realism by dropping this normative constraint, we would not be able to make adequate sense of our political commitments" (Fricker 1994: 100). As already mentioned, Fricker's proposal is consistent with the possible social/political use of truth we outlined earlier following Harding, which would account for this fully embraced non-naïve kind of realism.

In essence, we have argued for the viability of constructing a feminist notion of truth that does not revive epistemological and political frameworks that are outdated as they are harmful. We have done so by showing that although some foundational texts of standpoint epistemology are full of arguments against this endeavour, they are directed towards a very specific notion of truth, that of correspondence. We have also traced a "minimal realist intuition", or a realism of the non-naïve variety, in both Harding's and Haraway's work, to which a renewed notion of truth in feminist terms will provide depth, by acknowledging it head-on and also linking it to our feminist political commitments.

It still remains to be made clear why this reclaiming of truth is useful and beneficial for the tasks envisioned by feminist standpoint epistemology, both epistemological and political. On the epistemic front, a return to these specific, non-naïve types of realism and truth allows us to pursue knowledge claims that, while not strictly corresponding to reality or reflecting it like a mirror, are better accounts of it, being more truthful (nor merely less false) and less partial, distorted and perverse than those that do not have feminist interests at their core or that, in general, advocate for value neutrality. In Harding's words, we would achieve representations that "effectively map 'the world out there' and yet are also fully part of the culture" (Harding 2001: 513).

As for the political dimension and acknowledging Harding's observation regarding the potential social/political uses of the notion of truth, it becomes

evident that the conservative political framework is not the one tied to truth exclusively. It is indeed possible to reclaim the concept of truth for it to align with feminist commitments. As Harding asserted, “[t]his issue about the ‘politics of truth’ should not be ignored” (Harding 2006: 143), but while she proposed this statement to warn us about how truth as correspondence sets us on the road to inequality, we build on it to proclaim a different type of warning: we should not ignore the issue about the “politics of truth” because we risk leaving behind the great potential it would bring to feminist politics to uphold a renewed ideal of truth, that would go hand in hand with the defence of a non-naïve kind of realism. We shouldn't underestimate both the epistemic and political potential of possessing a set of realist arguments against rival, non-feminist accounts of the world out there.

References

- Fricker, Miranda. (2007) *Epistemic injustice*. Oxford University Press.
- Fricker, Miranda. (1994) “Knowledge as construct: Theorizing the role of gender in knowledge”, in: Lennon, Kathleen & Whitford, Margaret (eds.) *Knowing the Difference*, London: Routledge, 95-109.
- Haraway, Donna. (1998) “The science question in feminism and the privilege of partial perspective”, *Feminist Studies* 14 (3), 575-599.
- Harding, Sandra. (1995) “Strong objectivity. A response to the new objectivity question”, *Synthese* 104, 331-349.
- Harding, Sandra. (2001) “Comment on Walby's "Against Epistemological Chasms: The Science Question in Feminism Revisited": Can Democratic Values and Interests Ever Play a Rationally Justifiable Role in the Evaluation of Scientific Work?”, *Signs* 26 (2), 511-525.
- Harding, Sandra. (2006) “Are truth claims in science dysfunctional?”, in: Harding, Sandra. *Science and Social Inequality: Feminist and Postcolonial Issues*. University of Illinois press, 133-144.
- Martín Alcoff, Linda. (1987) “Justifying Feminist Social Science”, *Hypatia* 2 (3), 107-127.
- Martín Alcoff, Linda. (1996) *Real Knowing. New Versions of the Coherence Theory*. Cornell University Press

Was Wittgenstein Aiming at an Identity Theory of Truth?

Jimmy Plourde (Trois-Rivières, Canada)

Abstract

In *Wittgenstein's 'Notes on Logic'*, Michael Potter claims that Wittgenstein's picture theory was aiming to be an identity theory of truth. His view is partly based on his conjecture that this was the result of the influence of Frege's critique of the correspondence theory, a critique that stems from the 1897 draft of "Der Gedanke". Though sympathetic to Potter's conjecture and to his claim that it requires revision of our current understanding of the *Tractatus* conception of truth, I argue in the present paper that Potter's view is ill-founded. I then draw attention to remarks of the *Tractatus* which suggest an alternative outcome: Wittgenstein adopted a view of correspondence that is based on identity, and which is able to uphold Frege's critique of the correspondence theory.

1. Introduction

In *Wittgenstein's Notes on Logic*, Michael Potter addresses the issue of Wittgenstein's conception of truth. His view on the matter is based on a conjecture that has been hitherto neglected. According to Potter Wittgenstein knew Frege's critique of the correspondence theory, agreed with it, and thus ended up with the picture theory aiming to be an identity theory of truth:

It might well be argued then, that what is essential to the picture theory of the *Tractatus* is not its appeal to the picturing analogy on its own, but rather the manner in which the theory takes to heart Frege's admonition that correspondence comes in degrees but truth does not. What is essential, that is to say, is that it aims to be an identity theory, not a correspondence theory" (Potter 2011: 230)

The admonition Potter is here referring to is presented at length in the following passage of "Der Gedanke":

It might be supposed (...) that truth consists in the correspondence of a picture with what it depicts. (...) A Correspondence, moreover, can only be perfect if the corresponding things coincide and so just are not different things. (...) But this is not at all what people intend when they define truth as the correspondence of an idea with something real. For in this case it is essential precisely that the reality shall be distinct from the idea. But then

there can be no complete correspondence, no complete truth. So nothing at all would be true; for what is only half true is untrue. Truth does not admit of more and less. (Frege 1984: 352-3)

Despite the fact that “Der Gedanke” was published only in 1919, Potter points out that “the distinction Frege draws here between propositions and pictures is already present in 1897 draft of ‘Der Gedanke’” (Potter 2011: 230). Thus, Wittgenstein could arguably have come to know Frege’s point directly from him, something Potter believes is likely to have happened during Wittgenstein’s 1912 December visit to Frege. If this is right, standard interpretations of the tractarian conception of truth as a correspondence theory ought to be substantially revised.

But is it right? Could Wittgenstein have been so deeply influenced by Frege’s point that he was aiming with the picture theory at an identity theory of truth and should we therefore substantially revise our current understanding of the *Tractatus* on this issue? In order to address this question, I shall briefly present Potter’s case for his reading (2). I shall then argue that the thesis of identity of structure that Potter considers to be the basis of Wittgenstein’s account of truth can’t qualify as an identity relation of truth (3), and that Potter’s reading rests on inconclusive evidence (4). I then claim that Wittgenstein might nonetheless have been aware of Frege’s critique and that it might have played a role in the shaping of his conception of truth by helping him to account for correspondence in such a way that it is partly based on identity and is able to face Frege’s critique (5). In so doing, I shall limit myself to Potter’s reading and ignore other identity theory readings of Wittgenstein such as Johnston and Sullivan’s (2018).

2. Potter’s case for his Identity Theory Reading

Besides the historical conjecture according to which Wittgenstein knew of Frege’s admonition from one of his visits to Frege, Potter’s case for his thesis that Wittgenstein was aiming at an identity theory of truth rests on two pillars. The first one is the conjunction of the identity of structure thesis (IST) and the thesis that IST constitutes a central element of the picture theory. IST is the thesis that every symbolizing atomic fact and the fact it symbolizes have an identical structure. It was embraced by Russell in his Harvard lectures when he said “the structure of the symbol must be identical with the structure of the

symbolised” (quoted in Potter 2011: 225), and the picture theory would have been brought in by Wittgenstein essentially to account for this identity. This is why IST is explicitly endorsed in the *Tractatus* within the context of the picture theory and constitute one of its most basic tenets:

That the elements of the picture are combined with one another in a definite way, represents that the things are so combined with one another. (TLP: 2.15)

In order to be a picture a fact must have something in common with what it pictures.

In the picture and the pictured there must be something identical in order that the one can be a picture of the other at all. (TLP: 2.16-2.161)

According to Potter the fact that the picture theory is based on IST shows that it is aiming to account for truth in terms of identity and it therefore must be this idea of structural identity Wittgenstein had in mind when he said that the picture theory could “yield the nature of truth straight away” (NB: 7(4)).

Though Wittgenstein did not explicitly state the account of truth yielded by a picture theory based on IST, Potter considers that the picture theory nonetheless was meant to be an identity theory account of truth. Potter argues that we do have textual evidence that this is what Wittgenstein was aiming at in at least two different remarks of the NB. According to him Wittgenstein seems:

to have become quite clear that the representational form of the propositional sign does not merely correspond, but is identical, to the situation it represents (NB 18 Dec. 1914). ‘Non-truth,’ he noted some months later, ‘is like non-identity (NB, 27 Apr. 1915).’ And he could presumably have added, truth is like identity. (Potter 2011: 230).

Thus, the thesis that the picture theory essentially consists of IST constitutes the kernel of Potter’s case for his reading and the two remarks here mentioned in the quote constitute the textual evidence in favor of this reading of IST.

The second pillar is the rejection of what Potter calls the “eye-catching claim” of the picture theory. Basically, this is the view that the picture theory literally

says that a proposition is a picture, i.e. that it consists of representing some things being in a certain relation in virtue of itself consisting of its constituents being in that same relation. When coupled with Wittgenstein's understanding of sense as "a standard to which facts behave", something to which the reality may agree or disagree, it strongly suggests that Wittgenstein thought of the picture as something that may or may not correspond to reality and, therefore, that he conceived truth in terms of correspondence. But the eye-catching claim is mistaken according to Potter (Potter 2011: 228). Potter argues that the eye-catching claim thinks of pictures exclusively in spatial terms, but Wittgenstein was clear that pictures are logical pictures and not spatial pictures. Moreover the idea of a propositional sign's sense that would be a standard to which reality could be compared is wrong according to Potter and was not Wittgenstein's position in the *Tractatus*:

What a proposition expresses is already true or false, and if I assert the proposition, what I assert is just that -what it expresses. There cannot be an intermediate stage in the explanation where a proposition represents a possible situation while yet remaining neutral as to whether it represents as obtaining or not (Potter 2011: 229).

Thus, interpretations of the *Tractatus* being committed to a correspondence theory of truth are, according to this last pillar of Potter's position, derived from erroneous understandings of the picture theory and propositional meaning. Therefore, though it has not been explicitly stated, Wittgenstein was in the *Tractatus* aiming at an identity theory of truth and was not embracing a correspondence theory, and this, according to Potter, would be the result of Wittgenstein taking seriously Frege's admonition that truth does not come in degrees, i.e. it would be the result of the influence of one key argument in Frege's critique of the correspondence theory of truth.

3. Accounting for the aimed Identity theory of truth in terms of IST

Potter's view that IST is a central element of the picture theory is widely acknowledged. That this should be considered as a significant step towards an identity theory of truth is however much more controversial.

In order for IST to qualify as an identity theory of truth, it has to be committed to a relation of identity that obtains between *a sentence's meaning* and *a fact* (Baldwin 1991). Yet IST asserts identity between the structure of a symbol and

the structure of a fact. The identity of structure Wittgenstein recognizes at 2.15-2.161 does not therefore *per se* qualify as the identity relation of the identity theory of truth.

Now, this may seem not to be deeply problematic for Potter's reading. After all, Potter did not hold that Wittgenstein was defending a classic identity theory of truth and though the basic idea of an identity theory of truth is that 'to be true' just is for a sentence's content to be identical to how things are, this content does not have to be only conceived in terms of the sentence's meaning (e.g. Hornsby 2001). Couldn't one therefore argue that Wittgenstein was aiming at what was in fact a version of the identity theory of truth that conceived truth in terms of an identity relation between the structure of the symbolising fact and the structure of the symbolised fact?

There are however two important problems with this. First, contrarily to the relation of identity in any variant of an identity theory, the relation '... has the same structure as ...' is not a one-one relation, but a many-many relation. Indeed, 'aRb' has the same structure as the fact that *aRb*, but it also has the same structure as the fact that *aRc* and, say, the fact that *cRd*. But, of course, what 'aRb' says cannot be identical to the fact that *aRc*. Thus, IST's identity of structure between symbolizing facts and symbolized facts is a weaker relation than the identity relation in an identity theory of truth and it cannot, for that reason, be what truth consists of. In fact, it is, by itself, not sufficient to account for the truth of a proposition.

The other problem for Potter's position is that identity, in any variant of an identity theory, is taken as *numerical identity*, i.e. within such a theory, the *relata* of the identity relation turn out to be a single existing entity. Yet, the identity of structure Wittgenstein acknowledges is an identity of the structures of two different entities: the symbolizing fact and the symbolized fact. Unlike a sentence's content and a fact in an identity theory framework, the symbolising fact and the symbolized fact do not have the same constituents, and since the structure of anything consists in the arrangement of its constituents (*TLP*: 2.032, 2.15), the structures of these two facts are numerically distinct. They might be considered as two different tokens of one and the same structure type, but they can't be a single particular real entity as a sentence content and a fact are taken to be in an identity theory of truth. Now, if this is true, then

any attempt to construe IST as some variant of the identity theory of truth must fail.

4. Textual evidence examined

Even if IST does not qualify as an identity theory of truth, one may nonetheless hold that textual evidence Potter relies on might be sufficient to grant his claim that Wittgenstein was aiming at an identity theory. There are however important problems with this view.

The first one is that textual evidence in favor of Potter's reading is faulty. As a matter of fact, the *Notebooks* remark that has been translated by Anscombe as "Non-truth is like non-identity" reads "*Das Nicht-stimmen ist ähnlich wie Nicht-Identität*" (NB: 43(12)). Thus, it's not about "Non-truth", but about "incorrectness" ("*Nicht-stimmen*"), a term which is often used by Wittgenstein to characterize truth as correspondence (NB: 29(9), 33(11), 34(2), 38(3-4)). Regarding the remark of December 1918 where Wittgenstein is said to have acknowledged the identity of the representational form and the represented situation, the truth is that he only writes that the identity of the propositional sign and the situation consists in the fact that they have the same multiplicity. That this was intended by Wittgenstein to mean that the representational form is numerically identical to the situation and that this was his final position on the issue is clearly dubious, for only two days later he wonders how a situation may *correspond* to the propositional sign "p" and not correspond to the propositional sign "~p" (NB: 37(4)).

The second problem is that there is textual evidence in pre-*Tractatus* writings that Wittgenstein rejected the identity theory of truth, something Potter simply ignores. Thus, for instance, in the "Notes on Logic", Wittgenstein writes: "There are positive and negative facts: if the proposition 'This rose is not red' is true, then what it signifies is negative. (...) *Positive* and *negative* facts there are, but not true and false facts." (NB: 97). Rejecting true and false facts is an explicit rejection of a central component of Russell's identity theory of truth. Thus, though it seems to be the case that Wittgenstein once did consider the identity theory as a possible account for the nature of truth, it was to reject it.

Finally and alternatively, even if the analogy of the picture which suggests a correspondentist account was to be "kitsch", there is also ample evidence, *throughout* the pre-*Tractatus* writings and the *Tractatus*, that Wittgenstein did

conceive truth as a relation, and that he explicitly understood this relation as a correspondence relation (Plourde 2024) something which, again, Potter simply ignores.

5. Conclusion

What should we think of Potter's conjecture? Though the distinction between picture and proposition on which "Frege's admonition" is based is present in the 1897 draft of "Der Gedanke", the point itself is not there (Frege 1976). Moreover, as Wolfgang Künne has pointed out, Scholz mentions in his compilation of Frege's correspondence a letter written by Wittgenstein in October 1913 containing "important arguments against Frege's theory of truth" (Künne 2009: 28). Though this confirms that Frege and Wittgenstein talked about truth before Wittgenstein could read "Der Gedanke", it does not fit squarely with Potter's conjecture.

Looking more closely at Frege's critique, there is a point against correspondence which have gone unnoticed, but which seems to have found an echo in the *Tractatus*. According to Frege:

It is supposed to be possible to test the genuineness of a bank-note by comparing it stereoscopically with a genuine one. But it would be ridiculous to try to compare a gold piece stereoscopically with a twenty-mark note. It would only be possible to compare an idea with a thing if the thing were an idea too. (Frege 1984: 353)

The interesting point here is that Frege allows the possibility of correspondence under the sole condition that the corresponding elements be qualitatively identical in the way in which two bank notes may be said to be identical. What rules out such a possibility for a picture theory of correspondence however is that, being of different natures, sentences may not be pictures corresponding in this sense to real entities.

Now, immediately after saying that the proposition is a picture of reality (*TLP*: 4.01), Wittgenstein acknowledges that a proposition does not at first appears "to be a picture of reality", before adding: "these sign-languages prove to be pictures, even in the ordinary sense, of what they represent" (*TLP*: 4.011). This

seems to be a rejection of Frege's point that sentences cannot be considered as pictures of reality because they are of a different nature than facts. But what remains unclear here is how this could be possible.

The following remarks might indicate the basis of Wittgenstein's solution:

A picture depicts reality by representing a possibility of existence and non-existence of states of affairs. (*TLP*: 2.201)

The agreement or disagreement of its sense with reality constitutes its truth or falsity". (*TLP*: 2.222)

As a matter of fact, if the sense of a proposition such as 'aRb' is the possible situation that it represents (*TLP*: 4.031) or the possibility of existence and non-existence of states of affairs (*TLP*: 2.201+2.202) and if the depicted reality is the fact that aRb, or the actual existence or non-existence of states of affairs (*TLP*: 2.06), then, arguably, the represented situation and the depicted fact are like two qualitatively identical bank-notes. They both have the same constituents and structure, but since one is merely possible, while the other is actual, they are not numerically identical (Plourde: 2017). This would allow Wittgenstein to account for truth in terms of correspondence on the basis of Frege's suggestion without facing the problems of the difference of nature of the corresponding elements. According to that view, 'to be true' just consist for the sentence to correctly depict reality (*TLP*: 2.21), i.e. it consists of representing a situation (*TLP*: 2.201) which agrees to the depicted fact (*TLP*: 2.222) by being qualitatively identical to it.

Should this be accurate, then Potter would be right to say that Wittgenstein was aware of Frege's critique of the correspondence theory before 1919 and sense and the depicted reality.

References

Baldwin, Thomas (1991) "The Identity Theory of Truth", *Mind*, 100: 35–52.

Frege, Gottlob (1976) "Logik", in: *Nachgelassene Schriften und Wissenschaftlicher Briefwechsel*, vol. 1, Hamburg: Meiner, 137-163.

Frege, Gottlob (1984) "Thoughts" in: Brian McGuinness (ed.) *Collected papers on mathematics, logic, and philosophy*, Oxford: Blackwell, 351-372.

Hornsby, Jennifer (2001) "Truth: The identity theory" in: Michael P. Lynch (ed.) *The Nature of Truth*, Cambridge: MIT Press, 663–681.

Johnston, Colin & Sullivan, Peter (2018) "Judgements, facts and propositions: theories of truth in Russell, Wittgenstein and Ramsey" in: Michael Glanzberg (ed.) *The Oxford Handbook of Truth*, Oxford: OUP, 150-192.

Künne, Wolfgang (2009) "Wittgenstein and Frege's *Logical Investigations*", in: H.-J. Glock & J. Hyman (eds) *Wittgenstein and Analytic Philosophy*, Oxford: OUP, 26-62.

Plourde, Jimmy (2017) "Wittgenstein's Picture Theory and the Distinction between *Representing* and *Depicting*", *International Journal of Philosophical Studies* 25, 16-39.

Plourde, Jimmy (2024) "Truth and Correspondence in the *Tractatus*", *Synthese* 204, 1-27.

Potter, Michael (2011), *Wittgenstein's 'Notes on Logic'*, Oxford: OUP.

Wittgenstein, Ludwig (1979) *Notebooks 1914-1916*, 2nd ed., G. H. von Wright & G. E. M. Anscombe (eds.), with an English translation by G. E. M. Anscombe, Chicago: Chicago University Press. (NB)

Wittgenstein, Ludwig (1998) *Logisch-philosophische Abhandlung/Tractatus Logico-philosophicus*, B. McGuinness & J. Schulte (eds.), Frankfurt: Suhrkamp. (TLP)

Die Frage nach der Lebensform – Wittgensteins Begriff der Lebensform und die Begriffe Lebensmuster, Lebensgepflogenheit, Lebensschablone, Lebensteppich auf dem Prüfstand und ein religionsphilosophischer Exkurs

Moritz René Pretzsch (Kassel/München, Deutschland)

Abstract

Der Begriff der Lebensform polarisiert in der Wittgenstein-Forschung. Den einen gilt er als besonders charakteristischer Begriff für Wittgensteins Denken und wird mitunter sogar als *terminus technicus* seiner Philosophie verstanden (siehe etwa Garver 1984; Haller 1984; Savigny 1999; Glock 2010). Andere kritisieren diese Einschätzung und bestreiten, dass der Begriff Lebensform in der gegenwärtigen Wittgenstein-Forschung überhaupt richtig verstanden wird (siehe etwa Schulte 1999; Majetschak 2010, Moyal-Sharrock 2015, Martin 2018 und Trächtler 2023). Im Zentrum der gegenwärtigen Debatte scheint weniger die alte Frage zu stehen, ob wir Wittgensteins Begriff der Lebensform pluralistisch denken müssen, sondern mehr die Frage zu polarisieren, ob der Begriff der Lebensform, seiner Standardinterpretation gemäß, überhaupt mit sozio-kulturellen Systemen, in denen eine Sprache nebst ihrer Sprecher eingebettet ist, kurzgeschlossen werden darf. Dieser Beitrag möchte den Begriff Lebensform wie seine Nachbarbegriffe Lebensmuster, Lebensschablone und Lebensteppich zunächst genauer in den Blick nehmen und dabei – in streng exegetischem Prüfverfahren – anknüpfend an Wittgenstein auch die Frage diskutieren, auf welche Phänomene der Begriff Lebensform resp. Lebensmuster überhaupt angewendet werden kann, welches Potential in ihm angelegt ist und wie wir auch gegenwärtig philosophisch von ihm profitieren können – eine Frage die in der Forschung mit wenigen Ausnahmen oft außer Acht gelassen wird. Hierbei möchte ich insbesondere aus religionsphilosophischer Perspektive die Frage diskutieren, ob wir exemplarisch im Fall des problematischen und kontroversen Begriffs Religion den Begriff des Lebensmusters erfolgreich anwenden können.

1. Problemexposition

In den *Philosophischen Untersuchungen* betont Wittgenstein in Auseinandersetzung mit dem Begriff des Sprachspiels, dass „das Sprechen der Sprache ein Teil [...] einer Tätigkeit, oder einer Lebensform“ (PU: §23) ist. Der Begriff Lebensform bei Wittgenstein wird seiner Standardinterpretation (siehe hier exemplarisch Savigny 1999 und Glock 2010) gemäß als „eine Kultur oder soziale Formation“ verstanden, die „die Gesamtheit gemeinschaftlicher Tätigkeiten“ von Menschen umfasst, in die wiederum die Mannigfaltigkeit der „Sprachspiele eingebettet“ ist (Glock 2010: 200). Anders gesagt, Lebensformen werden als sozio-kulturelle Systeme gedeutet, in welche eine bestimmte Sprache samt ihrer Sprecher jeweils eingebettet ist (vgl. Savigny 1999: 136). Und damit wird eine „Lebensform“ letztlich als eine Kultur gedeutet, die

spezifische, in verschiedenen Kulturen ggf. unterschiedliche Sprachgebräuche inkludiert. Allein die Annahme, dass der Begriff der Lebensform – welcher von Wittgenstein an keiner Stelle näher expliziert wird – ebenso grundlegend für Wittgensteins Spätphilosophie sein soll wie der Begriff Sprachspiel, darf doch stark in Zweifel gezogen werden, kommt er etwa in den *Philosophischen Untersuchungen* lediglich dreimal vor. So heißt es:

Man kann sich leicht eine Sprache vorstellen, die nur aus Befehlen und Meldungen in der Schlacht besteht. – Oder eine Sprache, die nur aus Fragen besteht und einem Ausdruck der Bejahung und der Verneinung. Und unzählige Andere. – Und eine Sprache vorstellen heißt, sich eine Lebensform vorstellen. (PU: §19)

Das Wort „Sprachspiel“ soll hier hervorheben, daß das Sprechen der Sprache ein Teil ist einer Tätigkeit, oder einer Lebensform. (PU: §23)

„So sagst du also, daß die Übereinstimmung der Menschen entscheide, was richtig und was falsch ist?“ – Richtig und falsch ist, was Menschen *sagen*; und in der *Sprache* stimmen die Menschen überein. Dies ist keine Übereinstimmung der Meinungen, sondern der Lebensform. (PU: §241)

Problematisch ist die Standardinterpretation schon allein deshalb, da sich Deutung von Lebensformen als Sprachen einbettende sozio-kulturelle Systeme nicht durch andere Wittgensteinsche Verwendungen des Begriffs „Lebensform“ im Nachlass gestützt sieht (sieht man lediglich von einer einzigen Stelle ab (vgl. EPB: 202)). Man beachte, wie oft und geradezu leichtfertig der Lebensformbegriff in Untersuchungen mit dem Begriff Kultur kurzgeschlossen wird. Zudem tritt erschwerend hinzu, dass sich die Standardinterpretation auch nicht durch Bemerkungen Wittgensteins abgesichert weiß, in welchen er tatsächlich darüber nachdenkt, worin bestimmte Ausdrücke unserer Sprache – etwa „psychologische Begriffe“ wie „Hoffnung“, „Glaube“ oder „Kummer“ – eingebettet seien. Zwar verwendet Wittgenstein in diesen Kontexten tatsächlich die Formulierung „eingebettet“ (vgl. BPP: 16) nicht aber – wie man es von der Standardinterpretation erwarten dürfte – den Begriff Lebensform. Statt der Standardinterpretation muss dementsprechend ein anderes Verständnis des Begriffs Lebensform gelten. Eben dieses „neue“ Verständnis möchte ich, nebst

seinen familienähnlichen Begriffen, Lebensmuster und Lebensschablone, in diesem Beitrag näher diskutieren. Um anschließend, aus religionsphilosophischer Perspektive, seinem aktuellen philosophischen Potential nachzuspüren. Exemplarisch werde ich hierbei den bis heute problematischen und teils stark kontroversen Begriff der Religion in Augenschein nehmen.

2. Erste Zurückweisung der Standardinterpretation

Im Nachlass heißt es an einer Stelle, an der Wittgenstein über unterschiedliche Sprachen in interkultureller Perspektive nachdenkt:

Kämen wir in ein fremdes Land mit fremder Sprache & fremden Sitten, so wäre es [...] in manchen Fällen leicht, eine Sprach- & Lebensform zu finden [...], die wir Befehlen & Befolgen zu nennen hätten, vielleicht aber besäßen sie keine Sprach- & Lebensform, die ganz unserm Befehlen etc. entsprächen. So wie es vielleicht ein Volk gibt, das [...] keine unserm Gruß entsprechende Lebensform besitzt. /das nichts unserm Grüßen entsprechendes besitzt/ (MS 165: 110f.)

Auffällig an dieser Textstelle ist, dass Wittgenstein „Befehlen“, „Befolge von Befehlen“, „Gruß“ und „Grüßen“ explizit als Sprach- und Lebensformen deklariert. Lebensform, so zeigt sich an dieser Textstelle, meint etwas ganz anderes, als es die Standardinterpretation vorsieht. Bei dem Begriff Lebensform ist *nicht* das Sprachen und Sprecher einbettende sozio-kulturelle System gemeint, sondern es scheint vielmehr von Lebensmustern die Rede zu sein. Deutlich wird das, wenn Wittgenstein in einer anderen Version einer Bemerkung, in welcher von „Lebensform“ die Rede ist, statt „Lebensform“ (MS 144: 1) auch „Lebensmuster“ (LSPP: 365) schreibt. Stefan Majetschaks Deutungsansatz ist zuzustimmen: Wenn Wittgenstein in seiner Spätphilosophie von „Lebensform“ spricht, so ist damit nichts anderes als „Lebensmuster“ gemeint (vgl. Majetschak 2010 und 2020). Wie aber muss der Begriff Lebensmuster verstanden werden?

3. Betrachtung der Begriffe „Lebensform“, „Lebensmuster“, „Lebensschablone“ und „Lebensgepflogenheit“

Insbesondere in seinen Überlegungen zum „Lebensmuster“ zeigt sich einmal mehr Wittgensteins ganzer philosophischer Anspruch nach Klarheit und Übersicht. Denn gerade in seinem Nachsinnen über „Lebensmuster“ wird die

gewaltige philosophische Anstrengung überdeutlich, dem Leben selbst als „Getriebe“ (BPP: 625), als unregelmäßiges, in verschiedenen Rhythmen pulsierendes, „undurchschaubare[s] Gewirr von mannigfaltigen menschlichen Handlungsweisen“ (Schulte/Majetschak 2022: 288) beizukommen. Unbestritten ist, dass das Leben als „ganze[s] Gewimmel“ den „Hintergrund“ bildet, qua diesem wir „eine Handlung sehen“, „Urteil[e]“ fällen, „Begriffe und Reaktionen“ (BPP: 629) festlegen. Kurz, unsere Begriffe bezeichnen etwas in „diesem Getriebe des Lebens“ (BPP: 625). Und dies gelingt allein, da das Leben trotz aller Unordnung und Wirrnis „verschiedene Muster an den Tag legt, die wiederkehrende und wiedererkennbare Formationen von menschlichen Handlungs- und Verhaltensweisen darstellen und die die Sprecher/innen einer Sprache mit einem Begriff bezeichnen, wenn sie für ihr Leben signifikant erscheinen.“ (Schulte/Majetschak 2022: 288) Bekanntlich hat Wittgenstein solche wiederkehrenden Formationen im sogenannten „Lebensteppich“ (MS 144: 1; vgl. BPP: 673) unseres Lebens explizit als „Lebensmuster“ (BPP: 652) bezeichnet – so in seinen Bemerkungssammlungen zur Philosophie der Psychologie von ihm dargelegt und ab den 1940ern Jahren von ihm zunehmend verwendet – und geltend gemacht, dass unsere Begriffe auf sie bezogen und von ihnen abhängig sind.

„Lebensmuster“ (BPP: 652, 672 u.ö.) bezeichnen dementsprechend nach Wittgenstein Regelmäßigkeiten sprachlichen und nicht-sprachlichen Handelns, welche mit gewissen Variationen im Leben der Menschen wiederkehren, von ihnen als zusammengehörige Einheiten aufgefasst und darum mit einem Wort benannt werden. Orientieren wir uns in unserer sozialen Lebenswelt, so stellen wir, wie Wittgenstein sagt, „Verschiedenes zu einer ‚Gestalt‘ (Muster) zusammen“ (BPP: 651) und benennen selbiges Muster dann mit einem Wort, um auf eine spezifische Konstellation von Vollzügen sprachlichen und nicht-sprachlichen Handelns erneut Bezug nehmen zu können. Dementsprechend bilden „Lebensmuster die Grundlage für eine Wortverwendung“ (LSPP: 211). Wichtig ist in diesem Zusammenhang die „Unbestimmtheit“ die prinzipiell für Begriffe gelten muss, „[d]enn das Lebensmuster ist ja nicht genaue Regelmäßigkeit“ (ebd.), sondern vielmehr „im Teppich mit vielen andern Mustern verwoben“ (BPP: 673) Außerdem erscheint ein Muster des Lebens auch „nicht immer vollständig“, sondern im Gegenteil zumeist „vielfach variiert“ (BPP: 672). Doch „wir, in unserer

Begriffswelt, sehen immer wieder das Gleiche mit Variationen wiederkehren. So fassen's unsere Begriffe auf" (ebd.). Oder wie Wittgenstein an anderer Stelle bemerkt: Ein jeder bezieht sich auf „ein Muster, welches im Lebensteppich mit verschiedenen Variationen wiederkehrt“ (MS 144: 1).

4. Wittgensteins Beispiele für Lebensmuster

Ein von Wittgenstein selbst gegebenes Beispiel für ein Lebensmuster ist etwa „Kummer“, „das im Lebensteppich mit verschiedenen Variationen wiederkehrt“ (ebd.). „Kummer“ ist nicht etwas, das wir einer Person allein aufgrund seines inneren Zustandes zuschreiben, sondern evidentermaßen verwenden, wenn eben diese Person bestimmte Verhaltensweisen an den Tag legt, bestimmte Äußerungen tut oder gerade nicht tut etc. Wiederum ein anderes Beispiel, das Wittgenstein für ein Lebensmuster verwendet, ist „Hoffnung“, welches auf ähnliche Weise wie „Kummer“ als eine konstante Sprach- und Lebensform ins menschliche Leben eingelassen und verankert ist. Wir nennen eine Person „hoffnungsvoll“, wenn ihr sprachliches und nicht-sprachliches Handeln eine bestimmte Form resp. ein wiedererkennbares Muster erkennen lässt. Insbesondere beim Muster des „Grüßen“ wird deutlich – man denke nur an das bereits erwähnte Beispiel „Grüß Gott“ oder das Hongi-Begrüßungsritual in Neuseeland –, dass ein bestimmtes Muster nicht in allen Gesellschaften gleichermaßen vertreten ist. Was für Ausdrücke wie „Hoffnung“ oder „Kummer“ gilt, gilt auch für philosophische Termini: Wittgenstein hat sich vornehmlich in seinen umfangreichen Manuskripten und Typoskripten intensiv mit Sprachspielen in denen die Wörter „Denken“, „Verstehen“ oder „Geist“ vorkommen auseinandergesetzt, um auf grammatikalische Missdeutungen aufmerksam zu machen, wenn wir sie auf „innere Vorgänge“ und „Bewusstseinszustände“ beziehen.

5. Lebensmuster und Religion

Blickt man noch einmal auf die „Muster“ Kummer oder Hoffnung, so stellt sich die Frage, ob das, was für diese gilt, auch für den umstrittenen Begriff der Religion gelten kann. Beschreibt auch der Begriff Religion in Wittgensteins Sinne eine „Gestalt“ resp. ein „Muster“? Anders gefragt, ist auch die Religion „ein Muster, das im Lebensteppich mit verschiedenen Variationen wiederkehrt“ (PPF: 2)? Hierzu muss ich zunächst einmal kurz deutlich machen, warum der Begriff der Religion philosophisch umstritten ist und warum ich den Begriff Religion an dieser Stelle überhaupt ins Spiel bringe.

Der Begriff „Religion“ ist umstritten und nur schwer zu definieren. Zudem scheint es keinen univoken Begriff zu geben, sondern nur analoge und vielleicht auch teils lediglich äquivalente Religionsbegriffe. „Religion“ bezeichnet so unterschiedliche Phänomene wie den individuellen subjektiven religiösen Vollzug, die verschiedenen objektiven Erscheinungsformen von Religion oder Phänomene, die oft zunächst einmal nur im uneigentlichen Sinne als „Religion“ bezeichnet werden. Definitionen erscheinen nur im Hinblick auf einzelne Aspekte der geschichtlich-objektiven Erscheinungsform von Religion oder einzelne religiöse Akte wie etwa das Gebet und seine verschiedenen Formen, das Bekenntnis und die Vergebung von Schuld oder Initiations- und Bestattungsrituale möglich.

Noch schwieriger wird es bei einem genuin philosophischen Versuch, eine Definition des Grundbegriffs „Religion“ zu finden. Zum einen sieht sich die Philosophie mit dem Problem konfrontiert, dass sie bei einer philosophischen Definition von Religion möglicherweise dem religiösen praktischen Vollzug nicht vollkommen gerecht wird. Zum anderen sieht sie sich mitunter dem Verdacht der Vereinseitigung ausgesetzt, bei einer philosophischen Begriffsdefinition nicht vollkommen voraussetzungsfrei auftreten zu können und immer schon von einem (problematischen) Vorverständnis von Religion auszugehen. Ziel dieses Abschnitts ist, all diesen Schwierigkeiten und Problemen mit dem Wittgensteinschen Begriff des Lebensmusters beizukommen. Wenn mein Vorschlag funktioniert, besteht die Möglichkeit, Religion letztlich auch global zu denken. Meine zentrale These ist, wir können qua dem Begriff des Lebensmusters auch die Religion als Grundvollzug des Menschseins besser beschreiben. Bevor ich auf den Begriff Lebensmuster zusprechen komme, möchte ich kurz auf Wittgensteins Verhältnis zur Religion eingehen.

Wittgensteins Verhältnis zum religiösen Glauben wird von eigenen Forschern ins Zentrum seines Denkens gerückt. Dabei gibt es Versuche, Analogien herzustellen zwischen religiösen und philosophischen Ansprüchen, zum Beispiel der Unerklärbarkeit Gottes und Wittgensteins Auffassung, dass Sprachspiele und deren Regeln keine Erklärung oder Rechtfertigung haben können. In einer Fülle von persönlichen Äußerungen zeigt Wittgenstein, dass er nicht nur über religiöse Fragen nachdenkt, sondern sich selbst in bestimmter Hinsicht als religiös denkenden Menschen und als Christ

betrachtet. Es wird dabei häufig auf eine bestimmte Äußerung hingewiesen. In den Jahren, in denen Wittgenstein an den *Philosophischen Untersuchungen* arbeitete, sagte er in einem Gespräch zu seinem Freund Maurice Drury: „Ich bin zwar kein religiöser Mensch, aber ich kann nicht anders: ich sehe jedes Problem von einem religiösen Standpunkt“ (Drury 1992: 121). Erläuterungsversuche dieser ambivalenten Aussage dominieren die Wittgenstein-Forschung bis heute. Es geht wohl letztlich darum, wie wörtlich wir diese Äußerung verstehen können. Feststeht, bereits die frühen Tagebücher, die Wittgenstein in den ersten beiden Jahren des Ersten Weltkriegs als Soldat schrieb, enthalten eine Fülle religiöser Äußerungen. Es sind Gebete, in denen er Gott um Kraft, Beistand, Schutz und Erleuchtung bittet, oder in denen er sich in Gottes Willen fügt. Andere Äußerungen sind deklarativ; sie bezeichnen Gott zum Beispiel als die Liebe und das Christentum als „einzig sicheren Weg zum Glück“. Wir begegnen aber auch Äußerungen, die zeigen, dass Wittgenstein in Sorge und starken Zweifel um seinen Glauben an Gott ist. Seine religiösen Äußerungen in den Tagebüchern zwischen 1914 und 1916 sind allesamt inbrünstig, nicht distanziert, sondern tief gläubig. Dies bedeutet aber nicht, dass Wittgenstein in einem engeren, konfessionellen Sinn gläubig war. Er war zwar katholisch getauft, blieb es, ohne dies bei offiziellen Befragungen zu verleugnen, und wurde katholisch beerdigt. Praktizierender Katholik war er aber nicht. Wittgensteins gläubige religiöse Äußerungen aus der Zeit des Ersten Weltkriegs sollten wir auch nicht als zuverlässigen Nachweis einer bleibenden Einstellung betrachten. In späteren Notizen kommen die Zweifel und Unsicherheiten seines Glaubens zum Ausdruck. Seine Zweifel sind aber begleitet von tiefen Überzeugungen und religiösen Einsichten.

Es ist offensichtlich, dass Wittgenstein nicht nur von der christlichen Tradition geprägt ist, sondern intensiv über religiöse Fragen nachdenkt. Aber Wittgenstein sucht nicht nach religiösen Quellen oder Bestätigungen seines Denkens. Aus distanzierter Perspektive überlegt er, ob es gute Gründe für den Glauben an Gott gibt. Dabei zeigt er sich als Verteidiger des Glaubens gegen dessen Theoretisierung, etwa in der Theologie. Das Christentum sei keine Lehre und keine Theorie, sondern „eine Beschreibung eines tatsächlichen Vorgangs im Leben des Menschen“ (VB: 32). Die Vorgänge, auf die er hier ausdrücklich anspielt, sind zum Beispiel die „Erkenntnis der Sünde“, die

„Verzweiflung“ und die „Erlösung durch den Glauben“. Sie alle hält er offenbar für Tatsachen des religiösen Lebens. Für die Beschreibung der Religion als eines tatsächlichen Vorgangs im menschlichen Leben nutzt Wittgenstein auch seine sprachphilosophischen Konzepte. Zu denen ich nun kommen werde. Dabei möchte ich die Überlegung anstellen, ob Wittgensteins Konzepte vom Lebensmuster uns bei dem Definitionsproblem der Religion helfen können. Kurz, können wir mittels Lebensmuster die Religion als Grundvollzug des Menschseins besser beschreiben?

Im Folgenden ein Vorschlag: Statt platonische Wesensschau zu betreiben und nach einer invarianten Definition der Religion zu suchen, müssen wir die Religion als Lebensmuster verstehen und uns Klarheit über ihre Grammatik, im Sinne Wittgensteins, verschaffen. Wir müssen uns von einer irrigen Suche nach dem Wesen der Religion befreien. Die Religion ist keine Lehre und keine Theorie, sondern die Bezeichnung und Beschreibung eines tatsächlichen Vorgangs im Leben des Menschen. Darauf weist auch Wittgenstein immer wieder hin. Religion ist somit ein anthropologisches Phänomen, das sich empirisch seit den Anfängen der Menschheit nachweisen lässt. So etwa als Hoffnung auf eine andere, jenseitige Welt und auf ein mögliches Weiterleben nach dem Tod oder als betend-bittende Verehrung eines göttlichen oder heiligen Wesens. Religion ist ein großes Lebensmuster im menschlichen Lebensteppich. Und zerfällt in viele weitere Lebensmuster. Lebensmuster bilden die Grundlage für eine Wortverwendung und bezeichnen nach Wittgenstein Regelmäßigkeiten sprachlichen und nicht-sprachlichen Handelns, welche mit Variationen im Leben der Menschen wiederkehren, von ihnen als zusammengehörige Einheiten aufgefasst und darum mit einem Wort benannt werden. Dies gilt, so meines Erachtens, auch für das Konzept der Religion, welche einen Grundvollzug des Menschseins abbildet. Auch bei der Religion stellen wir „Verschiedenes zu einer ‚Gestalt‘ (Muster) zusammen“ (BPP: 651) und benennen selbiges Muster dann mit einem Wort, um auf eine spezifische Konstellation von Vollzügen sprachlichen und nicht-sprachlichen Handelns erneut Bezug nehmen zu können. Kurz, das Lebensmuster Religion zerfällt in unzählige (religiöse) Sprachspiele, wie etwa dem Segnen, Beten, Taufen etc. Wichtig ist in diesem Zusammenhang die „Unbestimmtheit“, die prinzipiell für Begriffe gelten muss, „denn das Lebensmuster ist ja nicht genaue Regelmäßigkeit“ (ebd.), sondern vielmehr

„im Teppich mit vielen andern Mustern verwoben“ (BPP: 673). Außerdem erscheint ein Muster des Lebens auch „nicht immer vollständig“, sondern zumeist „vielfach variiert“ (BPP: 672). Doch „wir, in unserer Begriffswelt, sehen immer wieder das Gleiche mit Variationen wiederkehren. So fassen’s unsere Begriffe auf“ (ebd.). Und ein jeder bezieht sich auf „ein Muster, welches im Lebensteppich mit verschiedenen Variationen wiederkehrt“ (MS 144: 1). Eben dies gilt auch für die Religion.

Eine der Krankheiten des Verstandes, die eine richtig verstandene Philosophie heilen soll, ist nach Wittgenstein das „Streben nach Allgemeinheit“ (BBB: 37) oder die Frage nach dem Wesentlichen oder Gemeinsamen (PU: §65). Und eine der Hauptquellen für das Streben nach Allgemeinheit ist „unsere Voreingenommenheit für die naturwissenschaftliche Methode“ (BBB: 39). Das heißt die Methode, die Naturerscheinungen auf eine möglichst kleine Zahl von Gesetzen zurückzuführen (vgl. ebd.). Die Methode der Lebensmuster ist eine Therapie für eine geistige Krankheit, die Wittgenstein als „Streben nach Allgemeinheit“ oder „verächtliche Haltung gegenüber dem Einzelfall“ (ebd.) bezeichnet. Zwei Tendenzen führen wesentlich zum Streben nach Allgemeinheit: „die Tendenz, nach etwas Ausschau zu halten, das all den Dingen gemeinsam ist, die wir gewöhnlich unter einer allgemeinen Bezeichnung zusammenfassen“ und unsere Voreingenommenheit für die naturwissenschaftliche Methode, d.h. die Methode, „die Erklärung von Naturerscheinungen auf die kleinstmögliche Zahl primitiver Naturgesetze zurückzuführen“ (ebd.).

Das Wort Religion könnte dazu verführen, nach dem Wesen der Religion zu fragen. Dagegen war Wittgenstein nach dem Zeugnis von Drury schon früh beeinflusst von William James Buch *The Varieties of Religious Experience*. Und die Kategorie der „Vielfalt“ hat in seinem Denken eine wichtige Rolle gespielt. So bemerkt er:

Die Weisen in denen die Menschen ihre religiösen Überzeugungen zum Ausdruck bringen mußten, weisen enorme Unterschiede auf. Alle echten Ausdrucksweisen der Religion sind wundervoll, auch die der wildesten Volksstämme. (Drury 1992: 138)

An die Stelle des geschlossenen Weltbildes einer Metaphysik oder der Naturwissenschaft rückt die unübersehbare, sich wandelnde Mannigfaltigkeit der Lebensmuster, die in der Naturgeschichte des Menschen wurzeln und auf denen Sprachspiele aufruhen können und dies gilt auch für die Religion. Wittgenstein lenkt den Blick auf die Vielfalt der Phänomene. Wie Wittgenstein die vielfältigen Lebensmuster beschreibt, so muss die Religionsphilosophie die vielfältigen Formen des religiösen Bewusstseins und seiner Äußerungen beschreiben. Eine Religionsphilosophie kann die Kenntnis des religiösen Phänomens nicht voraussetzen und sich auf dessen Reflexion beschränken; sie muss es in seinen vielfältigen Formen und Brechungen darstellen. Der Vielfalt des religiösen Phänomens qua Lebensmuster entspricht der Vielfalt der Möglichkeiten, es zu reflektieren. Unsere Aufgabe als Religionsphilosophen muss sein, die Vielfalt der Religion qua ihrer Lebensmuster in Form einer „Landschaftsskizze“, um einen Vergleich von Wittgenstein zu bemühen, zu beschreiben. Es spricht nichts dagegen, diese Herangehensweise als phänomenologisch zu beschreiben. Nur so lernen wir die Religion unvoreingenommen kennen und damit uns selbst.

6. Schlussbemerkung:

Insgesamt kristallisiert sich heraus: Lebensformen sind als Lebensmuster zu verstehen und nicht wie standardgemäß als Sprachen und Sprecher einbettende sozio-kulturelle Systeme. Wenn es in den *Philosophischen Untersuchungen* heißt, sich „eine Sprache vorstellen, heißt sich eine Lebensform vorstellen“ (PU: §19), so bedeutet dies, sich vorzustellen, „was für die Menschen, die die Sprache sprechen, eine rekurrente Lebensform resp. ein feststehendes Lebensmuster aus zusammengehörigen Handlungs- und Redeweisen bildet.“ (Majetschak 2019: 93) Und die Rede von „vorstellen“ deutet einmal mehr darauf hin, dass wir uns die Lebensmuster lediglich ausmalen und beschreiben, nicht aber erklären oder rechtfertigen können. Des Weiteren konnte erfolgreich gezeigt werden, dass der Begriff des Lebensmusters für die Bestimmung des umstrittenen Begriffs der Religion vielversprechende Dienste leisten kann. Insbesondere dann, wenn wir unsere Denkweise ändern und verstehen lernen, dass die Religion keine Lehre ist und keine Theorie, sondern eine Beschreibung eines tatsächlichen Vorgangs im Leben des Menschen. Religion ist somit ein anthropologisches Phänomen, das sich empirisch seit den Anfängen der Menschheit nachweisen lässt, etwa als

Hoffnung auf eine andere, jenseitige Welt und auf ein mögliches Weiterleben nach dem Tod oder als betend-bittende Verehrung eines göttlichen oder heiligen Wesens. Religion ist ein global angelegtes Lebensmuster im menschlichen Lebensteppich.

References

Drury, Maurice (1992) „Bemerkungen zu einigen Gesprächen mit Wittgenstein“, in: Rush Rhees (ed.) *Ludwig Wittgenstein: Porträts und Gespräche*, Frankfurt a. M.: Suhrkamp, 117-141.

Garver, Newton (1984) „Die Lebensform in Wittgensteins Philosophischen Untersuchungen“, *Grazer Philosophische Studien* 21, 33-54.

Glock, Hans-Johann (2010) *Wittgenstein-Lexikon*. Übers. von Ernst Michael Lange, Darmstadt: WBG.

Haller, Rudolf (1984) „Lebensform oder Lebensformen? Eine Bemerkung zu N. Garvers ‚Die Lebensform in Wittgensteins Philosophischen Untersuchungen‘“, *Grazer Philosophische Studien* 21, 55-63.

Majetschak, Stefan (2010) „Lebensformen und Lebensmuster. Zur Deutung eines sogenannten Grundbegriffs der Spätphilosophie Ludwig Wittgensteins“, in: Volker Munz, Klaus Puhl und Joseph Wang (eds.) *Language and World. Part One: Essays on the Philosophy of Wittgenstein*. Heusenstamm: Ontos, 265-290.

Majetschak, Stefan (2020) „Eine irreführende Parallele‘. Wittgenstein über Begriffsverwirrung in der Psychologie und die Semantik psychologischer Begriffe“, *Wittgenstein-Studien* 11, 169-182.

Martin, Christian (ed.) (2018) *Language, Form(s) of Life, and Logic. Investigations after Wittgenstein* Berlin/Boston: De Gruyter.

Moyal-Sharrock, Danièle (2015) „Wittgenstein on Forms of Life, Patterns of Life, and Ways of Living“, *Nordic Wittgenstein Review* 4, 21-42.

Savigny, Eike von (1999) „Wittgensteins ‚Lebensformen‘ und die Grenzen der Verständigung“, in: Wilhelm Lütterfelds und Andreas Roser (eds.) *Der Konflikt der Lebensformen in Wittgensteins Philosophie der Sprache*, Frankfurt a. M.: Suhrkamp 120-137.

Schulte, Joachim (1999) „Die Hinnahme von Sprachspielen und Lebensformen“, in: Wilhelm Lütterfelds und Andreas Roser (eds.) *Der Konflikt der Lebensformen in Wittgensteins Philosophie der Sprache*, Frankfurt a. M.: Suhrkamp, 156-170.

Schulte, Joachim; Majetschak, Stefan (2022) „Lebensform“, in: Anja Weiberg und Stefan Majetschak (eds.) *Wittgenstein-Handbuch*, Stuttgart: Metzler, 285-292.

Trächtler, Jasmin (2023) "Facts, Concepts and Patterns of Life—Or How to Change Things with Words", *Philosophies* 58.

Wittgenstein, Ludwig (1970) *Eine Philosophische Betrachtung*. In: *Schriften Band 5*. Rush Rhees (ed.), übers. von Petra von Morstein. Frankfurt a. M.: Suhrkamp, 117–282.

Wittgenstein, Ludwig (1984) *Philosophische Untersuchungen*. In: *Werkausgabe Band 1*. Frankfurt a. M.: Suhrkamp, 224–580.

Wittgenstein, Ludwig (1984) *Das Blaue Buch/Eine Philosophische Betrachtung (Das Braune Buch)*. Rush Rhees (ed.), übers. von Petra von Morstein. Werkausgabe Bd. 5. Frankfurt a. M.: Suhrkamp 1984.

Wittgenstein, Ludwig (1984) *Letzte Schriften über die Philosophie der Psychologie*. In: *Werkausgabe Band 7*. Frankfurt a. M.: Suhrkamp 1984, 347–488.

Wittgenstein, Ludwig (1998) *Vermischte Bemerkungen. Eine Auswahl aus dem Nachlaß/Culture and Value. A Selection from the Posthumous Remains*, Georg H. von Wright und Heikki Nyman (eds.). Neubearbeitung des Textes durch Alois Pichler. Übers. von Peter Winch. Revised second edition. Oxford: Oxford University Press 1998.

Wittgenstein's Nachlass. The Bergen Electronic Edition. Hg. vom Wittgenstein-Archiv an der Universität Bergen unter der Leitung von Claus Huitfeldt. Oxford: Oxford University Press 1998–2000.

Wittgenstein, Ludwig (2009) *Philosophy of Psychology – A Fragment*. In: *Philosophische Untersuchungen/Philosophical Investigations* Peter Michael Stephan Hacker und Joachim Schulte (eds.), übers. von Gertrude Elizabeth Margret Anscombe, Peter Michael Stephan Hacker und Joachim Schulte. New York 2009, 182–243.

Feminist anxieties and practical genealogy

Ronya Ramrath (Oxford, United Kingdom)

Abstract

Taking my lead from Amia Srinivasan's 2019 paper 'Genealogy, Epistemology and Worldmaking', I consider the phenomenon she labels 'genealogical anxiety', examining how ethical flavours of this angst might continue to play a role even when we leave mainstream epistemological concerns around genealogical scepticism behind and move into more political modes of critique. Orienting my discussion around Srinivasan's reading of Catharine MacKinnon's genealogy of male power, I consider the ethical anxieties feminists might have around forms of critique that accidentally foreclose more intersectional perspectives. The problem I want to articulate here is that even as the ethically anxious feminist tries to resist the totalising nature of a particular critique, the contravening impulse to open it up to yet *more* perspectives could itself be considered an expression of the problematic Archimedean urge for a more total point of view. In response, I turn to the realities of feminist praxis and consider how such ethical forms of genealogical anxiety might nonetheless prove fruitful. In particular, I discuss Srinivasan's critique of the 'sex-positivity' movement in *The Right to Sex* and her suggestion that we invite to scrutiny the preferences that we worry might have a contingent political formation at root, which I claim is a kind of critical genealogy in action. Though anxiety-inducing, this practice might serve to help us go beyond our political conditioning, *particularly* through the engagement with other, intersectional perspectives. I conclude by considering whether in this terrain we might find a vindication of MacKinnon's project.

I.

In her 2019 paper 'Genealogy, Epistemology and Worldmaking', Amia Srinivasan discusses a perennial philosophical concern she dubs 'genealogical anxiety': the fear that "the causal origins of our representations, once revealed, will somehow undermine, destabilize, or cast doubt on the legitimacy or standing of those representations" (2019: 128). This, I want to suggest, isn't a purely epistemic worry: whereas we might provisionally appease some sceptics with an appeal to something like *genealogical luck* (i.e., that we contingently happened upon the 'correct' set of beliefs about something) some of us are dogged by a more ethical flavour of genealogical concern – on what basis do we take ourselves to be more genealogically lucky than someone else with a contingently different formation? This is particularly relevant in the context of feminism, where it becomes clear that our representational worlds aren't just products of contingency, but of power and ideology: if we have had

little cause to question our genealogical luck, this is likely indicative of the fact that we inhabit a socially, and by extension epistemically, dominant position; one that is rarely confronted with disbelief and epistemic injustice.

There are two related concerns we might have here: firstly, that our ideologically-inflected representations of the world might affect *us* in detrimental ways, which is the sort of revelation generations of feminist theorists have been interested in carving out (the idea that we are ‘wrong’ about ourselves). The second worry, one that I think is at the root of the ethical anxiety provoked by the appeal to genealogical luck and reminiscent of some forms of standpoint epistemology, is that by virtue of our (contingent) social positions, our representations fail to capture certain politically salient features of other’s experience (the idea that we are ‘wrong’ about others).

Women are no strangers to genealogical anxieties. Inundated with talk of adaptive preferences and what is or isn’t acceptable under the aegis of a feminist self-conception, we’re left wondering at every errant inclination whether it is what we really want, or whether it’s a patriarchy-induced preference that forms part of the oppressive scaffolding of our society, only to then arrive at the paralysing thought that given our contingent patriarchal formations, perhaps there is no differentiating between the two. We might think that if the genealogy of the desire is ‘bad’, so is our genealogy as female subjects. This opens up a new set of questions answered differently by whatever school of feminism one chooses to adhere to – is it a feminist reclamation if we give in to the self-destructive temptation to wear high heels (to offer a rather mild example), or are we conceding the game? It is an anxiety that dogs the experience of the ‘enlightened’ female subject.

Then there is the other-directed variant: what about the fears around feminist organising, the guilt White feminists feel about previous iterations of the movement leaving others out? Here, we slip back into the explicitly ethical angst – who are we to impose our supposedly liberated perspectives on other women, who, perhaps weighed down by a different cultural history of patriarchy, appear to us as insufficiently progressive? Saba Mahmood has written acerbically about the limitations of European feminism and gives voice

to the common decolonial critique that imposing culturally specific standards of freedom or autonomy risks complicity in Western neo-imperial projects (2008: 82). What, then, does an ethical politics look like?

In the final sections of her 2019 paper, Srinivasan shifts her focus from epistemology to more politically pragmatic questions – what matters on this later account is not whether our representations are *true*, but what they *do*. The epistemic anxieties here seem to fade into the background. However, I claim, the *ethical* form of this anxiety persists even here. To get a better grip on it, I want to turn next to the role genealogy plays in radical politics.

II.

The emancipatory potential of genealogies is often located in the idea that by reflecting on our representations, we create the space for improving them: ‘How is the world?’ becomes ‘How do we want it to be?’ The role of genealogy is often seen as largely descriptive, charting the origins of oppressive systems and the ways in which certain practices sustain existing relations of power, so that we may recognise that what we thought was given is both contingent and constructed. This then clears the way for the normative business of combatting the crises of our time. However, this is not a clean division: genealogies themselves aren’t normatively neutral, and our positionality is built in to the narratives of power we construct. This I want to highlight as one of the key insights of Srinivasan’s paper: in her concluding discussion of Catharine MacKinnon and the feminist ‘slave revolt’ against male power, she remarks that “it is only in describing its impossibility—in describing the totality of male power—that feminist consciousness comes into being. [...] [T]o genealogize the world as a product of male power is already to worldmake. Genealogy and worldmaking, for MacKinnon, are one and the same” (2019: 149).

Here, I argue, a certain ethical anxiety reasserts itself – what about the voices that have as yet gone unheard? Projects like MacKinnon’s, while crucial, often risk obscuring other, more marginal struggles from view. Ruíz and Dotson note that despite MacKinnon’s “engaged commitment to intersectional feminisms” and methodological self-awareness, the “interstitial sites of those convergences [between her work and women-of-colour feminisms] are often fraught with logics and assumptions that still privilege the political projects and concerns of Anglo-American and liberal feminisms” (2017: 2). MacKinnon

herself is sensitive to this issue, unlike some of her readers – Ruíz and Dotson articulate their discomfort with “how MacKinnon’s work was being celebrated in ways that could be used to exclude women-of-color feminisms, and the seeming absence of critical awareness of this” (ibid.). Nevertheless, despite MacKinnon’s caveat that “[t]o look for the place of gender in everything is not to reduce everything to gender” (1989: xi), her work is nonetheless refracted primarily through this lens, and thus threatens to elide oppressions that don’t show up on the White feminist horizon. As Ruíz and Dotson point out, “[w]e only look for what we already theorize as missing, as ‘trackable’ by virtue of the prior cultural backdrop of interpretive resources that guide our theoretical projects and claims” (2017: 4). This cultural background is contingent and constructed; there is no Archimedean point.

The worry, then, is that even critique oriented toward liberation might succumb to a certain renewed totalitarianism. One might characterise the difficulty of philosophy as lying in its presupposition of our ability to do something that in fact eludes us, namely to stand outside the relationship between ourselves and the world (compare Srinivasan, 2017). The frustration of philosophising, then, comes from the desire to transcend this relationship and yet find oneself repeatedly thrown back on one’s own finitude (to appropriate Wittgenstein’s image, we continually find ourselves tracing the particular frame of our perspective). This aspiration to look beyond the frame, which Srinivasan styles as ‘the Archimedean urge’, has effectively been held in check by various forms of scepticism, of which genealogical anxiety is one variety. However, as I suggest above, when the sceptical impulse is brought into the service of the philosophical it again becomes susceptible to this Archimedean aspiration, which again sets the stage for a state of genealogical anxiety, a shying away from any position that might suggest a false or ungrounded totality.

To return to MacKinnon: on a certain understanding, she wasn’t just concerned with the possibility of a feminist discourse – she was concerned with writing ‘women’s experience into law’. But which women’s experience(s)? “What does it mean,” asks Wendy Brown, “to write historically and culturally circumscribed experience into an ahistorical discourse, the universalist discourse of law?” (2005, 90) Here, we come up against an inherent issue of critique – as emancipatory discourses are institutionalised and gain power,

making new worlds, they lose their radical potential and themselves risk becoming oppressive. Foucault writes of the risk of the “recodification and recolonization” of “disinterred knowledges” as soon as they are accredited and put into circulation by those “unitary discourses, which first disqualified and then ignored them when they made their appearance,” annexed and brought back into the fold (1981, 86). The will to truth resurfaces in the construction of a new liberatory narrative that usurps its competitors and risks obscuring once again those that don’t show up on the self-emancipating horizon: genealogical revelations of contingency and partiality of perspective are forgotten. Critique, once canonised, might seem to us just as worrisome as its object.

Where does this leave feminists torn between ethical anxieties and political aspirations? Although we might say that the ethical angst attendant on the fear of erasing difference is itself opposed to Archimedean aspirations, threatening as they do to subsume everything under one totalitarian gaze, it could itself nonetheless be considered an expression of that very same urge. The rejection of the totalising standpoints espoused by certain radical positions stems from a desire to reach a more maximally ethical stance, one that takes into account even these moral contingencies and guards against any possible exclusion, which is of course impossible by the same logic. Thus, we are caught in an endless cycle, whereby every attempt to step outside of the limitations of a certain position simply reproduces itself. Genealogy as a method seems to sit on the edge, vacillating between opening our worlds up and closing them again – at first, it may appear an antidote to the Archimedean urge, only for our interpretations to succumb to it once more, in a slippery dynamic whereby we both want to insist on the primacy of the non-neutral perspective and nonetheless feel caged by its inevitable blind spots.

III.

Given these Sisyphean frustrations, it is easy to forget that this ethical anxiety in itself needn’t be a debilitating thing and actually carries the promise of a sort of counter-ideological function. In particular, although it might on occasion seem tempting to give up all epistemic responsibility in the face of a looming threat of genealogical determinism (insofar as we might think we can never escape our contingent epistemic formations), it also generates the sort of awareness that seems indispensable to radical political efforts. Recall the

initial discussion of feminist anxieties, specifically the self-directed variety: yes, it is unlikely that any of us will ever fully deconstruct the conditioning of patriarchy, but perhaps there are ways of identifying other impulses that might take us beyond at least the primary layer of political conditioning.

Consider the following example: in *The Right to Sex*, Srinivasan critiques the nominally self-liberatory ‘sex-positivity’ movement, which, unmoored from certain ambivalence (or lacking, I want to suggest, in genealogical anxiety) risks “covering not only for misogyny, but for racism, ableism, transphobia and every other oppressive system that makes its way into the bedroom through the seemingly innocuous mechanism of ‘personal preference’.” (2021: 84) (This is precisely the oppressive effect of a self-consciously emancipatory but blinkered movement that engenders ethical anxiety even in cases of critique.) On, then, to Srinivasan’s counter-proposal: she suggests that rather than valorising our desires, we should question them, and invite to scrutiny the preferences (white, able-bodied, thin, cis, hereto) that we think might to some extent be contingent on our specific political formations (“‘Black is beautiful’ and ‘Big is beautiful’ are not just slogans of empowerment, but proposals for a revaluation of our values.” (2021: 90)). Although Srinivasan never explicitly draws the connection, this is a live example of critical genealogical practice. By genealogising our preferences, tracing their origins and development within the contingent and constructed political worlds we find ourselves in, we can defamiliarize them and wonder what might have been otherwise (although, of course, this ‘otherwise’ is a similarly ambiguous space).

This is, of course, an anxiety-inducing activity in the self-directed sense, but it allows us to consider whether our desires might in fact point us somewhere beyond the surface layer of contingent socialisation. The idea here isn’t to simply choose against politics in a kind of voluntarism, reminiscent perhaps of women who practice lesbianism on political grounds, although this is a frequent misunderstanding – at one point, Srinivasan asks: “what if the envy you feel for another woman’s body, her face, her charm, her ease, her brilliance, were not envy at all – but desire?” (2021: 97) A philosophical colleague of mine, horrified, compared this to asking a self-proclaimed gay man whether he wasn’t, after all, into women rather than men. But this seems to me a false equivalence – one is a question asked against the system of

power, the other asked with it. I don't think Srinivasan is here suggesting some kind of renewed moral authoritarianism that might itself threaten freshly anxiety-inducing exclusions: not all constellations of desire are available to us. Rather, the hope is that, as Srinivasan observes, "sexual desire doesn't always neatly conform to our own sense of it, as generations of gay men and women can attest. Desire can take us by surprise, leading us somewhere we hadn't imagined we would ever go, or towards someone we never thought we would lust after, or love. In the very best cases, the cases that perhaps ground our best hope, desire can cut against what politics has chosen for us, and choose for itself." (2021: 91)

This, if successful, might seem to us an antidote to at least one locus of feminist genealogical anxiety. However, the appeal to hope is an important caveat: we can recognise a preference as problematic while nonetheless being unable to shake it (the trappings of patriarchy aren't so easily removed, and our ability to peer beyond the frame of the given is limited). Furthermore, a certain anxiety persists: we might worry that this intervention helps us only in the self-directed case. Yes, as a practice it may ostensibly yield an improved sensitivity to others, but only insofar as our interrogation of our desires yields anything more inclusive or diverse than what politics has already established. There is no real insistence on an engagement with other perspectives here yet, nothing that *necessarily* takes us beyond the solipsistic introspection of the armchair philosopher to alleviate our other-directed concerns.

Then again, it seems that in order to truly reflect on our desires, we do need to seek out other perspectives in order to dislodge the perceived fixity of our own. It is certainly true, for instance, that while of course LGBTQ+ tendencies have always existed (as have people who recognised these tendencies within themselves against the mainstream articulations of appropriate sexuality), their newfound visibility has awakened a more prevalent sensitivity to such preferences. It takes exposure to a world in which these questions may be asked, exposure to people doing the asking, to inspire the appropriate introspection and attendant genealogical questioning of what we have always previously assumed – this is true of our sexual desires as much as of any other preferences, adaptive or not, that shape us as women. As much as 'consciousness-raising' may be occasionally maligned for the same reasons as any totalising critique – in that one party may implicitly assume a superiority

of perspective that once again fails to take into account its own partiality and contingency – there is something powerful about the practice of people talking to each other about the world they inhabit. This power doesn't lie in the establishment of a shared, totalising narrative or the imposition of a particular perspective, but rather in being confronted with a different way of seeing that may or may not trigger a shift of recognition in one's own, cultivating new forms of responsiveness to shared realities.

IV.

I wonder if perhaps, somewhere in this terrain, we may find a vindication of MacKinnon's worldmaking narrative of male power. It seems wrong to understand her 'worldmaking' here as connoting an appeal to transcendent normative standards (greater truth? Greater goodness?) that Srinivasan is quick to point out MacKinnon rejects as "all too immanently male." (2019, 149). As MacKinnon says, "[f]eminism criticizes this male totality without an account of our capacity to do so or to imagine or realize a more whole truth" (1989, 115) – there is no prior vision of the world beyond the patriarchy, because this would be impossible from within its totality (hence the self-directed anxieties of women worrying about their patriarchal formations). The 'worldmaking' at play here involves providing the conditions of the possibility of imagining a world beyond the totality of male power, in a consciousness-raising exercise that is enacted *through* the description of the world as male. This is a kind of contradiction, but a performative one that ruptures the apparent self-evidence of our contingent worlds without offering a full-fledged new representation, just a possibility of something else that we may or may not find ourselves resonating with (just like a non-mainstream expression of desire). Arousing our genealogical anxieties, MacKinnon may sensitise us to aspects of our experience that lie behind layers of political conditioning – not out of a fresh Archimedeanism, but to unsettle and provoke, without necessarily claiming a new authority.

Of course, it is not clear that this reading does much to assuage the ethical anxieties of MacKinnon's feminist readers, given that her work is nonetheless open to more worryingly totalising (and hence, exclusionary) interpretations and appropriations. What makes a work more or less conducive to ethical anxiety may, on this account, be defined more by how well it guards against appropriation by what Brown calls "easy polemics" (2005: 131). More

generally, it seems we have a low tolerance for thought-provoking ambiguity: Srinivasan, whose mastery of nuance is unparalleled in the popular philosophical literature, has nonetheless received criticism for how *The Right to Sex* foregrounds suggestive remarks over ‘clear theses’. Ethical genealogical anxiety, it seems, is here to stay – if only because the Archimedean urge is so pervasive, even in interpretation. Whether we choose to succumb to it ourselves is a different question.

References

Brown, Wendy (2005) *Edgework: Critical Essays on Knowledge and Politics*, Princeton: Princeton University Press.

Foucault, Michel (1981) *Power / Knowledge: Selected Interviews and Other Writings 1972 – 1977*, Colin Gordon (ed.), New York: Pantheon Books.

MacKinnon, Catharine (1989) *Towards a Feminist Theory of the State*, Cambridge, MA: Harvard University Press.

Mahmood, Saba (2008) “Feminism, Democracy, and Empire: Islam and the War of Terror.” *Women’s Studies on the Edge*: 81–114, New York, USA: Duke University Press.

Ruiz, Elena, and Kristie Dotson (2017) “On the Politics of Coalition”, *Feminist Philosophy Quarterly* 3 (2), <https://doi.org/10.5206/fpq/2017.2.4>.

Srinivasan, Amia (2017) “Genealogical Anxiety: An Interview with Amia Srinivasan”, *Oxford Review of Books*, <https://www.the-orb.org/post/genealogical-anxiety-an-interview-with-amia-srinivasan>.

Srinivasan, Amia (2019) “VII — Genealogy, Epistemology and Worldmaking”, *Proceedings of the Aristotelian Society* 119 (2): 127–56.

Srinivasan, Amia (2021) *The Right to Sex*, London: Bloomsbury.

Acknowledgement

Thanks to everyone at the SWIP Ireland Conference ‘World Building in Times of Violence’ for their helpful comments on earlier versions of this material. Thank you also to the anonymous reviewer for their constructive input, and finally to Stephen Mulhall and Kate Kirkpatrick for continually engaging with my ideas.

The method of grammatical analysis in Philippa Foot's early moral philosophy

Toni Rašić (Starčevo, Serbia)

Abstract

Philippa Foot counts along with Elizabeth Anscombe and Peter Geach among the earliest adopters of Wittgenstein's method of grammatical analysis within moral philosophy. In his book „Philippa Foot's Metaethics“ Hacker-Wright explores how this method shapes the later stages of Philippa Foot's moral philosophy (post Foot 2001). In this paper I shall argue that the method of grammatical analysis is just as present in Foot's early body of work in the late 1950ies. Wittgenstein did not just leave a methodological trace in Foot's Philosophy, though. Some of the key Insights of the PI are also integrated in her moral analysis. The idea that terms must have public criteria and cannot be given arbitrary meaning by the individual speaker (Hacker-Wright 2021b) figures directly into Foot's early moral investigations.

I begin by briefly explaining the meta-ethical position called noncognitivism and Foot's grounds for rejecting it. In the second part of the paper I will showcase the grammatical investigations Foot used to back up her criticisms of noncognitivism. In the third part I shall work out an understanding of the grammatical analysis as it is employed by Wittgenstein in the PI, the principal work of Wittgenstein's later philosophy that was available to Foot in the late 1950ies. I conclude the paper in the fourth part by comparing Foot's use of the grammatical method to Wittgenstein's own use of it and by tracing the seeds of ideas that would blossom into meta-ethical grammatical investigations in Foot's early writing back to passages of the PI.

1) Moral cognitivism

By the middle of the 20th century, moral Philosophy at Oxford was dominated by the moral subjectivism of Moore's and Ayer's and Stevenson's and Hare's prescriptivism (Hursthouse 2012: 182). As the linguistic turn was still at its zenith, these philosophers were primarily concerned with morality's expression through moral judgments. Ayer argued that moral judgments are expressions of the speaker's feelings. Hare understood moral judgments as prescriptions or imperatives. According to him, the moral judgment "it is good to..." leads the action-guiding imperative "I ought to..." (Hare 1952: 44). Foot calls such conceptions 'noncognitivist' because they see moral judgments not as descriptions of facts capable of having a truth value, but as expressions of feelings or attitudes or as imperatives, which thereby perform noncognitivist functions (Arrington 1989: 45). Accepting this construal means maintaining that there can be no moral knowledge, as the only judgments that can be known are those capable of being true (Arrington 1989: 1).

Noncognitivism seems most convincing in cases of disagreement. I might try my hardest to convince you of the environmental impact of the meat industry, and you might even agree with me on that, while disagreeing on the judgment that it is wrong to consume meat. We might have to concede that the differences in our judgments are based in our differing attitudes. Foot sees three problems in agreeing to disagree. Firstly, such a view leaves open a gap between the facts that a moral judgment bases itself on and the judgment itself (Foot 2003: 34). This gap leads to the above described instance in which, given the same set of facts two people agree on, they can still come to wildly different conclusions about what is just, right or good (Foot 2003: 34). The facts are real states of the world, such as the fact that during the Second World War nazis have executed innocent Jews, Romani, queer people, people with disabilities and political opponents in gruesome concentration camps. The moral judgments consist in what we make of those facts. Given the fact that people were being tortured and killed in concentration camps we'd judge that act as abhorrently wrong, whereas the nazis perpetrating the act would, without denying the fact that they were perpetrating it, deny that what they were doing is wrong (Foot 2003: 34). In extreme cases like this it becomes apparent that one side simply has to be in the wrong regarding their judgments. To Foot, noncognitivism is impermissible in light of such cases. Noncognitivist accounts, as stated by Foot, are governed by the thought that there is no logical connection between statements about the facts and statements about value, which leads each person to make up their own evaluation (Foot 1958: 504). But in Foot's view, "[m]orality just cannot be subjective in the way that different attitudes, like some aesthetic ones, or likes and dislikes, are subjective." (Foot 2003: 34). Were we to take up noncognitivist approaches, moral judgments would be reduced to simple expressions of one's own likes and dislikes (Foot 1958: 502) and the facts that moral judgments are about would be irrelevant to the judgments themselves (Hacker-Wright 2021b).

Secondly, the noncognitivist conception of moral reasoning contrasts with reasoning in non-moral contexts, such as in science (Arrington 1989: 23). Foot asks how it is that we do not accept someone's statement that the Earth is flat as simply their attitude towards the facts but we would, if we were noncognitivists, accept that we can have opposing stances on important moral issues (Foot 1958: 503).

And thirdly, while some noncognitivists, such as Hare, did acknowledge that moral judgments have a descriptive component, they thought it secondary to and dependent on its evaluative meaning (Arrington 1989: 52). Foot argues that noncognitivism implies that someone could accept a standard of goodness which incorporates any standards of goodness whatsoever, even if they were irrelevant to it (Arrington 1989: 52). A moral eccentric might consider an act as mundane as clapping your hands to be a good act (Foot 1959: 84). This struck Foot as odd, leading her to conclude that there must be constraints on what the object of moral approval can be (Arrington 1989: 52).

2) Foot's reply

Descriptive moral propositions, Foot argues, are descriptive on account that they do not offer the speaker's evaluation of goodness in form of commendation or condemnation (Foot 1958: 506). Passing a moral judgment in Foot's view is more akin to describing an objectively present set of facts rather than giving one's opinion on something. The premise of her cognitivism is that moral judgments are descriptive utterances capable of being true or false. In order to prove this, Foot conducts a grammatical investigation of the concept 'good' in her 1958 paper 'Moral Arguments'. She writes that 'good' is a descriptive word. There is an established practice of saying that it is a fact that something is a good action or that someone is a good person (Foot 1958: 505, 506). Similarly to the word 'good', the word 'rude' expresses the speaker's disapproval of a certain action and is used to discourage offensive behavior (Foot 1958: 507). The noncognitivists got that part right. However, if condemnations fulfilled only noncognitivist functions, anything at all the speaker thinks to be rude could be called rude. But that is not how 'rude' is used in language. Something can only be called rude when certain descriptions apply (Arrington 1989: 53). The condemned action has to be such that it causes offense by showing a certain lack of respect in face of generally accepted patterns of social conduct (Foot 1958: 507). The example that Foot gives is that of a man not taking off his hat while indoors (Foot 1958: 507). This action can only be considered rude if the patterns of social conduct are in place that would designate such behavior as rude. Indeed, nowadays, keeping on a baseball cap indoors would hardly be seen as an offense in most casual circumstances. The offended party's attitude or belief that the act of keeping on a hat indoors is rude may be present and it does figure in their

condemnation of that act, but it not a sufficient condition for deeming the act to be rude. Conditions for rudeness have to be met (Foot 1958: 508). Foot concludes that “either thinking something rude is not to be described in terms of attitudes, or attitudes are not to be described in terms of such things” (Foot 1958: 508). What is significant here, is that in order to condemn something as rude, the offended party cannot make reference to their subjective attitude, but instead has to use the same criteria as anyone else, since, if the condemned action meets the criteria for rudeness, it would be impossible to deny that the action is rude (Foot 1958: 509). According to Foot, the conceptual grammar of ‘rude’ requires the inference from a factual judgment to an evaluative judgment (Arrington 1989: 54). Were every speaker to refer to their own private idea of ‘rude’, people would not just constantly disagree on what was rude or not, such discussions could not even begin (Foot 1958: 509).

Foot infers from this that “[a]nyone who uses moral terms at all, whether to assert or deny a moral proposition, must abide by the rules for their use, including the rules about what shall count as evidence for or against the moral judgment concerned” (Foot 1958: 510). A person cannot make their own personal decision about what counts as evidence for moral judgments (Foot 1958: 510, 511). According to Foot, there must be a constraint on what the good-making properties can be. The relationship between these properties and moral approval cannot merely be external, so as that the approval could attach itself to any object (Arrington 1989: 52). Instead, the relation would have to be internal. The expressed attitude and a certain set of moral properties must be connected in a moral judgment (*ibid*). If something were to be accurately described as ‘good’, it would have to possess features which are built into the very notion of ‘goodness’ (Arrington 1989: 53). The descriptive content of ‘good’ is to a large extent fixed and cannot be entirely dependent on the variable feelings of the individual doing the evaluation (Arrington 1989: 53). Hence the descriptive content of ‘x is rude’ and the attitudes expressed by this judgment do not divide into two separate conditions that could be met independently, and even less is it the case that the descriptive content is determined by the attitudes of the speaker (Arrington 1989: 53).

In ‘Moral Beliefs (1959)’ Foot argues against the assumption that a person can, without making a logical error, base their moral beliefs on premises which no one would accept as giving evidence at all. To this end Foot conducts a

grammatical investigation of the word 'pride'. Propositions that express pride can be thought of as expressing the attitude of the speaker towards the object of their pride. However, there are limits to what one can feel proud of (Foot 1959: 86). If a person said they felt proud of the sea or the sky, whatever they felt, it could not be called pride, as their own use of 'proud' would be incongruent with the public use of 'proud' (Foot 1959: 87). For something to be able to be an object of pride, it must be one's own achievement (Foot 1959: 87). That is to say, a set of conditions must be fulfilled in order to profess one's pride in something. This investigation lets Foot conclude that "[...] there is no describing the evaluative meaning of 'good', evaluation, commending, or anything of the sort, without fixing the object to which they are supposed to be attached." (Foot 1959: 86). The concept 'pride' must therefore contain the conditions for its application on an object. Similarly, the concept of a good person already contains the criteria for its application (Arrington 1989: 53). Thus the relationship between the attitude and its object is an internal one (Arrington 1989: 53).

3) The grammatical Method

An important role of philosophy, according to Wittgenstein, is to elucidate concepts, as „[a] main source of our failure to understand is that we do not command a clear view of the use of our words. – Our grammar is lacking in this sort of perspicuity. A perspicuous representation produces just that understanding which consists in ‚seeing connections‘“ (PI §122). We gain insight into these connections by examining the different use-cases of a given concept. It is also what makes this method a 'grammatical' method. Wittgenstein's most basic conception of grammar is that it consists of rules which govern the use of words and which thereby constitute meanings of concepts (Forster 2004: 7). Grammatical rules constitute concepts (Forster 2004: 14). Since the meaning of a word is its use in language (PI §43), if we want to elucidate the meaning of a word we shall look at the way in which it is used in common parlance. 'Grammar' can thus also mean the act by which we describe the use of signs (PI §496). Grammar in the first meaning is the set of explicit or implicit rules of language, while grammar in the second meaning describes grammar in the first meaning (Forster 2004: 19). Examples of grammatical rules include "one plays patience by oneself" (PI §248) and "every rod has a length" (PI §251). Stating grammatical rules is like reading the rules

for playing a board game on the back of the package. The grammatical method is 'grammatical' in the second sense. It doesn't prescribe how we ought to use language. It is purely descriptive of natural language, as "[p]hilosophy may in no way interfere with the actual use of language; it can in the end only describe it" (PI §124).

I find it justified to speak of Wittgenstein's investigations in terms of a proper philosophical method, as, while there isn't one philosophical method according to Wittgenstein, „there are indeed [many] methods“ (PI §133). I claim that the grammatical method is one of them. In using the grammatical method Wittgenstein's interest is not in the structures of language for their own sake, but rather in the help that the appreciation of the grammar of our expressions could provide in resolving philosophical problems (Hacker-Wright 2021a: 5). The goal isn't to construct an argument made up of premises and conclusions, but to expose a philosophical position as confused, flawed or incongruent with its own assumptions. This is achieved by giving examples of language-use which cause such confusions. Wittgenstein writes that “[T]he fundamental fact here is that we lay down rules, a technique, for a game, and that then when we follow the rules, things do not turn out as we had assumed. That we are therefore as it were entangled in our own rules. This entanglement in our rules is what we want to understand (i.e. get a clear view of). Hence it is important to find and invent intermediate cases (PI §122). Wittgenstein conducts many such grammatical analyses throughout his later work. Consider for instance one of the main themes of the PI: the critique of the augustinian conception of language as consisting primarily of acts of naming (PI §1). An area in which the augustinian conception becomes particularly confused is that of sensations. Augustinus would explain our acquisition of pain language like this. We stub our toe and feel pain. We learn that the thing we feel is called 'pain'. We first learn on our own example what the word 'pain' means and only then do we infer from analogy that other people mean by 'pain' the same sort of thing as we do (Nielsen 2007: 72). In §257 Wittgenstein considers the use of the word 'pain' as a name for the sensation in order to explain how the act of naming an object is, contrary to how the augustinian picture understands it, not one of the elementary moves of language. Instead, a lot needs to be prepared in language in order for the act of naming to make sense. He shows that by simply naming a sensation we have accomplished nothing. We haven't

made a move in the language game. There needs to be an established set of rules that govern pain-utterances and situate them within a linguistic practice. A corollary of the Augustinian view is that sensations would be essentially private to the person having them. This leads Wittgenstein to analyze use-cases of the concept 'private'. Wittgenstein shows how confusions inevitably arise from calling sensations private, as "sensations are private" is being treated like an empirical proposition asserting something about the nature of sensations, while it's actually more readily comparable to a grammatical proposition (PI §248, 251).

The method of grammatical analysis can be summed up as the method by which philosophical positions are shown to be confused by comparing and contrasting their use of words and concepts with the use of words and concepts in natural language.

4) The Wittgensteinian legacy of Foot's cognitivism

Wittgenstein's philosophy had at first reached Foot indirectly, through the teachings of her close friend and tutor, Elizabeth Anscombe, who was a student of Wittgenstein's (Foot 2003: 34). At a later point in her life, Foot would recall how „[Anscombe] must have had discussions with Wittgenstein on topics like the ones that I was discussing with her. She didn't talk about Wittgenstein, but she was teaching me something of his way of thinking“ (Foot 2003: 34). Interestingly, Anscombe had not wanted Foot to read the PI (Hursthouse 2012: 183). It was on Norman Malcolm's behest that Foot had first read the *Philosophical Investigations* (Foot 2003: 35). Wittgenstein's influence on Foot's early moral philosophy can be felt in two ways. Firstly, Foot adopted the method of grammatical analysis and applied it in a more focused fashion on meta-ethical discussions, and secondly, some of her considerations can be traced back to ideas explored in the PI. Foot doesn't prescribe how moral language ought to be used. She simply points towards real use-cases of moral concepts and shows that cognitivist positions lead to a number of confusions. Foot shows through her grammatical investigation that concepts like 'rude' and 'pride' are both descriptive and evaluative (Arrington 1989: 65) and that the descriptive and evaluative meanings are fused and cannot be separated in such a way that the descriptive meaning would depend on the evaluative meaning (Arrington 1989: 72). The descriptive content of 'rude' and the attitudes expressed by judging something as rude do not divide into separate

conditions that could be met independently (Arrington 1989: 53). This is further reminiscent of Wittgenstein's writing in §22 of the PI. Wittgenstein criticizes Frege's distinction between the sense, or conceptual content, of a sentence and the assertoric force with which it is uttered. Foot's criticism of Hare bears similarities to Wittgenstein's criticism of Frege – uttering an assertion and asserting it aren't separate acts, just how uttering a moral judgment isn't a separate act from moral evaluation (Arrington 1989: 61). Furthermore, Foot maintains that there are certain linguistic rules in place for calling an action just, unjust, moral, immoral, good or bad. Only when a person adopts these rules for themselves in their moral judgment can we say that they hold moral principle in the first place (Foot 1958: 512). This itself is reminiscent of Wittgenstein's investigations into the possibility of private rule-following. To speak a language is to partake in a practice (PI §202). Words don't have meaning by ways of private association of name and object, but by adherence to publicly known rules. According to Foot, when passing moral judgments, we don't merely express our subjective views. The judgment is made meaningful by stating that an act meets the public linguistic conditions for being called rude, just or unjust.

References

- Arrington, Robert L. (1989) *Rationalism, Realism and Relativism*, London: Cornell University Press.
- Foot, Philippa (1958) *Moral Arguments*, *Mind* 67, 502–513.
- Foot, Philippa (1959) *Moral Beliefs*, *Proceedings of the Aristotelian Society* 59, 83–104.
- Foot, Philippa (2001) *Natural Goodness*, Oxford University Press, New York.
- Foot, Philippa (2003) *The Grammar of Goodness*. An Interview with Philippa Foot, *The Harvard Review of Philosophy* 11, 32–44.
- Forster, Michael N. (2004) *Wittgenstein and the Arbitrariness of Grammar*, Princeton University Press.
- Hacker-Wright, John (2021a) *Philippa Foot*, *The Stanford Encyclopedia of Philosophy* (Winter 2021 Edition), Edward N. Zalta (ed.), URL = <<https://plato.stanford.edu/archives/win2021/entries/philippa-foot/>>.
- Hacker-Wright, John (2021b) *Philippa Foot's Metaethics*, Cambridge University Press.
- Hare, R. M. (1952) *The Language of Morals*, Clarendon press, Oxford.

Hursthouse, Rosalind (2012) *Philippa Ruth Foot*, Biographical memoirs of Fellows of the British Academy 11, 179-196.

Nielsen, Keld Stehr (2007) *The Evolution of the Private Language Argument*, Routledge, London.

Wittgenstein, Ludwig (1953) *Philosophical Investigation*, Blackwell.

Giving body to hinge epistemology

Janyne Sattler (Florianópolis, Brasil)

Abstract

The purpose of this essay is to delineate a feminist epistemological concept of ‘situated knowledge’ from the perspective of the body as hinge around which epistemological constructions – and even those of first order knowledge – turn. It aims at answering the question of who the (supposedly universal) subject of true and justified belief (really) is and at furnishing a more inclusive and pluralistic grounding for the subject of knowledge. It departs from taking the ‘body’ as a founding concept for epistemic subjectivity and agency. This starting point is informed by both feminist epistemologies and a feminist reading of Ludwig Wittgenstein’s hinge epistemology. In the long term, the aim of my background project is to challenge the concept of propositional knowledge as the standard and the criterion for expressions of both epistemic subjectivity and agency, and epistemic authority and authorization. This may also require challenging the central role attributed to justification by canonical epistemology or else re-conceptualize ‘justification’ from the perspective of the ‘embodied situated knowledge’. For the mean time, I would like to approach the concept of ‘body’ by means of Wittgenstein’s *On Certainty* and to advance a possible interpretation concerning hinging qualities as attributed not only to propositions, but to concepts as well.

Over the last few years, my reflections on the philosophy of Ludwig Wittgenstein have coexisted with an ever more detailed and in-depth study in feminist philosophy – due to circumstances, but also due to the ballast of discoveries. At first, I thought I was moving away from my initial concerns and considered my path, if not bifurcated, at least fragmented. Recently, however, I realized that this variation constituted not a thematic detachment, but a qualified deviation, an excursion enriched by the look of plurality and by the call for political responsibility in our philosophical practices. The points of contact are found largely in the epistemological field, as in the methodological one, and a certain meta-philosophical perspective (already present in my previous inquiries) crosses both the diagnostic frameworks and the alternative proposals to the “exclusionary gestures” (as Naomi Scheman would put it: Scheman 2020: 31) of the philosophical norm towards a more democratic horizon – one that informs my criteria for moral, political and epistemological consideration. It is within this reflective context that I would like to think about the mutual contributions between hinge epistemology and feminist epistemologies.

But before starting, let me state some things that should count as my background and that I won’t be able to develop properly in this brief essay.

First, it should be noted that I am not interested in developing here Wittgenstein's philosophy in a fully exegetical manner. His philosophy constitutes more of a starting point for methodological and conceptual reflections that intersect with my aims in the field of feminist epistemology. Second, I am assuming a general understanding on the variety of propositions made by feminist epistemologies, but my main readings stand with Sandra Harding and Donna Haraway, and some of their crucial diagnosis and concepts (as 'situated knowledge' and 'experience'). Third, I'm trying to develop a suggestion made by Alana Ashton (2019) and some of Naomi Scheman's writings about a potential exchange between feminist epistemology and Wittgenstein's epistemological and methodological approach to knowledge and philosophy. Finally, there's Raimond Gaita approach to morality and his Wittgensteinian reading on corporeality and creatureliness. Gaita informs one of the crucial concepts that guide my current endeavor along this epistemological-ethical-political path by insisting on the need to "return the body to the realm of meaning" (Gaita 2004: 98).

It is the way this last sentence is expressed that leads me to think about the mutual contributions between hinge epistemology and feminist epistemologies, since it is not a novelty that the concept of 'body', in its various expressions, permeates the claims of feminist epistemologists – 'situated knowledge', 'localized knowledge', 'embodied or gendered knowledge', 'corporeal vision', 'positioning'. But it is not just the question of situation that interests me here, but the question of meaning. Otherwise put: the question of how to return the body to the realm of meaning seems to be a demand of feminist epistemologies that may be answered with an approach guided by hinge epistemology. The interest of this epistemological version lies in the possibility of finally dissolving the insistence on the "purely epistemological" character of canonical formulations and definitions and of knowledge-justification assertions – "dissolution" that contends, first and foremost, the misunderstanding of the question.

With this movement of repositioning the question concerning (not only attribution and legitimization but also) the justification of knowledge, I am taking a step back from the propositions to the concepts, imagining that conceptual hinges do count as the "certain things" [*Gewisses*] that are not put in doubt and that must be firmly in place (OC 1969: § 342-343) – whose

expressions seem to have been ignored by epistemology. I am aware that this does not necessarily agree with standard interpretation of hinges as propositional. But, as I said, I'm trying to unfold Wittgenstein's conceptualization to serve my own intentions.

Therefore, I would like here to talk about hinge concepts and, especially, of the body as hinge, and to propose their meanings as being detached from doubt, constituting the foundational basis for any epistemological manifestation and essential to that political, democratic and emancipatory horizon present in feminist philosophy.

I think of hinge concepts as those that may answer the question of who the socially situated subject is. For what there is prior to the subject and what there is prior to subjectivation, and even what there is prior to knowledge, is the body. We do not talk about self-consciousness, subjectivation, individuation, not even about personhood, or epistemic concessions, without the axial presupposition of corporeality. Even the terminological variation of our conversations is a variation that revolves around conceptual hinges that are thus essential to further understandings and visions about reality and life. This may help us to clarify why it is we cannot speak of the "purely epistemological" character of epistemic justification, the question being: What are the meanings that structure a body *as* a subject? It is at this point, it seems to me, that we can begin to talk about returning the body to the realm of meaning and begin to understand what the various technologies of situational inscription operate on.

Raimond Gaita talks about the realm of meaning as structuring morality, as he is interested in a (non-analytic) meta-ethical investigation that would allow us to escape from the bridles of normative theories. This is an eminently non-legalistic conceptual search. I share this interest with him, and I do understand it as epistemologically comprehensive. In this sense, the meanings that structure the body as a subject of knowledge are the same ones that structure it as a living body (the body of morality), a stance that allows for us to reject the claims of primacy of epistemology over ethics. It is, in fact, in this epistemological-moral imbrication that the meanings reside on which social, political, and cultural situation are added to qualify the subject. I am thinking here of those contextualized meanings, that might be differently read by

different cultures, but that share corporeality as one of those features that make an individual an irreplaceable inhabitant of a (perishable) world, whose vulnerability is crossed by sexuality, bonds, and affections, mortality, and the need to understand otherness. Gaita calls these concepts “shared assumptions” that must be “disciplined by critical concepts” (Gaita 2004: 92) (lucidity, sobriety, truth, responsiveness) in order to prevent “meaning-blindness” and exclusionary biases; the latter ones being, precisely, the result of racial, colonial and gender technologies and pressures. I would like to call them “hinge concepts”.

I could elaborate this differently by saying that the meanings that structure a body as a subject are the meanings assigned by our unquestioned responses to bodies as bodies – and not as exploitable things or resources. Responses that have to do with the attitude of the human mind (OC 1969: § 89), and that depend on the assumption or conformation (OC 1969: § 343-344) on which or from which we can effectively ask meaningful questions or develop epistemological answers. These answers should be interested not only in understanding ‘knowledge’, but also interested in understanding how definitions of ‘knowledge’ may constitute policies of exclusion and extermination. That is why the feminist stance that informs the present reflection on hinge epistemology asserts a commitment that is above all political. Things could not be otherwise, given the clarity of feminist epistemologies’ denunciation; they all agree on the fact that the subject of canonical (rational) knowledge is himself built on a determined and qualified body from privileged technologies of supposed superiority and superlative morality. In spite of himself, he is a body as well. The point is to understand what makes him a body whose presence and discourse performs a reputed universal omnipresence and normative institution, to the point that we were able to realize only recently that this disembodied characterization is not innocent, but a political coup aiming for control and domain. What structures even this subject as a body is his inescapable and unquestioned contextual and experiential bodily dependence. Power and privilege are themselves ready-made gestures that work as responses and reactions to the surface of corporeality.

This is to emphasize the inescapable links between epistemology and politics so that we can turn our attention to what really determines our knowledge

assertions and to the meanings that structure the body as a subject. Now, these meanings are constructed from our immediate, safe, and obvious attitudes and reactions, as they are immediately experienced as embodied reactions. They are expressed in unsuspected ways by means of a language that shows that we are all inhabitants of the same shared world. They do not find in skepticism – and not even in exacerbated forms of rationalism – a commonplace of communication.

Raimond Gaita insists on how the living body (human and non-human) is essential for the constitution of our concepts, “including our concepts of belief and knowledge” (Gaita 2004: 36), and on how this formation emerges from our “unhesitating interactions” (Gaita 2004: 31) with others, with other people in their bodies, and from our constituent reactions in response to the living world, no, he says, as a “consequence of ascribing states of consciousness to others”, but as a condition (*ibidem*). Here, I would like to argue that this condition can be understood and developed from two related points found in Wittgenstein which would allow us to think about how our beliefs, assumed to be true and justified, are determined by reasons that are always structured on hinge concepts such as body, surface, and experience.

“Thinking,” says Wittgenstein in *Philosophical Investigations*, “is not an incorporeal process which lends life and sense to speaking, and which it would be possible to detach from speaking” (PI 2009: §339), but not just from speaking, but from acting as a whole. After all, what is it that allows us to distinguish what a living body is, animated by action and expression, from “a thing”? It is our attitude presupposed in our behaviors and responses and reactions to the living bodies: “Our attitude to what is alive and to what is dead is not the same. All our reactions are different” (PI 2009: § 284). But not because the body “possesses” a soul that animates it. This language is still equivocal and dualistic, and when he says that “the human body is the best picture of the human soul” (PPF 2009: §25), it is to the picture of the living body that Wittgenstein refers to, one who “behaves like a human being” (PI 2009: § 283) and manifests his pains and his human behavior, which we are able to recognize because we recognize a soul: “my attitude towards him is an attitude towards a soul. I am not of the *opinion* that he has a soul” (PPF 2009: §22). There is no place to have opinions here as there is no place to trust or distrust in automaton’s animation or in children’s fiction about talking furniture and

sentient stones, because no “expression of doubt” about the ways of reading, seeing, interpreting and recognizing a human soul, a living body, has place in this language game (PI 2009: § 288). We could thus speak of a grammar of the living body in the context of which we construct our concepts related to sensations, as those related to belief and knowledge as well. So, we could then say, with Wittgenstein, that just as it is not the body that feels pain, but the living body (“the sufferer”, PI 2009: §286), so it is not the incorporeal soul that thinks, believes, and doubts, but the “soul which some body *has*” (PI 2009: §283); or: the subject as a body.

The epistemological “anteriority” does not reside, therefore, in the supposedly “purely epistemological” concepts that constitute the canonical definition of knowledge, but in the living body as an undoubted expression of subjectivation and knowledge. Without the body as hinge, as a record of all the propositional inscriptions that unfold into action and attitude, as well as into judgments and claims, we do not have an anchor point for the “acting, which lies at the bottom of the language game” of justification. (OC 1969: §204). And when we cannot effectively provide a reason for or against a decision concerning the truth or falsity of a proposition about certain beliefs (OC 1969: §200), this may be due to our mixing up language games in the course of our attempt to understand the ways our concepts and epistemic assertions are constructed (skepticism being, in this sense, undeniably a symptom of wasted philosophical confusion, as it is the dogmatic need for the inviolability of certainty). After all, what are the good reasons that support my “true belief”? “I act with *complete* certainty,” says Wittgenstein, “but this certainty is my own” (OC 1969: § 174). There is no justification for this belief. Or rather, at this point the question of justification is based on the (epistemologically and conceptually prior) condition of corporeality and of our unquestioned and unhesitating attitudes and reactions that allow us to act “surely, acting without any doubt” (OC §196).

None of this needs to be understood in the sense of foundationalist claims, although we could call the concepts that I named as hinge concepts as foundational concepts as well, but only because they are structural or structuring, especially if we take the meaning of ‘hinge’ or ‘fulcrum’ seriously as a point of anchorage for the propositional, systematic frameworks of the grammar of belief and knowledge – and morality – as belonging to the

grammar of the living body. In this grammar, our speeches, actions, reactions, and attitudes compose a system of propositions that support each other (OC 1969: §142) where the body works as a crossroads (OC 1969: § 225) and manifests not only our mutual comprehensive communication, as well as the possibility of understanding those changes brought, displaced or imposed on the field of meaning.

We can think of some propositions that involve the nature of our vulnerability and of our interdependence as relevant examples to the grammar of the living body, and of situational inscriptions as those acquiring new or old meanings according to the transformations that take place in our social and political practices and ways of living life. Let's think about how the vulnerability of old living bodies carries with it a constructed view, contextualized by culture, that is sometimes expressed by means of a compassionate language, but sometimes by means of an excluding one. Think of how the importance attributed to interdependence is modified by linguistic practices informed by feminist readings of what it means to bodily depend on other bodies. The grammar of the living body is thus liable to adjustments and change because our system of images and propositions (the totality of our judgments (OC 1969: § 140)) is open to change and to the critical scrutiny; so we can assess differently those concepts that are compositional to our collective experiences and that may guide our ethical and political practices. In this sense, the fixity of the body or its character as a socially situated subject (the subject as body), is not that of immutability, above all in terms of meaning (and this is precisely what it is about), but it is that of "shared assumptions" around which our judgments about the world, and about the construction of knowledge itself, revolve. This is how, says Wittgenstein, the fixed is also feasibility of movement:

OC 1969: § 152 – I do not explicitly learn the propositions that stand fast for me. I can *discover* them subsequently like the axis around which a body rotates. This axis is not fixed in the sense that anything holds it fast, but the movement around it determines its immobility.

This means that although a foundational – structuring – concept of our epistemological-linguistic – and moral – practices, the body is a hinge, and not irreducible sediment.

Thus, I would like to think that we can get rid of both dualistic epistemological strands and the foundationalist proposal by proceeding in the way of returning the body to the realm of meaning, taking it as an epistemic condition for our knowledge assertions, and for the determination of our beliefs assumed as true and justified by reasons that are always structured in the context of the grammar of the living body. Body, surface, and experience may count as constituting this grammar for us to be finally able to characterize the subject's situation in a contextualized way.

The experience of corporeality is the common ground for any epistemic communication, including or especially that of offering justifications and reasons in the processes of knowledge construction, a process that is mutually delimited and corrected by the scope of this language-game. A process of knowledge construction that is mutually delimited and corrected also by a criterion of objectivity that is itself built on the shared framework of epistemological-political-moral hinge concepts, and that is opposite to any claim of primacy or "purity" concerning our epistemological language. The horizon of this understanding is political because epistemic demands are all political, especially when supposedly and *per impossibile* claimed as a-political. It is from the understanding of material corporeality as epistemically axial that we can develop or unfold the understanding of situation as inscribed on its unquestioned surface – inscriptions that are the effects of gender, sexual, racial, colonial terminologies and technologies. And it is from this path that we can proceed to corrections and adjustments on semantic blindness, responsibility, authority, and epistemic injustice.

This still need to be done. Just as the loose ends of this essay still need to be tied up and all contributions to this are most welcome.

References

- Ashton, Alan (2019) "The case for a feminist hinge epistemology", *Wittgenstein-Studien*, v. 10, n. 1, 153-163.
- Gaita, Raimond (2004). *The Philosopher's Dog*. New York, London: Routledge.
- Haraway, Donna. (1985/2016). *A Cyborg Manifesto: Science, Technology, and Socialist Feminism in the Late Twentieth Century*. University of Minnesota Press.

Haraway, Donna (1988) "Situated Knowledges: The Science Question in Feminism and the Privilege of Partial Perspective", *Feminist Studies*, vol. 14 n.3, 575-599.

Scheman, Naomi (1995) "Feminist Epistemology", *Metaphilosophy*. vol. 26, n. 3, 177-90.
Published By: Wiley.

Wittgenstein, Ludwig (1969) *On Certainty*. Translated by Denis Paul and G.E.M. Anscombe.
Oxford: Basil Blackwell.

Wittgenstein, Ludwig (2009) *Philosophical Investigations*. Revised 4th edition by P.M.S. Hacker and Joachim Schulte. Wiley-Blackwell.

The Threat of Solipsism: Understanding the Net with Iris Murdoch and Susan Stebbing.

Francesca Scapinello (Lisbon, Portugal)

Abstract

This article explores Iris Murdoch's and Susan Stebbing's reactions to two aspects of Ludwig Wittgenstein's *Tractatus logico-philosophicus*: its reliance on logical symbolism and its "solipsistic conclusion" (Stebbing 1933a: 24). Murdoch and Stebbing, despite having different interests and conceptions of philosophy, make a similar claim against Wittgenstein's TLP: its use of symbolism leads to the wrong conclusion that "solipsism (...) coincides with pure realism." (TLP: 5.64) Murdoch and Stebbing mean different things when they accuse Wittgenstein of symbolism: Murdoch approaches the term from an artistic perspective, whereas Stebbing refers to its logical aspect. Still, they both think that symbolism presupposes a reduction of reality and denies the validity of common experience, which ultimately results in solipsism - a position that they reject. To do so, Stebbing resorts to commonsense knowledge of the existence of the world and to the notion of directional analysis; Murdoch instead argues that solipsism is an unfaithful representation of our moral life.

Without effacing their differences, in this article Murdoch and Stebbing are presented as virtual allies in contrasting some philosophical methods. More in general, they show the importance of debunking the assumptions we make in philosophy, and they voice the importance of retaining our ordinary experience when philosophising.

Introduction

"When you've been most warmly involved in life, when you've most felt yourself to be a man, has a theory ever helped you?" The utterer of this line is Annandine, one of the two characters of *The Silencer*, a book published by Jake Donaghue, the protagonist of Iris Murdoch's novel *Under the Net* (1954). Tamarus, Annandine's interlocutor, had just asserted that human beings are "rational animals in the sense of theory-making animals" (1954: 89) and the rest of the exchange between the two revolves around the role of theory in attaining knowledge of truth. Tamarus doesn't let go of the idea that theory is what makes life intelligible, and Annandine firmly holds the opposite: theory leads us astray and what we should do is "crawl under the net." (1954: 90) The limits of philosophical theory in relation to human life is something that Murdoch was deeply interested in. A part of her philosophical activity can be described as the effort to make us aware of our assumptions when we theorise, but also as a warning: some presuppositions compromise a whole system, a way of thinking, as it is the case, Murdoch thought, with existentialism, linguistic analysis, behaviourism, and structuralism. Susan

Stebbing, a generation-older philosopher, whom was known to Murdoch, manifested a similar uneasiness: she thought that the idea of philosophy as a theory should be discarded and that common sense is the beginning and the end of the activity of analysis. Murdoch and Stebbing are more idiosyncratic than similar: Murdoch rejects the idea that philosophy is only about analysis, and that there is a neutral, value-free description of facts. She thought that philosophy comprises a two way movement, analysis and construction (1997: 299), and that we see different worlds through different moral concepts (1997: 82). On the contrary, Stebbing believed in philosophy mainly as a clarification of our ordinary understanding, and we don't find the same kind of emphasis on the connection between ethics and epistemology as we do in Murdoch. Yet, both philosophers - Stebbing explicitly, Murdoch implicitly - raised quite the same issue about Wittgenstein's *Tractatus logico-philosophicus*: the theory (pace Wittgenstein) of that book worried Murdoch and Stebbing because of, in Stebbing's words, its "solipsistic conclusion."

The aim of this article is to explore and compare Murdoch's and Stebbing's reactions to solipsism as expressed by Wittgenstein in the *Tractatus*. Section I is dedicated to Stebbing's 1933 lecture titled *Logical Positivism and Analysis* (1933a) which contains a criticism of Wittgenstein's philosophy (the main allegation: its use of symbolism) and a description of what Stebbing thinks is the proper way of practicing analysis. Section II carves out Murdoch's position by considering some papers published during the 1950s and her 1953 book *Sartre. Romantic Rationalist*: the problem of Wittgenstein's symbolism is taken up by Murdoch too (although in a different way than Stebbing), together with the denounce of solipsism as an unacceptable conclusion. Section III contains a comparison between Murdoch's and Stebbing's attitudes towards common sense in order to contrast solipsism as (they thought) Wittgenstein took it.

Section I

The 1934 Winter Issue of the journal *Analysis* opens with three papers on the topic of solipsism. The authors are J. Wisdom, M. Cornforth and Stebbing and they all, despite their different approaches, argue against the solipsistic thesis that "no person other than myself exists" (1933b: 13) (this is how R. B. Braithwaite, and Stebbing after him, formulated it). In particular, Cornforth and Stebbing are responding to a paper (published the year before in the same journal) titled "Solipsism and the 'Common Sense View of the World'" (1933b),

where Braithwaite, the author, made the claim that some solipsistic statements are compatible with common sense. Braithwaite's stance was in turn formulated as a reply to Stebbing's 1933 Henriette Hertz Lecture to the British Academy. In that occasion, Stebbing had declared that "any philosophical view which leads to the conclusion that what solipsism *means* or *intends* to say is quite correct is, in my opinion, obviously false." (1933a: 25) The reason behind this exchange is Wittgenstein's *Tractatus*, more precisely proposition 5.62 which reads "This remark provides the key to the problem, how much truth there is in solipsism. For what the solipsist *means* is quite correct; only it cannot be *said*, but makes itself manifest. The world is *my* world: this is manifest in the fact that the limits of *language* (of that language which alone I understand) mean the limits of *my* world". For some, like Cornforth and Stebbing, proposition 5.62 epitomises Wittgenstein's "solipsistic conclusion" (1933a: 24). Following Stebbing's reconstruction of the *Tractatus*, which, she specifies, is informed by the reading of the logical positivists, the connection Wittgenstein draws between language and world responds to the 'principle of verifiability'. Stebbing's view can be summarised in this way: since "the meaning of a proposition is the method of its verification" (1933a: 18) (this is, for Stebbing, entailed by proposition 4.031), and given that ultimately verification depends on 'content', that is, facts, then the solipsistic conclusion follows: "every significant proposition which I assert says what could be verified only by my own direct experience, either present or future". (1933a: 18)

Stebbing wants to resist this picture, and she does so by resorting to G. E. Moore's philosophical practice. Stebbing challenges the solipsistic thesis that no other person beside myself exists by asserting that she *knows* it to be false. In doing so, she appeals to the validity of some common sense propositions, as Moore did in *A Defence of Common Sense* (1925), and justifies her position through the notion of directional analysis. Stebbing thought that there are two ways in which one can do analysis: there is same-level analysis, through which we attain clarity on what we say, and then there is directional analysis, which aims at clarifying what we are talking about. For Stebbing, what we talk about are facts which ultimately depend on basic facts, the ultimate elements of our reality: as indicated by Chapman (2013) and Coliva (2021), when delivering the 1933 lecture Stebbing was exploring the approach of metaphysical analysis,

which admitted the stipulation of metaphysical entities and didn't reject the possibility of metaphysics altogether, as the logical positivists did. Thus, Wittgenstein and the logical positivists were wrong in denying that there is a world of facts out there and in thinking that the only thing philosophy has to worry about is language and its logical constructions. Stebbing thought that these faults had to do with the fact that they didn't develop the very notion of analysis enough: not only they completely lack the directional sense of analysis, but they also don't distinguish between what Stebbing calls "analytic definition of symbolic expression" (1933a: 29) and "analytic clarification of a concept." (1933a: 29) Such inaccuracy, combined with the conviction that "all facts are linguistic facts," (1933a: 33) lead to the supposition that "the first problem of philosophy is to determine the principle of symbolism, and *from these principles* to draw limits with regard to what we *can* think." (1933a: 33) What Stebbing is opposing is the idea that we can fully rely on logical symbols in order to explain how our language works: if the aim of our philosophical exercise is to understand the limits of language, to do so with symbols means to rely on a priori assumptions about the nature of our language. Here lies the problem: for Stebbing we need to start from facts we know, not from logical deductive systems about language. Philosophy aims at the clarification of our thoughts, but it cannot be a linguistic investigation: "We cannot clarify our thoughts by thinking about thinking, nor by thinking about logic. We have to think *about* what we *were* thinking about." (1933a: 36) Stebbing thought that Wittgenstein's solipsistic conclusion is the direct consequence of "a theory with regard to the way in which language may become a significant symbolism" (1933a: 24): that is, a theory that allows for the idea that knowledge of the world only depends on one's direct experience, and thus the idea that each proposition is meaningful in so far as it refers to one's perception.

Section II

About 27 years later, in 1961, and without public reference to Stebbing's work, Iris Murdoch returned to the relation between Wittgenstein and symbolism. In *Against Dryness* (1961), Murdoch sets herself the task to illustrate the reasons behind the impoverishment of the 'idea of human personality' employed in the novels of her time. To do so, Murdoch looks at twentieth century novels by means of a comparison with the general tendency of the literary prose

produced in the eighteenth and nineteenth centuries. Twentieth century literature inherits, Murdoch notes, the unresolved issues raised by the Enlightenment: alongside with the works of the eighteenth century, “we retain a rationalist optimism about the beneficent results of education, or rather, technology.” (1997: 290/1) The human condition was addressed by ‘rationalistic allegories and moral tales’, and was grappling with an idea of the individual as ‘unitary and single’. Such endeavour was then swept away by nineteenth century novels, which connected “the idea of person with the idea of class” (1997: 291): their main concern was to address the struggle of the individual in a pre-modern Liberal society. What we find in twentieth century novels is, again, a portrayal of the individual as “stripped and solitary” (1997: 291) and, differently from the eighteenth century ones, a “crystalline or journalistic” (1997: 291) style. Murdoch describes the former as a “small quasi-allegorical object” (1997: 291) emptied out of the structure of characters developed in nineteenth century novels, and the latter as a “large shapeless quasi-documentary object” (1997: 291) which seems to be able to talk only about empirical facts. The crystalline style, Murdoch observes, is reminiscent of “the ideal of ‘dryness’ which we associate with the symbolist movement” (1997: 291/2). In the sentence that follows we find an unexpected figure associated to the literary movement of Symbolism: the one of Wittgenstein. As odd as it might seem, we might get a glimpse of the reason behind Murdoch’s connection when the concept of dryness is further spelled out as conveying ideas of “smallness, clearness, self-containedness” and when it is reported to be “the analogue of the lonely self-contained individual.” (1997: 292) The connection Murdoch is tracing is slightly similar to Stebbing’s: a certain way of approaching and expressing reality (might it be literary or philosophical) is connected to an implicit presupposition about what reality is. The idea that a pure, clear and self-contained symbol can represent human personality presupposes that the latter is eligible to be conceived of in these terms - a conclusion that Murdoch doesn’t want to accept.

In a previous article, *The Sublime and the Beautiful Revisited* (1959), Murdoch had traced some lines of continuity between the Symbolists, the linguistic empiricists and the existentialists in their abhorrence of messiness, in their aspiration to “precision and clear definition” (1997: 273), and in their deference to necessity over contingency. The symbol used by Symbolists

appears to be a distortion of what it wants to represent: it takes the unicity of the object of representation and it strips it out of its contingency. These are grave accusations, both towards Symbolism as an artistic movement and towards those philosophical postures associated to it: they fail in being truthful to the obscure and chaotic reality of our existence and divert our attention from what really matters, that is the realistic understanding of ourselves and of our morally coloured engagement with others. (To get a taste of the tone of the offence, Murdoch goes as far as saying that Plato would have approved symbolist art.) Murdoch's accusations towards existentialism and linguistic empiricism go further, and in Stebbing's direction, to the point of asserting that they both have "a terror of anything which encloses the agent or threatens his supremacy as a centre of significant. In this sense both philosophies tend towards solipsism." (1997: 269)

This is not the first occasion in which Murdoch draws a straightforward connection between existentialism, linguistic empiricism and solipsism. In *Existentialism as a Political Myth* (1952), Murdoch depicts the "solipsistic picture" (1997: 134) along the lines of what she takes to be Sartre's conception of individual consciousness: something "non-historical, non-social, and non-determined," (1997: 134) and she then argues that the logical analysts (a term used as a synonym for logical empiricists) assume the same picture of the individual. Other adjectives that Murdoch has for recounting the existentialist subject are "isolated" (1997: 134), "lonely" and "empty." (1997: 136) These traits, just mentioned in the paper, are deepened in Murdoch's book *Sartre. Romantic Rationalist* (1953), where Murdoch compounds Sartre's literary and philosophical works and their relation to his political efforts. Murdoch describes Sartre's man as painfully caged between the aspiration towards a non-existing God and the conscience of his impotence in changing reality; as free in the self-referential way admitted by such premises and as the bearer of the highest moral value, which is self-awareness. Others are not part of the constitution of the Sartrean existentialist subject: meaning, value and freedom are attained without reference to the world. Murdoch puts it this way: Sartre's individual "has a *dream* of human companionship, but never the experience. He touches others at the fingertips." (1953: 63) In this framework, other aren't considered as "objects of knowledge", but as "object to be feared, manipulated and imagined about." (1953: 129)

Such a vision of human individuality is, to Murdoch, conducive to a bad conception of morality, which results in being concerned only with oneself, instead of also including an outward-looking attitude, inclusive of other people. What is missing, Murdoch argues, is the appreciation of the multifarious aspects of human life and the acceptance of its muddy contingency. Morality must deal with this, not with unrealistic, as in partial, pictures of the individual. Values, for Murdoch, must be thought of as pervasive of reality, they cannot be severed from facts. The architecture of Wittgenstein's *Tractatus* is the epitome of what Murdoch thinks is wrong in philosophy (Murdoch talks about this in Chapter 2 of *Metaphysics as a Guide to Morals* (1992)): since values need to be left untouched, we can only speak about facts; the subject we can speak about isn't a moral (perhaps not even a human) being, but an "extensionless subject of experience" (1992: 27). Murdoch doesn't explicitly refer to the *Tractatus* when calls Wittgenstein a Symbolist, but the description she gives of the solipsistic picture of individuality fits the one presented in (or deduced from) the *Tractatus*: there is no history, no human relation, no specification of the subject. It is true that Wittgenstein talks about the metaphysical subject: this, however, doesn't constitute a good counter argument for Murdoch - metaphysical pictures of the subject should still mirror properly the reality of individuality. Wittgenstein's doesn't, and to Murdoch this seems enough of a reason to think of it as a solipsistic picture which should be discarded because untruthful.

Section III

Wittgenstein (at least according to some scholars like E.M Lange (2017) and H. Putnam (2008)) would have probably pushed back Murdoch's and Stebbing's allegations of solipsism: the *Tractatus* doesn't speak of the principle of verifiability in the way Stebbing and the Logical Positivists interpret it; a logical investigation on language is objective enough to prove that there is something else existing apart from myself; to say that solipsism coincides with realism is 'just' to point out that the subject is the necessary medium between language and world. Would these replies have tamed Murdoch's and Stebbing's criticism?

Probably not, as both philosophers saw serious problems in Wittgenstein's methodology, regardless of the solipsistic conclusion of the *Tractatus*. They mistrust the way in which Wittgenstein 'clarified' our philosophical problems:

his reliance on symbolism (in its logical sense for Stebbing and in the artistic connotation for Murdoch) leads us astray from what we know to be reality. For Stebbing, this fault is due to the idea that we can assume that logical symbols represent our language. We should instead rely on our common sense knowledge about the existence of certain facts, and then analyse our propositions in order to clarify some aspects of those facts. Wittgenstein's symbolism is thus a misrepresentation of reality, as it starts from the wrong premises. (As described by Chapman (2013), Stebbing's philosophical activity underwent some change, especially from being directed towards and academic to a public audience. Although she abandoned some ideas expressed in the *Lecture*, what Chapman shows is that Stebbing's interest towards ordinary use of language and common sense remained constant.) In Murdoch we find a different argument: the theoretical stylisation proper of symbolism comes from a lack of what she calls moral vision. Without it, philosophy produces theoretical frameworks which reduce the complexity of our lived experience, ending up trapping us in a false picture of ourselves and of what surrounds us. But what we need to resort to isn't just analysis: for Murdoch we must develop an attentive posture, which will make us see reality and others more clearly. We need philosophy to reflect this attitude: only in this way, as Murdoch wants, it will be able to explore moral concepts and enlarge our moral vocabulary.

The dissimilarities between Murdoch's and Stebbing's philosophies are undeniable: the different ways in which they use the term symbolism is alone a good example to signal their distance in methods and interests. Yet, they virtually join forces in denouncing those theoretical assumptions which reduce our ordinary reality to a symbol, thus effacing its complexity and validity. Both philosophers show that any philosophical activity cannot lose sight of our common experience: if we misrepresent (Murdoch) or deny (Stebbing) our common sense in dealing with morality (Murdoch) and with perception and use of language (Stebbing), we will end up with false conclusions, like solipsism. What they seem to suggest is that if we want to get rid of the wrong theories, we must do so trusting the common grounds upon which we crawl.

References

- Braithwaite, R. B. (1933b) "Solipsism and the 'Common Sense View of the World'", *Analysis* (1): 13-15.
- Chapman, Shioban (2013) *Susan Stebbing and the Language of Common Sense*, Basingstoke: Palgrave Macmillan.
- Coliva, Annalisa (2021) "Stebbing, Moore (and Wittgenstein) on common sense and metaphysical analysis", *British Journal for the History of Philosophy*: 1-21.
- Cornforth, Maurice (1934) "Is Solipsism Compatible With Common Sense", *Analysis* (2): 21-26.
- Moore, G. E. (1925) *A Defence of Common Sense*, in J. H. Muirhead (ed.) *Contemporary British Philosophy*, reprinted in Moore G. E. (1959) *Philosophical Papers*, London: George Allen and Unwin Ltd.
- Murdoch, Iris (1953) *Sartre. Romantic Rationalist*, London: Vintage.
- Murdoch, Iris (1954) *Under the Net*, London: Vintage Classics.
- Murdoch, Iris (1992) *Metaphysics as a Guide to Morals*, London: Penguin Books.
- Murdoch, Iris (1997) *Existentialists and Mystics, Writings on Philosophy and Literature*, Peter Conradi (ed.) New York: Penguin.
- Stebbing, Susan (1933a) "Logical Positivism and Analysis", *Proceedings of the British Academy* (XIX): 1-37.
- Stebbing, Susan (1934) "Concerning Solipsism: Reply to R. B. Braithwaite", *Analysis* (2): 26-28.
- Wisdom, J. O. (1934) "Solipsism", *Analysis* (2): 17-21.
- Wittgenstein, Ludwig (1921) *Tractatus logico-philosophicus*, Oxford: Blackwell.

Acknowledgement

This work was supported by FCT – Fundação para a Ciência e Tecnologia, I.P. with the scholarship UI/ BD/154631/2023 and in the scope of project UIDB/UIDP00310/2020 with DOI identifier 10.54499/UIDB/ 00310/2020.

Scepticism and „Moorean facts“ in *On Certainty* – a deflationist reading

Lauritz Seeberg (Leipzig, Germany)

Abstract

„What is your aim in philosophy? – To shew the fly the way out of the fly-bottle. (PI §309)

The fly bottle is the right metaphor for philosophy insofar as here too the philosophers (flies), irritated by the promise of knowledge in the form of the bait, seek an exit other than the one they have used as an entrance. However, the only exit that exists is the entrance. So the only option to exit the fly-bottle is to return right back the way you came in. The most interesting twist in the fly-bottle metaphor is that the exit (i.e. the entrance) was never closed and should have been known to all those who found their way in, so that they could leave the glass at any time. For the flies that recognize the entrance as the exit, the fly-bottle is no longer a prison. They can leave it at any time and return to it whenever they want to. „The real discovery is the one that makes me capable of stopping doing philosophy when I want to.“ (PI §133.) In this paper, I use Wittgenstein's "treatment" of scepticism and G.E. Moore's "facts" in *On Certainty* to demonstrate his method of doing philosophy without being "bewitched" by it. Through these examples, I argue in defence of a reading of Wittgenstein's works that leads to a deflation of traditional philosophical questions and problems:

„And so it is not to be wondered at that the deepest problems are really no problems.“ (TLP 4.007)

„Das Kind kann die Namen von Personen gebrauchen, lang ehe es in irgendeiner Form sagen kann: „Ich weiß, wie Dieser heißt; ich weiß noch nicht, wie jener heißt.“ (ÜG 1984: 543)

„Das Kind muß die Verwendung von Farbnamen lernen, ehe es nach dem Namen einer Farbe fragen kann.“ (ÜG 1984: 548.)

Philosophers learn to use language just like all other people. In his *Philosophical Investigations* Wittgenstein once describes the process of learning the first language as „Abrichtung“ („training“, but not used for humans, but for dogs for example. It's an involuntarily form of training).

„Ein wichtiger Teil der Abrichtung wird darin bestehen, dass der Lehrende auf die Gegenstände weist, die Aufmerksamkeit des Kindes auf die lenkt, und dabei ein Wort ausspricht; z.B. das Wort ‚Platte‘ beim Vorzeigen dieser Form. (Das will ich nicht ‚hinweisende Erklärung‘, oder ‚Definition‘, nennen, weil ja das Kind noch nicht nach der Benennung fragen kann. – Ich sage, es wird einen wichtigen Teil der Abrichtung bilden, weil es bei

Menschen so der Fall ist; nicht, weil es sich nicht anders vorstellen ließe.)
[...]“ (PU 1984: §5-6)

This quote and also the quotes from *On Certainty* quoted at the beginning show that the process of learning your first language is something that happens to you rather than something you do on purpose.

Re-educating philosophers back to ordinary language

In contrast to "trained" ordinary language, philosophising, considered as the use of philosophical language games, is something that can be learnt or not learnt. Traditional philosophical learning primarily involves the acquisition of a specialised language, just as in other subject disciplines. The special feature of philosophical language, however, is that ordinary language itself is one of the objects of examination, which is why it could also be described as a "metalanguage". "Meta-language", though, does not mean that there is a second language, independent of ordinary language, for speaking about ordinary language, but that ordinary language is used contrary to its originally "trained" ordinary use in order to speak about itself. If I understand Wittgenstein correctly, he considers this very special kind of use to be potentially harmful, which is why he also speaks of "misuse" (PU 1984 §6) in this case. The problem with this is the absence of language game rules to talk about the rules of language games, because this is simply no longer part of ordinary language. Away from ordinary language, a jumble of language games therefore emerges that attempt to provide answers to questions that only arise in these philosophical language games. The harm that arises from the misuse is that within the philosophical language games there is the opposite of a basic consensus on the way the language games are used: A dissent about the way of use that is impossible to end. The "Verhexung" (ÜG 1984: 109.) ("bewitchment") of language consists in the traditional philosophy mistakenly seeing its task as establishing hegemonic ways of use in this doomed field. From Wittgenstein's point of view, taking a step back would be necessary here: The philosophical language games should not be reworked further and further from a Meta-language perspective, but should be analysed as a whole from the ordinary language perspective in order to dissolve them by returning them to the ordinary way of use.

What Wittgenstein does in his works is, on the one hand, to make the nonsensical nature of traditional philosophical problems visible by comparing ordinary and philosophical language games and, on the other hand, teaching this method by demonstrating it through examples. This teaching is essentially a "re-education" of philosophers to ordinary language. Based on the ordinary use of language, a "deflation" of philosophical problems can be brought about by precisely analysing language games. In ÜG, the main object of this method is the expression "I know" and related expressions.

Dissolving nonsense – deflation

Wittgenstein would characterise most traditional questions or problems in philosophy as nonsensical. The reason for this is that there can be no "real" metalanguage, i.e. one that is detached from ordinary language. Every language must be based on ordinary language because the meaning, i.e. also the standards ("Maßstäbe") for truth and falsity, lie in the use of ordinary language because, as explained in the beginning, it precedes the acquisition of philosophical (and all other) language games. If this fact is ignored, it inevitably leads to philosophers discussing the form of language representation, which can only be shown through its use, as content. For this reason, however, the resulting supposed insights and problems do not offer the desired "certain" truths about the nature of reality, but rather aberrations guided by the peculiarities of language - nonsense.

Philosophers create their own problems by using abstract words that lie outside ordinary language use. In contrast, Wittgenstein assumes that "understanding" refers to what you mean when you say "I understand", and "knowledge" refers to what you mean when you say "I know" and so on. In other words, he assumes that the meaning has always already been understood. It is imperative that the existing definitions of these terms are not redefined, as this can lead to self-referential nonsense. Instead, philosophers should accept the definitions of these terms already established through their use and not impose new meanings on them.

Wittgenstein's method

Wittgenstein's method is therefore more humble in its aspirations. His philosophy considers its object to be the investigation of the structure of language in the manifold language games, especially those of traditional

philosophy. He therefore does not want to say anything new, only to make the use of language, as every child is already "trained" to use it, visible and understandable. He does not articulate a theory (in the scientific sense), but shows us his method. By "show" he does not mean an evocative experience that is beyond the "limits of the sayable", but a demonstration of his method of analysing propositions of traditional philosophy and contrasting them with ordinary language by means of examples. These examples - Wittgenstein calls them "Erläuterungen" (explanations) - are neither theories nor "anti-theories". Essentially, they are approaches to (dis-)solving by someone who engages in philosophical language games but does not allow himself to be "bewitched" (PU 1984: §109) by them, but instead brings them back into the ordinary language consensus. His method may be described as surprising, but the results of the method should not come as a surprise. Almost all far-reaching theories that refer to Wittgenstein therefore fall short of him in this respect. Of course, there are many brilliant theories that interpret Wittgenstein correctly, draw on him and legitimately extend him. Nevertheless, they do not do him justice in this respect and fall back into the never-ending philosophical language games.

„Komme ich nicht immer mehr und mehr dahin zu sagen, daß die Logik sich am Schluß nicht beschreiben lasse? Du mußt die Praxis der Sprache ansehen, dann siehst du sie.“ (ÜG 1984: 501.)

On Certainty

Wittgenstein does not articulate a theory in *On Certainty* either. Nevertheless, his explanations revolve around two explicitly named poles: Moore's alleged knowledge and the sceptical alleged non-knowledge. I think that Wittgenstein uses the method described in ÜG to take a close look at G.E. Moore's alleged "knowledge" on the one side of the extremes and sceptical alleged non-knowledge on the other. The sceptic claims the impossibility of knowledge as long as no ultimate justification is given. Moore asserts "Here is one Hand" (Moore 1993: p. 166) and similar statements as a matter of common sense.

However, Wittgenstein is less interested in the content of the sentences than in their peculiar language games. To say it in advance: I think that in ÜG he wants us to realise that the language games are so alienated from ordinary language

that they have no content at all, are meaningless. What Wittgenstein does not write, but what we can draw from his analysis, is that the sceptical sentences cannot be true because they undermine themselves and the Moorean sentences cannot be false because they express such fundamental truths that they would also call into question the truth of every other truth. I will explain this view below, beginning with my understanding of Wittgenstein's concept of knowledge.

Ordinary use and philosophical misuse of the concept of knowledge

We can speak of "knowledge" where it can be both true and false. Not at the same time of course, but there has to be the possibility for it to be different: i.e. the true sentence must be theoretically possible to be false and vice versa. This follows from Wittgenstein's anti-psychologistic stance and excludes private "truths" (see the remarks on the impossibility of private language from PU §243 onwards) as well as metaphysical revelations ("Gott offenbart sich nicht in der Welt." LPA 1984: 6.432).

Sentences that are necessarily true (tautologies) or necessarily false (contradictions) disqualify themselves as something that Wittgenstein would call "knowledge" because they do not describe anything in the world. At first glance, this may sound like a philosophical theory in the form of a definition. However, as I described earlier, such definitions that deviate from ordinary language are exactly what Wittgenstein is criticizing. In fact, I believe that with these "criteria" Wittgenstein is simply reflecting a certain tendency in the everyday use of the concept of knowledge. Namely, the basically uncontroversial assumption that we should provide evidence when we assert something (see also ÜG 504.).

Here, one may feel reminded of Plato's third definition of knowledge in the Theaetetus dialogue (to which Wittgenstein explicitly refers in a different context in PU §46). It states that knowledge is true, justified belief and, even if it is refuted in the dialogue itself, represents a milestone in the (so called) Western development "from mythos (μῦθος) to logos (λόγος)". My idea would be to interpret Wittgenstein's concept of knowledge as a recourse to this method of understanding the world through rational (empirical) investigation, without metaphysical, mystical belief, which has been so formative for our culture. Of course, Wittgenstein does not give us an explicit definition of the

concept of knowledge, because this simply does not exist in ordinary language. Through his numerous examples, however, he demonstrates to us family resemblances in ordinary usage in which one property remains constant: "knowledge" are statements for which there is the possibility that they could potentially turn out to be both true and false (notable) for everyone.

Sceptical sentences and Moorean sentences

Now back to the explanation of why I claimed that sceptical sentences are contradictions and Moorean sentences are tautologies.

1. A sceptical sentence, such as "I cannot be sure that the world existed before I was born", is first of all not an ordinary sentence and does not stand in any ordinary context. Perhaps with a lot of imagination it would be possible to find a context in which the sentence could be used in an ordinary way. But there is no such context here. The sentence represents a categorical mistake because it confuses a property of language form with a sayable content. The statement of this sentence itself uses the very form of language that it casts doubt on. No matter which sceptical or non-sceptical sentence one states as long as it is "sinnvoll", i.e. follows the rules of the ordinary language game, one thereby demonstrates the acceptance of the validity of all other rules of ordinary language games. (This is also how I understand Wittgenstein's first sentence in *ÜG*: "Wenn du weißt, daß hier eine Hand ist, so geben wir dir alles übrige zu.") The latter includes, of course, that we would not say something like "I cannot be sure that the world existed before I was born". The utterance of the sentence, i.e. the act of uttering it, is already the counterevidence to the assertion of the sentence. You could roughly compare it to a person who says, "I caught a living elephant in here" and presents a shoebox. We don't need to examine the contents of the shoebox to know that the sentence is false, because the form of the box already makes it impossible to be true. Similarly, we do not need to examine the sceptical sentence because its assertion does not fit the form of the shoebox, i.e. the language game that it demonstrates. He demonstrates it by uttering the sentence, just as someone demonstrates a shoebox. The sentence asserts A, but the language game that the sentence presupposes, i.e. also simultaneously asserts, asserts not-A. Such a sentence cannot be true and is therefore a contradiction.

2. A Moorean sentence such as: "I know that the world existed before I was born" (Moore 1993 p. 167) is also not a normal sentence not in an everyday context. („Ich kann mich darin nicht irren“ ist ein gewöhnlicher Satz, der dazu dient, den Gewißheitswert einer Aussage anzugeben. Und nur in seinem alltäglichen Gebrauch ist er berechtigt.“ ÜG 1984: 638.)

Here again, with a lot of imagination, it might be possible to find a context in which the sentence could be used in an ordinary way. However, there is no such context here either. This sentence also represents a categorical mistake because it expresses such a profound truth that all knowledge presupposes it, so it cannot be false. This may also sound like an abstract theory at first, but in my opinion, it is again just a representation of how we normally speak about "knowledge". How we certainly don't talk about knowledge is the way Moore does it: listing individual unrelated sentences without context and without realistic use. Rather, we talk about knowledge as an interconnected network of truths (Wittgenstein calls it a "Nest von Sätzen" (ÜG 1984: 225.) (nest of sentences)) that are mutually dependent, have each other as their foundation or build on each other.

Should this Moorean sentence be subject to honest doubt in a realistic scenario (which is impossible), the entire "Weltbild" (ÜG 176.) (worldview) would also be in doubt, because almost all knowledge that we believe we have is directly or indirectly related to this sentence. After all, what else could we say we know if we were wrong about something so central?

As I have already shown above (1.), such a sceptical doubt is not actually possible because it undermines itself. The Moorean sentence cannot be "sinnvoll" doubted (ÜG 1984: 8.), for the reasons already mentioned. „Wenn einer bezweifelte, ob die Erde vor 100 Jahren existiert hat, so verstünde ich das *darum* nicht, weil ich nicht wüßte, was dieser noch als Evidenz gelten ließe und was nicht.“ (ÜG 1984: 231.) Insofar as the sceptical sentence is considered a contradiction, the Moorean sentence must also be considered a tautology, as it cannot be proven false. The same applies to other Moorean sentences.

Taken together with Wittgenstein's examples of how "knowledge" is normally used, that is, not for sentences that are necessarily false or necessarily true, it turns out that his "problem" or perhaps rather his "fascination" with Moorean sentences is that they do precisely that anyway. I believe that Wittgenstein

wants to point out that Moores defence against scepticism by using the "I know ..." -sentences fulfil exactly the opposite of their intended purpose: They make scepticism possible in the first place.

This may sound controversial again, but it is only logical: for Wittgenstein, Moorean sentences would only become doubtful if they were given the status of "knowledge", because, as mentioned several times, "knowledge" is doubtful for him by definition. Moore thus puts his own sentences up for questioning, which is a convenient opportunity for the sceptic. For Wittgenstein, these sentences must be *out of question* in the literal sense, just as they are in ordinary circumstances in everyday life.

Moorean sentences are therefore for Wittgenstein a prime example of how traditional philosophy inadvertently provides its own work and creates its own problems. If Moore did not engage in deviating from ordinary language, the philosophical problems he wishes to treat would not exist at all. Wittgenstein, on the other hand, attempts to demonstrate a form of therapy that is not at the same time the illness.

„Das Seltsame ist, daß, wenn schon ich es ganz richtig finde, daß Einer den Versuch, ihn mit Zweifeln in dem Fundamente irrezumachen, mit dem Wort ‚Unsinn!‘ abweist, ich es für unrichtig halte, wenn er sich verteidigen will, wobei er etwa die Worte ‚Ich weiß‘ gebraucht.“ (ÜG 1984: 498.)

The „I“ in „I know“

A second critique by Wittgenstein focuses on the formulation of Moorean sentences: The emphasis on the "I" in "I know...". The fact that Moore constantly expresses himself as if his knowledge is a property that primarily belongs to him personally, at least hints that it is subjective knowledge.

(„Zu wem sagt Einer, er wisse etwas? Zu sich selbst oder zu einem Andern. Wenn er's zu sich selbst sagt, wie unterscheidet es sich von der Feststellung, er sei *gewiß*, es verhalte sich so? Es gibt keine subjektive Sicherheit, daß ich etwas weiß. Subjektiv ist die Gewißheit, aber nicht das Wissen.“ ÜG 1984: 245.)

This formulation similarly paves the way for something Moore is actually trying to get rid of: Idealism and psychologism. Wittgenstein therefore already emphasises his anti-psychologistic stance in the second remark of ÜG:

(„Daß es mir – oder Allen – so *scheint*, daraus folgt nicht, daß es so *ist*.“ ÜG 1984: 2.)

Just as the rules of language are negotiated in a constant public game, "knowledge" must also be public. Just as the meaning of a word cannot be created privately, the correctness or incorrectness of a proposition cannot be judged subjectively. "Private meaning" and "subjective knowledge" are oxymorons. They are a good example of how quickly the anarchy of rules in philosophical language can trap us in unsolvable (illusory) questions.

"Die tiefsten Probleme [sind] eigentlich keine Probleme.“ LPA 4.003)

References

Kienzler (2005): Kienzler, Wolfgang, *Was ist Philosophie?* In: Rentsch, Thomas (Hrsg.) *Einheit der Vernunft? Normativität zwischen Theorie und Praxis*. Paderborn: Mentis. S. 13-24.

Kienzler (2016): Kienzler, Wolfgang, *Ein Sack Rosinen: Über Wittgensteins Arbeitsweise und seine Vorstellung vom Ziel seines philosophischen Arbeitens*, Wittgenstein-Studien, Bd. 7, Heft 1 (2016): S. 15-38.

Moore (1993): Moore, George Edward, Baldwin, Thomas (Hrsg.) *Selected Writings*, London: Routledge.

Wittgenstein, Ludwig (1984a): *Werkausgabe Band 1. Tractatus logico-philosophicus/ Tagebücher 1914–1916/ Philosophische Untersuchungen*, Frankfurt am Main: Suhrkamp.

Wittgenstein, Ludwig (1984b): *Werkausgabe Band 6. Bemerkungen über die Grundlagen der Mathematik*, Frankfurt am Main: Suhrkamp.

Wittgenstein, Ludwig (1984c): *Werkausgabe Band 8. Bemerkungen über die Farben/Über Gewißheit/Zettel/Vermischte Bemerkungen*, Frankfurt am Main: Suhrkamp.

Wittgenstein, Ludwig (2015–): *Wittgenstein Source Bergen Nachlass Edition (WS-BNE)*. Edited by the Wittgenstein Archives at the University of Bergen under the direction of Alois Pichler. In: Wittgenstein Source, curated by Alois Pichler (2009–) and Joseph Wang-Kathrein (2020–). (N) Bergen: WAB.

A Feminist Approach to Verbal Humour

Mira Magdalena Sickinger (Vienna, Austria)

Abstract

In this paper I investigate whether verbal humour can serve as a subversive tool for feminist aims. I examine standard theories by Raskin and Attardo that regard opposing scripts as the essential property of verbal humour. Challenging script opposition and shifting the focus on conceptual re-evaluation, I refer to an alternative theory, proposed by Bing and Scheibman. Their model of blended spaces builds on the cognitivist theory of conceptual blending, where frames from established domains are combined to yield hybrid frames. By laying emphasis on the generation of new meaning in humorous communication, this model proves to be suitable for a purposeful investigation of subversive humour. I critically address the question how fictional narratives of verbal humour may have an impact on real social circumstances and argue that the politically relevant aspect of verbal humour lies the oscillation between conceptual play and linguistic efficacy. Though not all subversive humour is feminist or politically productive, the conclusion shows that joking, as a practice of re-signification, allows women and other marginalised groups, who are typically excluded from active participation in the language game of joking, to playfully challenge the status quo. Possibly initiating changes in the shared understanding of our concepts.

Introduction

Verbal humour is a common component of communication. Linguistic and cognitivist positions have examined the forms and reception of humorous communication. Feminist positions on verbal humour are rare and differ in their attitudes: on the one hand, there are accounts that point to the risks of humorous speech reinforcing prejudices and stereotypes; on the other hand, there are accounts that see humour as a potential force that may subvert oppressive structures. This paper is a contribution to the latter accounts. By drawing on linguistic and cognitivist contributions to verbal humour (Raskin, Attardo, and Coulson), as well as feminist accounts of humour and language (Bing and Scheibman, Marvin, Kramer, Kotthoff, and Hornsby), I investigate the receptive processes elicited by humorous communication, i.e. the processes of reception, comprehension, interpretation, and response to verbal humour, which can entail “re-evaluations of our own cherished (and preidentified) rules, heuristics, and conventional categorizations”. (Kramer 2020, 173) It is my hypothesis that verbal humour can indeed serve as a subversive tool for feminist aims.

Verbal Humour

As a linguistic variant of the theory of incongruity, the *script-based semantic theory of humor* (SSTH) argues that for a text to be humorous, it must support two overlapping and contrasting scripts. (Raskin, 1985) A script, for Raskin representing a “large chunk of semantic information surrounding the word or evoked by it. (...) Every speaker has internalized rather a large repertoire of scripts of ‘common sense’ which represent his/her knowledge of certain routines, standard procedures, basic situations, etc., for instance, the knowledge of what people do in certain situations, how they do it, (...)”. (Raskin 1985, 81) Jokes following this structure typically begin by evoking one script, while switching to another in the punchline. This is where the boundaries between the overlapping scripts become apparent. The infamous “doctor-wife-patient-joke” by Raskin (see Joke Appendix) serves as an example for the analysis of the scripts’ interplay.

What is particular to jokes, is the effect caused by the recognition and resolution of the incongruity perceived, which might illicit a response of amusement. Very often jokes use characterizations or stereotypes to develop an efficient narrative; drawing on the ways in which humans engage with or perceive the world in a simplified or exaggerated manner. These stereotypes might get twisted or thwarted in the punchline, however, by drawing on common places they facilitate a “general understanding”.

As an extension and revision of Raskin’s SSTH, Attardo and Raskin put forward a *general theory of verbal humor* (GTVH), focusing on verbal jokes as its most representative subset and determining six knowledge resources informing the joke, namely “script oppositions, logical mechanisms, situations, targets, narrative strategies, and language”. (Attardo and Raskin 1991, 293) Both Raskin and Attardo refer to Grice’s pragmatic considerations of successful communication (Grice 1975) and conclude that verbal humour does not adhere to Grice’s cooperative principle. (see Attardo 1990) They, however, hold that jokes can be communicated successfully, although a different mode of communication is required; a “non bona-fide” mode. (see Raskin 1985)

Verbal Humour as a Subversive Tool

Challenging the idea of verbal humour deriving from script opposition, Bing and Scheibman (2014) propose a model of blended spaces that transcends

generic scripts. The authors argue that the type of humour resulting from blended spaces “can be subversive because of the creation of new possible worlds—situations that suggest alternatives to the status quo”, investigating how it “challenges and subverts existing norms that marginalize some groups”. (Bing and Scheibman 2014, 14) Bing and Scheibman’s model builds on the cognitivist theory of conceptual blending, explaining how “frames from established domains (known as *inputs*) are combined to yield a hybrid frame (a *blend* or *blended model*) comprised of structure from each of the inputs, as well as unique structure of its own”. (Coulson 2001, 115) This view points to the potential of the receptive processes elicited by verbal humour to transcend established representations, subvert normative expectations, and find or form new meaning. Coulson, who investigates the connections between cognitive models and joke comprehension claims that people “play with each other’s ability to update their representations adaptively when they tell each other jokes”. (Coulson 2001, 32)

There are different possible responses to humour. Humour can elicit “funny ha-ha”, “funny huh”, (see Morreall 1987, or Hurley et al. 2011) or “funny aha!”. (see Kramer 2020, or Sickinger 2023) While “funny ha-ha” might be a response to straightforward, basic forms of humour, “funny huh” a response to a feeling of unusualness or strangeness, and “funny aha!” a response of spontaneous humour comprehension, there are also cases where perceived humour may lead to a process of conceptual reflection that goes beyond immediate amusement; possibly framed as something like “funny (...) oh”. Some instances of verbal humour bring unspoken assumptions to light that are presupposed in the way we think and talk, (see Yablo 2019, 3) perhaps initiating a deeper enquiry of certain linguistic, conceptual or societal issues or conflicts. Some jokes can have “disquieting” effects, (see Wittgenstein, PI 2009: §111) destabilising our “scripts of common sense”. The unsettling impact that this effect could have on the individual recipient might be directed outwards and understood as a productive and collective disruption with regard to the subversive joke. I propose that the potential of subversive humour for feminist aims is not confined to the individual’s sense of disquietude but rather lies in a collective communicative practice of socially related language users. For an argument how “communicative speech acts”, a concept missing from malestream accounts of linguistic meaning, could allow for a fruitful

exploration of feminist (or other political) themes, one might refer to Hornsby (2000).

Bing and Scheibman propose the following joke by performer and activist Robin Tyler as an example of a potentially subversive humorous conceptual blend:

If homosexuality is a disease, let's all call in queer to work. "Hello, can't work today. Still queer."

In our interpretation of this joke, we do not necessarily derive amusement from the incongruity between the "illness" script and the "queer" script. What the joke seems to accomplish instead, by blending cultural pathologization of "homosexuality as a disease" with labour legislation and the activity of "calling in sick," is to show that homosexuality is indeed inconsistent with the social implications of the notion of disease. "In this way, then, the blend undermines the authority of homo-phobic expectations. Not only is a new meaning created in the blend, the meaning is performative due to its activist message." (Bing and Scheibman 2014, 19)

A critical voice might legitimately ask at this point, how fictional narratives of verbal humour may penetrate reality. This is indeed an important issue, which I have addressed extensively elsewhere, referring back to the pragmatic considerations of Raskin and Attardo. (Sickinger forthcoming) I generally hold that it is important to make pragmatic distinctions between the communicators involved in the performance of verbal humour, i.e. joke teller and recipients, and the fictional narrative/dialogue featured within the joke content. Although joke tellers carry the responsibility of evaluating the appropriateness of the jokes they are telling and the environment they are telling it in, an essential property of joke telling lies in the way it allows taboo topics to be addressed without necessarily attaching an explicit personal judgement.

Verbal humour seems to move between fictionality and reality in an interesting way that cannot simply be dismissed with positions arguing that there is no overlap between humour and reality or positions that ignore the pragmatic implications of joke communication. The politically relevant aspect of verbal humour could lie precisely in this oscillation between, perhaps

provocative, conceptual play and linguistic efficacy. This thought is shared by Kramer, who emphasises that the aim of subversive humour is “to meaningfully connect a funny fictionalization with a serious reality (...) the meaning and intentions in our playfully constructed scenarios can bleed into reality in such a way that listeners can be persuaded to see things from a fresh perspective”. (Kramer 2020, 169)

Political subversive use of jokes has a tradition that demonstrates its impact on real social circumstances. Think of jokes about the Stasi (the repressive *Ministry for State Security* of East Germany, active in the years from 1950 to 1990) told in the GDR (DDR) during the communist regime. Such regime-critical jokes were viewed by the Stasi as subversive propaganda and could lead to imprisonment, which are in fact real-life consequences of humour. This means that such jokes were indeed regarded as having potential political impact. Although these are examples with especially complex pragmatic conditions, even in such cases we can hold on to the claim that the pragmatic boundary of fictionality may not be dissolved, however, the motivation and possible effects of telling such jokes clearly transgresses fictionality. If we tell jokes about Trump, we may create fictional narratives around the figure, however, the joking practice may also serve to poke fun at the actual political person.

Kotthoff expounds how verbal humour, though often consisting of fictional narrative, can intervene in our social reality. “Humorous communication plays an important role in the production of normality and normativity (...) By violating norms and creating unconventional perspectives, humor certainly influences norms. It creates new, unusual perspectives on the object and thereby communicates sovereignty, creative power, and the freedom to intervene in the world. Such demonstrative displays of subjectivity and their potential to define normality have been less accepted by women.” (Kotthoff 2006, 5)

Verbal Humour as a Subversive Tool for Feminist Aims

Humour is undeniably a male-dominated communicative practice, with the stereotype of a joke teller as a figure resembling the likes of Woody Allen. Though we are seeing more non-male comedians emerge, the presumed intellectual superiority, political controversy or sexist vulgarity displayed in many instances of verbal humour actively prevent women and other

marginalised groups from participating in the performance of verbal humour or the practice of joke telling, given the social expectations placed on them. Because of their systematic oppression they are typically considered humourless, presuming two things: Women and other marginalised groups are too involved in their oppressed state, therefore too serious for joking. Women and other marginalised groups are too fragile to participate in the practice of humour communication and possibly utilise it for political and subversive purposes. Feminist philosophers and theorists have pointed out that humour, as a mechanism of oppression “can reinforce patriarchy and silence women’s voices”. (Marvin 2023, 166) Here, Marvin refers to publications by Bergmann (1986), hooks (1992) and MacKinnon (1979). Marvin, who has also authored an insightful overview of feminist philosophy of humour (2022), instead places the focus on the work that humour can do as a “weapon against oppressive language and concepts”. (Marvin 2023, 163)

Claiming the performance of subversive verbal humour as a feminist instrument may be an effective tool against forms of oppression. Especially considering the fact that humour opens up safe spaces of speech that can be disguised as “we are joking”. By actively participating in the language game of joking, women and other marginalised groups may take the risk of telling jokes that blend commonplace concepts, playfully challenging the status quo and possibly initiating a shared process of new meaning generation. While it is certainly true that not all subversive humour is feminist or politically productive, joking, as a practice of re-evaluating and re-signification can, e.g. lead to a change of the shared understanding of a certain concept or script. This effect may be comparable to the re-appropriation of slurs, pejorative expressions, or hate words. In this line, Hornsby argues how “uses of hate words, in enough new contexts, may serve to counteract the possibility of their hateful use, and thus to change their common understanding.” (Hornsby 2000, 99)

By raising awareness for the “instability” of subversive humour, Marvin points to the risks involved in the practice of subversive humour, which might be received or interpreted in ways other than intended. As the effects of and responses to humorous communication cannot be fully controlled, it seems important to evaluate how our attempts at subversive humour could be taken up. “When language, concepts, and affects are recontextualized and

rechannelled, the chips do not always fall where we hope or intend.” (Marvin 2023, 183) However, and I agree with the author’s position, Marvin also suggests that we might understand the risks involved in subversive humour as a value, leaving its effects open and complex. Against the background of such a differentiated understanding of the pragmatic dynamics of verbal humour, it appears possible to employ humour as a subversive tool for feminist aims.

Conclusion

The examination of standard theories of verbal humour by Raskin and Attardo has shown that script opposition falls short of recognising the creative receptive processes of conceptual re-evaluation and new meaning generation that may be elicited by humorous communication transcending normative scripts. Instead, Bing and Scheibman’s model of blended spaces has proven to be suitable for a purposeful investigation of subversive humour. The analysis of a joke-example by Robin Tyler has demonstrated that subversive humour may generate new and political meaning via blending the concepts conveyed in the joke. The critical examination of the question as to how fictional humorous narratives can affect social reality has shown that while acknowledging the pragmatic implications of joke communication is necessary, subversive humour can transgress the boundaries of fictionality by prompting recipients to engage with new perspectives. I have addressed the problem that joke practice is inherently male dominated, traditionally excluding women and other marginalised groups from participating actively in this language game. I have argued that oppressed groups gain the opportunity to challenge the status quo in a playful way by appropriating the practice of joke telling. With reference to Marvin, I have pointed out the risk that reception and interpretation of subversive verbal humour may not always coincide with a politically progressive intention. Accepting these risks and keeping the effects of subversive humour open and complex, however, has emerged as a meaningful approach. Though not all subversive humour is feminist or politically productive, the conclusion shows that joking, as a practice of re-evaluating and re-signification, can serve as a subversive tool for feminist aims.

Joke Appendix

“Is the doctor at home?” the patient asked in his bronchial whisper. “No,” the doctor’s young and pretty wife whispered in reply. “Come right in.” (Raskin 1985, 100)

In response to this infamous joke, Bing and Scheibman polemically argue how Raskin’s extensive analysis of the contrasting, but overlapping “doctor” script and “lover” script, or, how I would rather frame it “health” script and “sexualised” script, fails to note that “in addition to the script opposition, the joke has created a new blended space, a hypothetical world in which pretty young wives are ready and eager to have sex with any male who asks (...). This particular blend is one that contains the joke message that women, particularly young pretty ones, are no more than sex objects, a message sometimes interpreted as bona fide by male college students”. (2014, 28)

In reference to Attardo’s and Raskin’s analysis of 7 variations of a light bulb joke *“How many Poles does it take (...)”* (1991, 295), below I am listing 7 light bulb jokes that play with conceptions of feminism. (The first 2 jokes have been told by Sally Haslanger and Stephen Yablo in an online discussion following Yablo’s talk on “Deep Jokes”, LLC Lecture University of Turin, 10/12/2020; the others are circulating on the internet, e.g. on Reddit.) I do not necessarily consider these jokes to be subversive. This list is meant to be a joke itself in its form and reference. Perhaps its striking revelation and accumulation of feminist commonplaces could have a political effect after all.

(1) *“How many feminists does it take to screw in a light bulb?” “It’s not the light bulb that needs changing.”*

(2) *“How many feminists does it take to screw in a light bulb?” “That’s not funny!”*

(3) *“How many feminists does it take to screw in a light bulb?” “You don’t need a light bulb if you have a glass ceiling.”*

(4) *“How many feminists does it take to screw in a light bulb?” “It’s not the light bulb that needs screwing.”*

(5) *“How many feminists does it take to screw in a light bulb?” “One, to hold the bulb to the socket and wait for the world to revolve around her.”*

(6) “How many feminists does it take to screw in a light bulb?” “Trick question, feminists don’t change anything.”

(7) “How many feminists does it take to screw in a light bulb?” “12. One to screw it in, one to excoriate men for creating the need for illumination, one to blame men for inventing such a faulty means of illumination, one to suggest the whole “screwing” bit to be too “rape-like”, one to deconstruct the lightbulb itself as being phallic, one to blame men for not changing the bulb, one to blame men for trying to change the bulb instead of letting a woman do it, one to blame men for creating a society that discourages women from changing light bulbs, one to blame men for creating a society where women change too many light bulbs, one to advocate that lightbulb changers should have wage parity with electricians, one to alert the media that women are now “out-light bulbing” men, and one to just sit there taking pictures for her blog for photo-evidence that men are unnecessary.

References

Attardo, Salvatore (1990) “The Violation of Grice’s Maxims in Jokes”, in: David J. Costa (ed.) *Proceedings of the Sixteenth Annual Meeting of the Berkeley Linguistics Society*, Berkeley, CA: Berkeley Linguistics Society, 355–362.

Attardo, Salvatore; Raskin, Victor (1991) “Script theory revis(it)ed: joke similarity and joke representation model”, *HUMOR*, 4(3-4), 293–348.

Bergmann, Merrie (1986) “How Many Feminists Does It Take to Make A Joke? Sexist Humor and What’s Wrong with It” *Hypatia* 1(1), 63–82.

Bing, Janet; Scheibman, Joanne (2014) “Blended Spaces as Subversive Feminist Humor”, in: Delia Chiaro and Raffaella Baccolini (eds.) *Gender and Humor: Interdisciplinary and International Perspectives*, New York: Routledge, 13–29.

Coulson, Seana (2001) *Semantic Leaps. Frame-Shifting and Conceptual Blending in Meaning Construction*. Cambridge: Cambridge University Press.

Grice, H. P. (1975) “Logic and Conversation” in: Peter Cole and Jerry L. Morgan (eds.) *Syntax and Semantics* 3, New York: Academic Press, 41–58.

hooks, bell (1992) *Black Looks: Race and Representation*. Boston, MA: South End Press.

Hornsby, Jennifer (2000) “Feminism in philosophy of language: Communicative speech acts”, in: Miranda Fricker and Jennifer Hornsby (eds.) *The Cambridge Companion to Feminism in Philosophy* Cambridge: Cambridge University Press, 87–106.

Hurley, Matthew; Dennett, Daniel; Adams, Reginald (2011) *Inside Jokes. Using Humor to Reverse-Engineer the Mind*. Cambridge, London: MIT Press.

Kotthoff, Helga (2006) "Gender and humor: The state of the art" *Journal of Pragmatics* 38(1), Special Issue: Gender and Humor, 4–25.

Kramer, Chris (2020) "Subversive Humor as Art and the Art of Subversive Humor" *The Philosophy of Humor Yearbook* 1, 153–79.

MacKinnon, Catherine A. (1979) *Sexual Harassment of Working Women: A Case of Sex Discrimination*. New Haven, CT: Yale University Press.

Marvin, Amy (2022) "Feminist philosophy of humor", *Philosophy Compass* 17(7), e12858. <https://doi.org/10.1111/phc3.12858>

Marvin, Amy (2023) "Oppression, Subversive Humor, and Unstable Politics", *The Philosophy of Humor Yearbook* 4(1), 163–186.

Morreall, John (1987) "Funny Ha-Ha, Funny Strange, and Other Reactions to Incongruity", in: John Morreall (ed.) *The Philosophy of Laughter and Humor*, Albany: State University of New York Press, 188–207.

Raskin, Victor (1985) *Semantic Mechanisms of Humor*. Dordrecht: Reidel.

Sickinger, Mira Magdalena (forthcoming) "The Joke-Coop. Pragmatic Issues of Jokes and Joking", *The Philosophy of Humor Yearbook* 6(1), 31–50.

Sickinger, Mira Magdalena (2023) "'Aha!/Haha! – That's a good one!' On the Correlation of Laughter and Understanding in Joke Reception", in: Daniel O' Shiel and Victoras Bachmetjevas (eds.) *Philosophy of Humour: New Perspectives*, Leiden: Brill (Value Inquiry Books Series 389) 80–91.

Wittgenstein, Ludwig (2009) *Philosophical Investigations / Philosophische Untersuchungen*. Hg. von P. M. S. Hacker und Joachim Schulte, übersetzt von G. E. M. Anscombe, P. M. S. Hacker und Joachim Schulte. New York: Wiley. [Deutsch mit englischer Übersetzung]

Yablo, Stephen (2019) "Deep Jokes" (Philosophy Club Lunch Talk MIT), https://www.academia.edu/42179228/Deep_Jokes.

Revealing World-Pictures: Cordova, Wittgenstein, Culture, and Alterity

Daniel Smith (Albuquerque, New Mexico, United States)

Abstract

This essay examines Viola Faye Cordova's use and interpretation of Ludwig Wittgenstein's conception of world-pictures. In doing so, it shows how Cordova employs this facet of Wittgenstein's thought to support a unique vision of the roles played by culture and alterity in philosophical inquiry. Both thinkers view philosophy as an activity meant to induce transformations of vision whereby the seeming necessity of certain philosophical problems is seen to be the result of fixation upon a philosophical picture (Bild). Both also recognize that culture influences the forms problems take for people. The bases of culture, argues Cordova, are what Wittgenstein calls "world-pictures" (Weltbilder), inherited backgrounds which shape our lives and inquiry. Philosophy must, then, reveal and examine world-pictures if it wishes to change how we see certain problems. Cordova suggests that encountering alterity at the level of world-pictures reveals the limits of our own world-pictures. Such encounters, then, are a way of revealing our world-pictures to ourselves. She argues that philosophers should both use Wittgenstein's methods of imagination, whereby we can envision people who do not share our world-picture as well as conditions in which world-pictures other than our own could conceivably arise, and seek out actual encounters with people from cultures which do not share our world-picture. In these ways, aspects of our world-pictures are revealed to us, and our vision, actions, and lives can be transformed.

In his later years, Ludwig Wittgenstein came to worry about the impact "culture" has on an individual's vision of the world (Monk 1990: 533). Wittgenstein's philosophical methods aim to "change [our] way of looking at things" (PI 2009: §144), but, Wittgenstein recognized, culture conditions how we look at things in a manner which is outside of philosophy's control. This recognition, which Stanley Cavell calls Wittgenstein's recognition of our "fatedness to a culture" (Cavell 1979: 112), led Wittgenstein to an attitude of "pessimism about the effectiveness of his work" (Monk 1990: 533)

For the Jicarilla Apache philosopher Viola Faye Cordova, Wittgenstein's reflections on culture come to a head in his descriptions of "world-pictures" (Weltbilder) found in *On Certainty* (hereafter "OC"). In 1992, Cordova became the first Indigenous American in the United States to earn a PhD in Philosophy. Much of her work, which aims to establish Native American philosophy as a tradition in its own right, is motivated by an interpretation and deployment of Wittgenstein's remarks on world-pictures. My purpose in this essay is to show how Cordova's use of the concept of world-pictures

(alongside other methods and reminders taken from Wittgenstein's later works) supports a unique vision of the roles played by culture and alterity in philosophical inquiry.

To begin, let's mark Cordova's and Wittgenstein's shared conception of philosophical inquiry. For both thinkers, philosophy is not a "body of doctrine" or theory, but an "activity" (TLP 2002: §4.112; Cordova 2004: 33). The aim of philosophical activity is represented by Wittgenstein's image of showing "the fly the way out of the fly bottle" (PI 2009: §309; quoted in Cordova n.d(d): 1). This entails showing the seeming necessity of a philosophical problem to result from fixating upon a particular philosophical picture (Bild)—e.g., one in which words' meanings are seen as analogous to objects' essences, a picture which renders necessary the problem of identifying how a word retains the same meaning across various uses (PI 2009: §116)—to the exclusion of all others. In this state of fixation, the boundaries of the picture do not appear. Just as the fly perceives the clear borders of the bottle as boundless space, running up against them in an attempt to escape, the person fixated upon a philosophical picture fails to see the picture as a picture, remaining trapped within it. Wittgenstein's philosophical methods aim to induce a shift of vision (or "aspect" (see RPP vol I 1980: §415, §1029; RPP, vol II 1980: §39, §436)) whereby the picture is seen as a picture, and thus as one among many, thereby stripping the problem of the intractable character it once possessed.

Cordova notes that, to show the fly the way out of the fly bottle, we must first undertake the more difficult task of showing there to be a fly bottle in the first place (Cordova 2007: 57). For her, the bottle is not just a particular philosophical picture within which a certain problem arises, but an entire world-picture. To make this point, she draws from OC, where Wittgenstein defines the world-picture as "the inherited background against which I distinguish true and false" (OC 1969: §94; quoted in Cordova 2007: 61). This background "gives our way of looking at things, and our researches, their form" (OC 1969: §211; quoted in Cordova n.d(a): 8). If world-pictures provide the form for our way of looking at things, and philosophy is an activity meant to change how we look at things, then philosophy must reveal and examine world-pictures. This is no simple task—hence Wittgenstein's pessimism. To understand the difficulty of this endeavor, we should first look closer at Cordova's and Wittgenstein's explanations of how world-pictures are inherited

and how they serve as a background for inquiry and life. In this way, we will arrive at a better understanding of what world-pictures are and how they shape our way of looking at things.

Cordova argues that world-pictures form as a people's response to their historical, environmental, and social circumstances (Cordova n.d(b): 12). They provide answers to basic questions concerning the human, the world, and the role of the human in the world (Cordova n.d(b): 12). These answers are codified in a "tradition" (e.g., the "Western tradition" (Cordova 1987: 13)) and instantiated in the "beliefs," "attitudes" (Cordova 2007: 57), "customs" (Cordova 1987: 7), and "practices" (Cordova 2007: 67) of a people. Thus, we inherit a world-picture by being born into and brought up in a particular human community. Therefore, world-pictures serve, says Cordova, as a "source of cultural distinctions" (Cordova 1987: 38). At the same time, world-pictures can encompass many cultures. For example, Cordova feels justified speaking for the "Native American" world-picture, which encompasses over 500 distinct Indigenous nations, because this world-picture formed in these various communities around the same time in response to similar circumstances (Cordova 2007: 3, 102). Since we inherit them, the features of our world-picture are not taught to us as explicit "rules" (OC:1969: §95; quoted in Cordova 2007: 61), but rather are "picked up" (Cordova n.d(c): 2) from our environment, community, parents, and so on. Our world-picture is something which we absorb and into which we ourselves are absorbed. As Wittgenstein puts it: "Tradition is not something one can learn; not a thread one can pick up when one feels like it; any more than one can choose one's own ancestors" (CV 1980: 76 (translation slightly modified)).

It is because our world-picture is inherited, because it is something by and in which our lives are formed, that it serves as the background of living and inquiring while itself standing "apart from the route traveled by enquiry" (OC 1969: §88; quoted in Cordova, n.d(a): 7). As the "scaffolding" of thought and action (OC 1969: §211, quoted in Cordova n.d(b): 12), that which allows us to make judgements about what is true or false, it would make no sense to say of our world-picture that it is true or false (OC 1969: §205)—just as it makes no sense to say that the meter stick in Paris is or is not one meter long (PI 2009: §50). We do not make measurements of the meter stick in Paris: it is that which allows us to make measurements, the background of all measurements.

Likewise, we do not make judgments of our world-picture: it is that which allows us to make judgments, the background of all judgements.

Crucially, this means that world-pictures exert an influence on what problems arise for a people: if world-pictures give us the form in which we look at and inquire into things, they condition what it is we see and what it is we inquire about. Cordova provides the following example: the concept of perpetual “progress,” and the concomitant problem of how to secure it by subduing nature to serve human ends, are central to the “Euro-American” world-picture in both its religious and secular manifestations (Cordova n.d(b): 5, 10). (For what it is worth, Wittgenstein wrote something strikingly similar in 1930 when he laments that “progress” has become the “form” of Western “civilization” (CV 1980: 7e), which suggests that progress, as a Western civilization’s form, is a concept which belongs to the Western world-picture.) Cordova demonstrates that this problem “could never have developed” in the Native American world-picture (Cordova 2007: 119). For in this world-picture, the concept of perpetual progress and the problem of how to secure it are replaced by the problem of how to maintain “balance” (Cordova 2007: 120) and “harmony” (Cordova 2007: 123) with the natural world, a world in which humanity is but one node in a system of equally important, interrelated nodes (Cordova 2007: 211). From this difference in world-pictures all sorts of different philosophical problems follow: for instance, the Euro-American tradition of philosophy is preoccupied with problems of “time” in relation to Humanity’s (capital “H”) projects (and progress) while the Native American philosophical tradition is more preoccupied with problems of “place” and the responsibility specific human communities have regarding their relations with those beings present in their particular natural surroundings (Cordova 1987: 35; see Deloria 2003: 61).

To recognize the influence world-pictures have on the forms which philosophical problems take for people from different cultures is to take Wittgenstein’s worry about the influence of culture on philosophical inquiry to its root. We can now see that the difficulty in revealing and examining world-pictures comes from the fact that they are something we inherit without initially realizing it and shape our inquiry without us usually being cognizant of it. As our inherited background, our world-picture does not usually show up to us. How, then, are world-pictures to be revealed and made available for philosophical inquiry?

Cordova offers the insightful reminder that world-pictures are revealed “when two people from different cultures” (Cordova 2007: 67), particularly from cultures “which share no histories, no experiences, previous to their initial encounters” (Cordova n.d(a): 2) and thus do not share the conditions which lead to the formation of a shared world-picture, genuinely encounter one another in their alterity. Such people will “find it difficult to communicate—their frames of reference do not meet. Again,” she continues, “Wittgenstein has an enlightening comment on this situation” (Cordova 2007: 62); here she quotes *Philosophy of Psychology-A Fragment* (hereafter “PPF”) §325, where Wittgenstein points to experiences in which we cannot “find our feet” with (or “find ourselves in” [Wir können uns nicht in sie finden]) people in or from a “strange [fremdes, alien] country,” even if we know their language (PPF 2009: §325; quoted in Cordova 2007: 62). The hesitation, the breakdown of communication, which can occur when encountering someone from a culture which does not share a world-picture with one’s own serves as an invitation to see the differences between the foundations of each culture, which means seeing, “perhaps for the first time,” as Anna Boncompagni says, “an axis of rotation of one’s life” (Boncompagni 2016: 208). These sorts of encounters are frequently imagined in Wittgenstein’s work. Think, for example, of the people imagined in the *Remarks on the Foundations of Mathematics* who measure the price of wood by how much ground it covers (RFM 1964: §149; quoted in Cavell 1979: 115), or of the imaginary tribe in OC who confronts G.E. Moore with the belief that “he has come from somewhere between the earth and the moon” (OC 1969: §264). If in these cases we try to teach these people our way of pricing wood or tell the tribe that we know Moore has never left the surface of the earth, “[w]e should be trying to give [them] our world-picture” (OC 1969: §262). If we are able to see that the giving of a world-picture is what would be required to persuade these other people to see things our way, then these imagined encounters with alterity too serve the purpose of revealing some facet of our world-picture to us. (Cavell underscores that this alien other often is, in Wittgenstein’s examples, a child from one’s own culture (Cavell 1979: 125); Cordova too calls the child “an unknown factor,” an alien presence who must, through acculturation, be brought into a world-picture (Cordova 2004a: 177).)

Through encounters with alterity (real or imagined) which exist at the level of word-pictures, the fly bottle is revealed to the fly. The first thing Cordova thinks we see when we catch a glimpse of our world-picture is that it constitutes a “bounded space,” that our cultural bottle has boundaries (Cordova 2007: 195). Consider Cordova’s commentary on PPF §325: “There is a context to language. [This context] consists of beliefs as well as history and physical circumstance [. . .] I am not speaking here of [the contexts in terms of] untranslatability nor even of incommensurability. The alien context would make no sense” (Cordova n.d(c): 2). When Cordova says that the alien context would make no sense, she is highlighting what occurs in the moments when we realize that there are other bounded, coherent systems of sense or “webs of significance”—as we might say, following Clifford Geertz (Geertz 1973: 5)—which may initially appear to us as nonsense. The boundedness of our own system of sense is thereby exposed—there is a place (many places) beyond which it cannot reach. The recognition of the boundedness of our world-picture is an invitation to see certain philosophical pictures—and the problems to which they give rise—as bounded by our world-picture: again, think here of the Euro-American insistence on progress and the resultant philosophical problems of time as opposed to the Native American preoccupation with balance and the resultant philosophical problems of place. The picture bounds the problem and the picture is in turn bounded by a world-picture.

Thus, if philosophy aims to change the way we see certain problems, then it must not only examine the conceptual and practical contexts within which certain problems arise—the pictures—but the cultural contexts as well—the world-pictures. For Cordova, this entails “the examination and analysis of belief systems offered up as ‘answers’ to the dilemmas that human beings face” (Cordova 2007: 50). We can see in this vision of philosophy’s responsibility to and for culture the presence of Wittgenstein’s admonishment that philosophical inquiry ought to be turned around on the axis of our “real need” (PI 2009: §108). World-pictures, says Cordova, must “have some grounding in a people’s factual circumstances in order to acquire validity or credibility” (Cordova 2007: 101)—so, to examine a world-picture is to examine those problems which arise out of a people’s real needs.

Lest this sensitivity to differences in world-pictures be taken as “relativism” and an abuse of Wittgenstein’s work (Cordova nd(c): 2), Cordova reiterates that

encounters with alterity are not fodder for theorization. Rather, they are an occasion to acknowledge that, as Wittgenstein says, the true “difficulty is to realize the groundlessness of our believing” (OC 1969: §166; quoted in Cordova n.d(a): 2). It is difficult to simultaneously realize that our world-picture grounds certain problems and that our world-picture is no ground at all. Rather, it is a particular background with particular boundaries, what Wittgenstein calls a *Lebenselement* (OC 1969: §105) within which problems have their lives and without which they do not live (see Boncompagni 2016: 181). This difficulty is, according to Cordova, the difficulty of philosophy, the struggle and work of philosophy. It is the difficulty concurrent with encountering the foundations of the alterity of the Other. It is a difficulty of vision, of what we are able to see or able to let ourselves see, and a difficulty of action, of what we are prepared to do and let happen to us, as much as it is a difficulty of thought (Cordova n.d(a): 2).

Cordova observes that philosophers are adept at subverting this difficulty. Oftentimes philosophers “come to the alien perspective armed with [their] own concepts and attempt to find something in the other culture that matches [their] concept” (Cordova 2007: 54). If they cannot find this match, the alien perspective is dismissed as “primitive” or “simply wrong” (Cordova n.d(a): 12). This tendency to valorize “similarities” and devalue differences most often serves to “establish some vast system of conformance by which those unlike ourselves can be ranked in some hierarchical order of ‘proper’ degrees of ‘rightness’ and ‘correctness’” (Cordova n.d.: 4). Cordova encountered this attitude as a PhD student in resistance from faculty toward her work comparing the concept of monism in Navajo philosophy with that found in Spinoza. A faculty member even told her that her desire to make such a comparison showed that she had “mistaken Spinoza for a primitive” (Cordova 1992: 110). Since the philosopher must reveal and examine world-pictures, and since it is through encountering genuine alterity that aspects of world-pictures are revealed, the philosopher must learn “to concentrate, not on similarities, but on differences. It is by contrasting notions that one learns about the distinction between the self and the Other. Those differences are what make the Other an Other” (Cordova 2007: 60). We have seen that what this encounter shows us—the boundedness of our world-picture—is difficult to accept. Cordova suggests here that overcoming this difficulty requires the cultivation

of a mode of attention which allows for the revelation and examination of alterity as alterity. This entails training ourselves to resist the tendency to search among differences for some underlying similarity, some fundamental “commonality of concepts” (Cordova 2007: 59).

Once again, Wittgenstein is the philosopher to whom Cordova turns to “teach [us] differences,” (the quotation from *King Lear* which Wittgenstein considered as a motto for his *Philosophical Investigations* (Monk 1990: 537)). She is particularly fond of quoting PPF §365-366, where Wittgenstein invites us to “imagine very general facts of nature to be different from what we are used to” so that we can see that our own “concepts” are not “absolutely correct” or the only “intelligible ones” (quoted in Cordova n.d(d): 7). Since she thinks world-pictures form in part as responses to and explanations of the general facts of nature faced by a particular people in a particular time and place, imagining such facts to be different than what we are used to allows us to conceive of the conditions in which a world-picture different from our own might form. She sees her work on Navajo philosophy as an application of this method: “The Navajo view of the universe is an example of someone ‘imagining’ certain ‘very general facts of nature to be different’ from what [Euro-Americans] are used to” (Cordova 1987: 7). In philosophical activity involving the revelation of world-pictures, “[d]ifferences will be presented as examples of the diversity of human thought patterns” (Cordova 2007: 3), and presented in such a manner that the distance of difference is not closed, but maintained as a source of personal transformation. This is a transformation of vision, but also of action, and even of how we live. As Lars Leeten puts it, for Wittgenstein, “our whole life can be converted if our vision of the world is transformed” (Leeten 2023: 193). And as Wittgenstein himself says: “What is called an alteration in concepts is of course not merely an alteration in what one says, but also in what one does” (RPP vol I 1980: §910).

Through Cordova’s work, we can more easily see how to follow Wittgenstein in imagining a people with practices, beliefs, and languages different from our own is to imagine a context within which those practices, beliefs, and languages have their life. As Wittgenstein puts it, “to imagine a language means to imagine a form of life” (PI 2009: §19). Even more to the point: “What belongs to a language game is a whole culture” (LC 1967: 8). What Cordova calls for—and what she undertakes in her own work—is the concrete

application and extension of Wittgenstein's methods of imagination in the form of "comparative philosophy" (Cordova 2007: 67), a mutli-cultural philosophy which can teach us differences.

References

Boncompagni, Anna (2016) *Wittgenstein and Pragmatism: On Certainty in Light of Peirce and James*, New York: Palgrave Macmillan.

Cavell, Stanley (1979) *The Claim of Reason: Wittgenstein, Skepticism, Morality, and Tragedy*, Oxford: Oxford University Press.

Cordova, Viola Faye (n.d) "A Matter of Language", Albuquerque: Center for Southwest Research, Box 2, Folder 4.

Cordova, Viola Faye (2004) "Approaches to Native American Philosophy", in: Waters, Anne (ed.) *American Indian Thought: Philosophical Essays*, New York: Blackwell, 27-33.

Cordova, Viola Faye (1987) "Conceptual Frameworks as a Source of Cultural Distinctions", Albuquerque: Center for Southwest Research, Box 2, Folder 1.

Cordova, Viola Faye (2004a) "Ethics: The We and the I", in: Waters, Anne (ed.) *American Indian Thought: Philosophical Essays*, New York: Blackwell, 173-181.

Cordova, Viola Faye (2007) *How It Is: The Native American Philosophy of V.F. Cordova*, Jojola, Ted; Lacy, Amber; Peters, Kurt; Moore, Kathleen Dean (eds.) Tucson: University of Arizona Press.

Cordova, Viola Faye (1992) *The Concept of Monism in Navajo Thought*, Albuquerque: Center for Southwest Research, Box 1, Folder 1.

Cordova, Viola Faye (n.d(a)) "The Structure of Conviction", Albuquerque: Center for Southwest Research, Box 2, Folder 4.

Cordova, Viola Faye (n.d(b)) "Untitled Essay", Albuquerque: Center for Southwest Research, Box 2, Folder 4.

Cordova, Viola Faye (n.d(c)) "Untitled Manuscript", Albuquerque: Center for Southwest Research, Box 2, Folder 2.

Cordova, Viola Faye (n.d(d)) "Wittgenstein and a Conceptual Framework", Albuquerque: Center For Southwest Research, Box 2, Folder 4.

Deloria, Vine Jr (2003) *God is Red: A Native View of Religion*, 30th Anniversary Edition, Golden: Fulcrum.

Geertz, Clifford (1973) *The Interpretation of Cultures*, New York: Basic Books.

Leeten, Lars (2023) "Wittgenstein and the Art of Conversing with Oneself: The Philosophical Investigations as a Book of Exercise", in: Guidi, Lucilla (ed.) *Wittgensteinian Exercises: Aesthetic and Ethical Transformations*, Munich: Wilhelm Fink Verlag, 187-206.

Monk, Ray (1990) *Ludwig Wittgenstein: The Duty of Genius*, New York: Penguin Books.

Wittgenstein, Ludwig (1980) *Culture and Value*, Chicago: The University of Chicago Press.

Wittgenstein, Ludwig (1967) *Lectures and Conversations on Aesthetics, Psychology, and Religious Belief*, Barret, Cyril (ed.), Berkeley: University of California Press.

Wittgenstein, Ludwig (1969) *On Certainty*, Anscombe, G.E.M; von Wright, G.H. (eds.), translated by Denis Paul and G.E.M Anscombe, New York: Harper and Row.

Wittgenstein, Ludwig (2009) *Philosophical Investigations*, translated by G.E.M Anscombe, P.M.S. Hacker, and Joachim Schulte, Oxford: Blackwell.

Wittgenstein, Ludwig (2009a) *Philosophy of Psychology-A Fragment*, translated by G.E.M Anscombe, P.M.S. Hacker, and Joachim Schulte, Oxford: Blackwell.

Wittgenstein, Ludwig (1964) *Remarks on the Foundations of Mathematics*, Anscombe, G.E.M; Rhees, Rush; von Wright G.H. (eds.), translated by G.E.M Anscombe, Oxford: Basil Blackwell.

Wittgenstein Ludwig (1980) *Remarks on the Philosophy of Psychology, Volume I*, Anscombe, G.E.M; von Wright, G.H. (eds.), translated by G.E.M Anscombe, Chicago: The University of Chicago Press.

Wittgenstein, Ludwig (1980a) *Remarks on the Philosophy of Psychology, Volume II*, Nymann, Heikki; von Wright, G.H. (eds.), translated by C. G. LuckHardt and M. A. E. AUE, Chicago: The University of Chicago Press.

Wittgenstein, Ludwig (2001) *Tractatus Logico-Philosophicus*, translated by D. F. Pears and B. F. McGuinness, New York: Routledge Classics.

Acknowledgement

I would like to thank Professors Sarale Ben-Asher and Paul Livingston for extensive feedback on earlier drafts of this essay. I would also like to thank my colleagues Rui Teng Phoebe Mak and Kedar Petwary for both feedback on earlier drafts of this essay and countless enlightening conversations. Finally, I would like to thank Naomi Jacobi for introducing me to the work of Viola Faye Cordova.

On Irigaray's and Wittgenstein's Critiques of Formal Logic

Kim Solin (Uppsala, Sweden)

Abstract

Luce Irigaray and Ludwig Wittgenstein are both critics of formal logic. Irigaray speaks of the desire for uniformity, while Wittgenstein refers to the craving for generality—tendencies that hinder our ability to attend to the nuances of life and language. In Irigaray's case, her critique is intimately connected to the task of revealing and articulating the Feminine. For Wittgenstein, it is closely related to his philosophy of mathematics. Both thinkers direct their critiques not only at the specific details of formal logic but also at its overarching use and influence. In this paper, I will use Marjorie Hass's essay 'Fluid Thinking: Irigaray's Critique of Formal Logic' (Hass 2002) as a starting point to explore Irigaray's objections to formal logic alongside Wittgenstein's. Methodologically, the paper is an interwoven commentary on Hass's paper, informed by Rush Rhees's understanding of Wittgenstein's philosophy of mathematics and directly by the Remarks on the Foundations of Mathematics.

Luce Irigaray and Ludwig Wittgenstein are both critics of formal logic. Irigaray speaks of the desire for uniformity, while Wittgenstein refers to the craving for generality—tendencies that hinder our ability to attend to the nuances of life and language. In Irigaray's case, her critique is intimately connected to the task of revealing and articulating the Feminine. For Wittgenstein, it is closely related to his philosophy of mathematics. Both thinkers direct their critiques not only at the specific details of formal logic but also at its overarching use and influence. Despite this, their critiques of formal logic have received significantly less attention compared to other aspects of their thinking. Yet, one could argue that their critiques are central to their thought, respectively.

Irigaray and Wittgenstein have numerous interpreters, with varying perspectives on their work. However, I propose that a comparative analysis of their critiques of formal logic can deepen our understanding of both thinkers. In this paper, I will use Marjorie Hass's essay 'Fluid Thinking: Irigaray's Critique of Formal Logic' (2002) as a starting point to explore Irigaray's objections to formal logic alongside Wittgenstein's. My own understanding of Wittgenstein's philosophy of mathematics is primarily based on Rush Rhees's interpretation. Methodologically, the paper is an interwoven commentary on Hass's paper, informed by Rhees's understanding of Wittgenstein's philosophy of mathematics and directly by RFM (by RFM, I shall henceforth refer to the 1978 edition).

*

Marjorie Hass notes the following: “Irigaray’s critique of logic has received less attention than the challenges she makes to law and language. Yet it is in many ways her most important” (Hass 2002: 71). Something similar could be said about Wittgenstein’s thinking about the philosophy of mathematics and its relation to mathematical logic: although it was central to Wittgenstein, it has received much less attention than other parts of his philosophy, as the editors of RFM were well aware of (Solin 2023). Perhaps the root causes are the same for both the reception of Irigaray and of Wittgenstein, something similar to what Hass writes: “Logic is the symbolic structure that we hold up as the most neutral of all, as essentially definitionally neutral. It serves as an exemplar of neutrality itself. [To] unmask logic is to call the very possibility of neutrality into question. But for this very reason, the case against logic is also the hardest to make plausible” (Hass 2002:71). Hass notes that “Irigaray’s approach to logic is the most theoretical and associative of her treatments of formal systems, requiring her to use her full range of philosophical techniques.” (Hass 2002: 71-72). I think that something similar could be said about Wittgenstein’s approach to the philosophy of mathematics, and his critique of mathematical logic. He attacks our thinking where it seems almost impossible and uses a whole battery of different techniques.

*

Many have been perplexed by what Wittgenstein is actually up to in the RFM. Rush Rhees writes the following in a letter to Georg Henrik von Wright (kept at the von Wright and Wittgenstein Archives in Helsinki):

Wittgenstein was not pretending to do mathematics. He was describing [something] which could be mathematics. (In this way it is similar to his construction eines Sprachspiels or of the Lebensweise eines Stammes, whose grammar differed in important ways from ours, for whom 'intelligibility' and 'understanding' were different from the intelligibility and understanding we seek; and so on. And this could illuminate, or separate certain features of our speech and questions and reasonings.) Wittgenstein is not suggesting that any mathematical problem would be solved through the development of [these] ways of calculating.

What he does suggest, I think, is that if one pays attention to the possibility of a variable so conceived [...] then one may be less prone to 'mythological' or 'fantastic' interpretations of the forms of calculation mathematicians do use. (Rhees to von Wright, 4 April 1973)

So Wittgenstein is not doing mathematics himself, but only trying to make us think about it differently. To make us see what we are actually doing more clearly. It seems to me that it is a similar thing that Irigaray is doing, as described by Hass:

She acts as both prophet and muse urging us to a land just slightly out of reach, just beyond the next horizon. And it is in just this mode that Irigaray returns again and again to the logic of negation, identity, and generality to illustrate the monosexuality of our theoretical paradigms, to gesture toward new paradigms, and to reconstruct relationships. (Hass 2002: 72)

This gesturing towards new paradigms seems to me akin to Wittgenstein's use approach as understood by Rush Rhees. Wittgenstein wants us, according to Rhees, to find a new way of doing philosophy of mathematics and a new way of viewing what we do as mathematicians (Solin 2023).

For sure, much of what Irigaray says about formal logic might also strike one as incomprehensible if one has started out as a logician. As Hass notes, "[a]pproaching formal logic through this lens is unusual and, from the perspective of traditional philosophy of logic, perhaps impossible" (Hass 2002: 73). But I believe that reading Irigaray and Wittgenstein in tandem might make it easier to understand both philosophers. Their immediate aims are different, in that Irigaray's main goal is to articulate sexuate difference, but their methods and their striving to clarity about differences, as it were, seem related to me. Irigaray, just like Wittgenstein, insists on "the fundamentality of difference, particularly sexual difference, and [...] that genuine difference is ignored even when we think we are focusing on it" (Hass 2002: 72).

*

Here's Hass again:

In Irigaray's work, a variety of supposedly neutral structures, then, are revealed as expressions of a cultural desire for sameness. A desire that must be revealed and interpreted before difference can be thought. (73)

Recall that Wittgenstein in the Blue Book pointed out that our "craving for generality" is related to many philosophical confusions and to a "contempt for the particular case" (BBB: 17; see also Stenlund 2002).

And as one can see from RFM, Irigaray and Wittgenstein share the following:

When Irigaray confronts logical theory, then, she is looking for at least two things: the elements of logical reality that exceed formal logic's representational capacity and the desires revealed by the marginalization of these elements. (Hass 2002: 74)

Wittgenstein, too, wants to show that mathematical reality exceeds mathematical logic, as seen in for example the following remark:

"The disastrous invasion" of mathematics by logic. [...] The harmful thing about logical technique is that it makes us forget the special mathematical technique. (RFM: 281, § 24)

And Wittgenstein wants to understand why we are driven by these desires, or at least to make it clear to us that we are (Rhees 2006: 263; Solin 2013: 33-34).

Wittgenstein and Irigaray are similar in this dual critique of formal logic: they simultaneously consider the technical details *and* the existential leanings driving us towards formal logic. Especially that latter, the existential aspects, are too easy to forget when it comes to Wittgenstein on mathematics and formal systems (Floyd & Mühlhölzer 2020; Solin 2021).

*

So consider then the following by Hass:

While logicians focused their attention exclusively on the formal properties of logical concepts, Irigaray's psychoanalytic perspective *includes* (my emphasis) the Imaginary meaning of logical connectives. In other words, while she is interested in the literal meaning of, for example, the negation symbol, she is also attentive to the associative, symbolic function of contradiction and, at a

deeper level, the way that negation functions in the cultural Imaginary. (Hass 2002: 74)

And related to this:

The conception of limit that Irigaray formulates cannot be directly articulated in the language of formal logic. ... in one sense then, the relationship that Irigaray describes is *nonlogical*. But this is not to say that it is unreasonable or irrational. At the center of Irigaray's work it's the task of drawing our attention to an as yet unsymbolized possibility. She uses a variety of techniques for expanding our thinking, for drawing us just enough beyond our current representations that we can begin to create a new representational possibility. ... As I suggested earlier Irigaray is interested in the symbolic associations binding culturally significant concepts. In the case of negation, for example, the logical rules that given the negation operator, carry with them a force that goes beyond the framework of formal logic per se. Logic and its operations come to represent rationality, meaning, and sense of themselves. (Hass 2002: 78)

It is a similar thing, I think, that Wittgenstein does in his remarks on negation (RFM: 102 ff.) Not only is he making remarks to the point of nuancing our understanding of negation, creating new representations ("ne"), but just like Irigaray he includes making us aware of how negation functions in what above was described as the cultural Imaginary. For example, in a striking parentheses he later talks about "[t]he superstitious dread and veneration by mathematicians in face of contradiction" (RFM: 122). Indeed, already in his war diaries Wittgenstein was concerned with unifying, on the one hand, his existential inquiries, with, on the other hand, his more technical, logical work (see Jareño Alarcón 2023).

*

Let us now consider, then, the law of identity, that is that for every x , x is equal to itself, or in symbols $(\forall x)(x=x)$. Hass argues that the law is enabled by differences specified within the law itself (Hass 79 ff.). To represent an object's identity, two instances (or tokens) are needed. Identity is never complete; there is always a leftover difference, a residue. Irigaray's views on identity have garnered significant attention, particularly her claim that the law of

identity does not apply to women. While others, like Heidegger and Deleuze, acknowledge the persistent difference beyond the law of identity, Irigaray's analysis is focused on sexual difference. She highlights the link between this residual difference and femininity. In the context of the law of identity, where diversity is both shown and overlooked, femininity represents what cannot be fully defined. Irigaray's critical readings of Western rationality reveal that femininity symbolizes excess, a concept also evident in everyday portrayals of femininity.

One can compare this to Wittgenstein when he writes the following about the law of identity:

Frege calls it 'a law about what men take for true' that 'It is impossible for human beings ... to recognize an object as different from itself'. – When I think of this as impossible for me, then I think of *trying* to do it. So I look at my lamp and say: "This lamp is different from itself". (But nothing stirs.) It is not that I see it is as false, I can't do anything with it at all. (Except when the lamp shimmers in sunlight; then I can quite well use the sentence to express that). One can even get oneself into a thinking-cramp, in which one *does* someone trying to think the impossible and not succeeding. Just as one can also *do* someone trying (vainly) to draw an object to himself from a distance by mere willing (in doing this one makes e.g. certain faces, as if one were trying, by one's expression, to give the thing to understand that it should come here.) (RFM: 89, § 132)

Wittgenstein's remark can be related to Irigaray in two ways. On the one hand, one could say that Irigaray is trying to think the impossible. On the other hand, one could say that Irigaray is talking about identity in a way that Wittgenstein says makes sense when talking about the shimmering light. I think it is the latter that would be the correct reading of Irigaray: when she says the law of identity does not apply to women, this should be understood as her wanting to make us see something like the lamp in sunlight, she wants us to pay attention to aspects of what it is to be a woman that have not been articulated. And, perhaps, it has not been articulated partly because we have been too concerned with the law of identity in the typical formal sense.

But the excess and the residue here still seems to me to be a contentious issue: why is it associated with femininity? On this point, I suspect Wittgenstein and Irigaray would differ, although they both recognize that there *is* a residue, and that this residue is important, neglected and yet meaning-giving.

Wittgenstein remarks that:

The propositions of logic are 'laws of thought', 'because they bring out the essence of human thinking'— to put it more correctly: because they bring out, or shew, the essence, the technique, of thinking. They shew what thinking is and also shew kinds of thinking. (RFM: 90, §§ 133-134)

Consider in relation to this the following from Hass:

Margaret Whitford writes, "the practical value of these principles, without which rationality would be inconceivable, is so evident that it appears unquestionable. The logic of identity is the prerequisite of any language or society at all. However, the point is that there will always be a residue which exceeds the categories, and this excess is conceptualised as female". [...] Irigaray has identified a sense, then, in which the law of identity does not "apply" to women. Femininity represents the excess that surrounds the very possibility of self-substituting identity. [...] Without their own specific form of identity, women cannot function as a limit to the male sex, they cannot resist the totalizing impulse of the logical structure of negation and identity. (Hass 2002: 80-82)

And perhaps one could argue, even from Wittgenstein's perspective, that Irigaray tries to show us a different "kind of thinking". And that the thinking which falls under the feminine has been concealed, also in the case of logic.

*

Let us return to Hass:

Irigaray's work calls the neutrality of logic into question, arguing that the standard formalism is capable of expressing only distorted and partial interpretations of negation identity and generality. ... Thinking of difference,

identity, and generality in the new ways that Irigaray asks of us requires that we expand on the representational possibilities offered to us by the standard formalism. (Hass 2002: 84)

This can be understood so that those interested in standard formal structures can ignore Irigaray's critique as long as they don't apply it to sexuate difference. Or that they can try to conceive of new formal structures, related to standard ones, for expressing sexuate difference. But Irigaray's and Wittgenstein's critiques do not end there. Hass writes:

Irigaray, however, insists that there can be no new formalism, that the problems she has identified are problems *with formalism itself*. (Hass 2002: 85)

And this is completely in tune with Wittgenstein's critique of mathematical logic:

The curse of the invasion of mathematics by mathematical logic is that now any proposition can be represented in a mathematical symbolism, and this makes us feel obliged to understand it. Although of course this method of writing is nothing but the translation of vague ordinary prose. [...] Mathematical logic has completely deformed the thinking of mathematicians and of philosophers, by setting up a superficial interpretation of the forms of our everyday language as an analysis of the structures of facts. Of course in this it has only continued to build on the Aristotelian logic. (RFM: 299-300)

This is one reason why Wittgenstein in the PI (§ 124) can say that no discovery in mathematical logic can advance philosophy. And, could one not, perhaps, say that no discovery in formal logic really advances our understanding of sexuate difference?

*

Famously, Wittgenstein wrote that:

Work on philosophy – like work in architecture in many respects – is really more work on oneself. On one's own conception. On how one sees things. (And what one expects of them). (CV: 24)

And this is precisely what Irigaray asks of us when it comes to formal logic:

Irigaray asks the logician to interrogate his own desire for closed logical system. Given that language and experience overflow the boundaries of logic, what pleasures are found in the narrowing and restricting of attention that formal logic demands? Asking into this level of desire is fundamentally foreign to the logician who focuses on instrumental ends and deemphasises even the aesthetic pleasures associated with an elegant proof or a spectacular technical solution. The relationship between desire and logic has already been opened by Wittgenstein in his *Investigations*. One open question, then, is the extent to which Irigaray has provided a sexed basis for this relationship, showing that the deep pleasures, desires, and frustrations that give rise to our pursuit of formal logic are the same ones that structure our sense of masculinity: closure, unity, and solidity. (Hass 2002: 85)

Note that this is the only reference to Wittgenstein in Hass's text, but the PI is not in the references and Wittgenstein is not in the index of names of the volume. It seems out of place to me, and I'm not certain what parts of PI she is referring to; the request of a reviewer, perhaps. But one can certainly agree with her, and as I hope to have already shown, this relationship is even more visible in RFM. However that may be, just like Hass notices and as I said above, the contentious issue is how and whether the perspective opened up by Wittgenstein and Irigaray relates to the feminine and to the masculine.

*

Hass quotes Rosi Braidotti:

Irigaray defends the notion of 'difference' in a conditional mode. This means that woman does not yet exist and that she will be unable to come into being without woman's collective efforts. ... Sexual difference that's the difference that women make has to be constructed, and, for Irigaray, it is the task of the women's movement to set the conditions of possibility for this becoming. (Braidotti 1994: 250, as quoted by Hass 2002: 72)

And comments:

For Irigaray, sexual difference – indeed, difference itself – has yet to be thought at the deepest and fullest sense in Western culture; it has yet to enter into the symbolic constructions that constitute our culture. Moreover, thinking

sexual difference is not simply a matter of adding women into the existing constructions. The structures themselves resist these additions since they are in fact constituted to resist difference. (Hass 2002: 72-73)

In RFM, Wittgenstein makes a related remark:

The sickness of a time is cured by an alteration in the mode of life of human beings, and it was possible for the sickness of philosophical problems to get cured only through a changed mode of thought and of life, not through a medicine invented by an individual. [...] Think of the use of the motor-car producing or encouraging certain sicknesses, and mankind being plagued by such sickness until, from some cause or other, as the result of some development or other, it abandons the habit of driving. (RFM: 132)

Both emphasise that we need to change the way we live, for certain things to even become comprehensible. Irigaray's project can be described as utopian, as slowly working towards a non-totalising relation between the sexes that, acknowledging the irreducible two, would not allow one to dominate the other. This, in turn, would foster a humility crucial for a common humanity (Hass 2002: 82; Irigaray 1996: 106-107). (Does the Irigarayan perspective represent a "regressive biological essentialism" (Hass 2002: 81)? I think that would be a very simplistic reading of Irigaray, but that is the topic of another paper (see, however, Hass 2002: 81 ff.).)

*

In this paper I have shown, by close commentary, that there are similarities between Irigaray's and Wittgenstein's discussions of formal logic. I have focused on their general approach and emphasized that they both consider the technical aspects as well as the more existential aspects of formal logic. I think that developing this outline further would be rewarding for Irigaray studies and Wittgenstein's studies, respectively, and that one can compare Irigaray and Wittgenstein on negation, identity and generality in much more detail. I hope to have the opportunity to do so in future work.

References

Braidotti, R. (1994) *Nomadic Subjects: Embodiment and Sexual Difference in Contemporary Feminist Theory*, New York: Columbia University Press.

Hass, M. (2002) 'Fluid Thinking: Irigaray's Critique of Formal Logic', in Falmagne, R.J. and Hass, M. (eds.) *Representing Reason: Feminist Theory and Formal Logic*. Rowman & Littlefield.

Floyd, J. and Mühlhölzer, F. (2020) *Wittgenstein's Annotations to Hardy's Course of Pure Mathematics: An Investigation of Wittgenstein's Non-Extensionalist Understanding of the Real Numbers*. Berlin: Springer.

Irigaray, L. (1996) *I Love to You: Sketch for a Felicity within History*. Translated by A. Martin. New York: Routledge.

Jareño-Alarcón, J. (ed.) (2023) *Ludwig Wittgenstein: The Meaning of Life*. Hoboken, NJ: Wiley Blackwell.

Rhees, R. (2006) *Wittgenstein and the Possibility of Discourse*, 2nd Edition. London: Blackwell.

Solin, K. (2013) *The Mathematician as Mathematics: Theories of Computation in Light of Wittgenstein's Thought*. PhD thesis. Uppsala University.

Solin, K. (2021) 'Book Review: Juliet Floyd and Felix Mühlhölzer, Wittgenstein's Annotations to Hardy's Course of Pure Mathematics: An Investigation of Wittgenstein's Non-Extensionalist Understanding of the Real Numbers', *Nordic Wittgenstein Review*, 10:157-162.

Solin, K. (2023b) 'The Revision of Wittgenstein's Remarks on the Foundations of Mathematics', in Wallgren, T. (ed.) *The Creation of Wittgenstein: Understanding the Roles of Rush Rhees, Elizabeth Anscombe and Georg Henrik von Wright*. London: Bloomsbury Academic, pp. 105-136.

Stenlund, S. (2002) 'The Craving for Generality', *Revista Portuguesa de Filosofia*, 58(3):569-580.

Wittgenstein, Ludwig (2009): *Philosophical Investigations / Philosophische Untersuchungen*. Hg. von P. M. S. Hacker und Joachim Schulte, übersetzt von G. E. M. Anscombe, P. M. S. Hacker und Joachim Schulte. New York: Wiley.

Wittgenstein, Ludwig (1998): *Preliminary Studies for the "Philosophical Investigations". Generally Known as The Blue and Brown Books*. Oxford: Basil Blackwell.

Wittgenstein, Ludwig (1978): *Remarks on the Foundations of Mathematics*. Hg. von G. H. von Wright, R. Rhees und G. E. M. Anscombe, übersetzt von G. E. M. Anscombe. Third edition, revised and reset. Oxford: Basil Blackwell.

Standpoint Epistemology and the Problem of Reflection.

Carla Spielmann (Bonn, Germany)

Abstract

Emerging from a critique of traditional conceptions of knowledge, particularly their claims to universality and neutrality, Standpoint Epistemology foregrounds the socio-political situatedness of the knowing subject, thereby integrating ethical concerns into epistemic and epistemological practice. Its insistence on taking subjective perspectives seriously in processes of conceptualizing and evaluating knowledge has, however, provoked accusations of epistemic relativism. In response to such allegations, Sandra Harding introduced the concept of *Strong Objectivity*, arguing that Standpoint Epistemology, far from lapsing into epistemic relativism, actually produces more objective knowledge. *Strong Objectivity* rests on the idea of *strong reflexivity*, understood as the capacity for critical self-reflection exercised within, and through, a heterogeneous community of epistemic subjects.

In this paper, I argue that Harding, along with those who have defended her position, lacks a sufficiently robust account of critical reflection, one that would be compatible with the feminist commitment to social constructionism regarding the epistemic subject. Consequently, their response to the charge of relativism remains unconvincing. More recent attempts to clarify the possibility and function of critical reflection within Standpoint Epistemology also fail to adequately address this crucial issue of reflexivity.

1. Introduction

Standpoint epistemology grew out of the feminist critiques of standard concepts of knowledge, its objectivity, and the knower. Feminist epistemologists accuse traditional, ‘western’ philosophy, going back to philosophers such as René Descartes, John Locke, and Immanuel Kant, of having false idols of “pure objectivity and value-neutrality” (Code 1993: 16f., 33), which presuppose the complete detachment of the epistemic subject from everything concrete and contingent, i.e., in their view, from everything subjective (ibid.). Convinced that every epistemic subject is necessarily situated, i.e. conditioned by the socio-political, historical, and physical situation of the subject, traditional epistemology is accused of ignorance and a false universality. Insofar as the goal of complete detachment is impossible, traditional epistemology is said to tacitly posit a white, male, or heterosexual perspective as the ideal, universal, and, most importantly, objective epistemic perspective (ibid.: 22f.).

Standpoint epistemology claims to thereby reveal hidden power dynamics within epistemology that reflect sexist, racist, or Eurocentric structures of

society. In sum, mainstream epistemologists are not only ignorant of the situatedness of others, but also of their own and the power dynamics behind it. As a result, mainstream epistemology ignores the ethical dimension of knowledge, perpetuates an illegitimate epistemological dominance, and fails in its own responsibility (Haraway 1988: 580ff.).

Based on the social constructionist assumption that power structures shape the situatedness of and thus produce epistemic subjects, feminist epistemologies such as standpoint epistemology argue that epistemic practice can be discriminatory when hierarchising epistemic situations. Epistemology is thus faced with the imperative of becoming aware of its epistemic responsibility and acting on its behalf by actively reducing discriminatory practice by integrating plural epistemic perspectives.

A key reference for the ethical dimension of epistemological issues is Miranda Fricker's theory of epistemic injustice, which introduces two different kinds of injustice within epistemic practice: testimonial and hermeneutic injustice. While the former refers to the denial of the credibility of epistemic practices, i.e. the questioning of their rationality, the latter refers to the dominance of certain epistemological structures and norms that make the intelligibility of different experiences impossible and thus deprive individuals of making sense of the world (Fricker 2017: 53f.).

The feminist emphasis on the intertwining of ethics and epistemology argues not only for ethical problems in epistemology, but also for epistemological problems arising from moral shortcomings. The exclusion of both epistemic and epistemological perspectives outside the dominant research landscape leads to a limited understanding of knowledge, its objectivity, and the knower. Consequently, feminist thinkers argue that epistemology should not aim for neutrality, but for a pluralist moral account that defines the epistemic subject and objectivity in different ways: "The knowing self is partial in all its guises, never finished, whole, simply there and original; it is always constructed and stitched together imperfectly, and therefore able to join with another, to see together without claiming to be another. Here is the promise of objectivity: a scientific knower seeks the subject position, not of identity, but of objectivity, that is, partial connection" (Haraway 1988: 586).

However, this focus on subject positions and the emphasis on their partiality is not meant to imply epistemic relativism and the abandonment of objectivity, but rather the opposite. By “taking subjectivity into account”, to use Lorraine Code’s words, standpoint epistemology aims for more objective knowledge than with standard, traditional epistemology. Sandra Harding’s *Strong Objectivity* is probably the most prominent account of this view.

2. Strong Objectivity

Harding sees the epistemic subject as partial, limited, and dependent on social structures, i.e. social structures condition epistemic capacities. In her view, knowers are created and as creations they support existing social structures. Therefore, she envisages the creation of knowers that support better, more plural, less absolute and dominant social structures (Harding 2006: 255). This social constructionist epistemological and political vision inherits the claim that epistemic subjects are limited and that becoming aware of this limitation will lead to less subjective relativity and, consequently, increase the objectivity of knowledge. Against this background, Harding’s famous concept of ‘strong objectivity’ introduces standpoint epistemology as a tool to defeat epistemic relativism by explicating the limitation of singular knowledge claims on the one hand, and enlarging the validity of knowledge claims by increasing their plurality on the other.

She defines a ‘standpoint’ as an epistemic achievement. To have or take a standpoint requires prior reflective as well as practical engagement; which is also why, standpoint epistemology is closely related to political activism. Assuming that by actively trying to live the way another one lives, i.e., to do what they do and believe what they believe, one takes another one’s standpoint. The same accounts for oneself: only by working on understanding the social, material, and historical location that situates one’s own thinking, one comes to have a standpoint (Harding 2009: 194f.).

As such, standpoints require what Harding calls ‘strong reflexivity,’ i.e., to reflect not only on the object, but also on the subject of knowledge, to maximise objectivity by uncovering value, motive, historical, and social position (Harding 1993: 71). Strong reflexivity thus aims at both self-understanding and an understanding of the other, insofar as it sets out to fully

acknowledge one's own situatedness and thus epistemic limitations, which in turn relies on exchange with and an understanding of the others. As Harding says: "[u]nderstanding ourselves and the world around us requires understanding what others think of us and our beliefs and actions, not just what we think of ourselves and them" (ibid.: 72). In other words, strong reflexivity requires taking the other's position, their social location, which conditions how they think about themselves and others.

As standpoint theories are more rigorous when it comes to reflecting their methods, aims, and solutions, Harding argues that their scientific standards are higher than the dominant ones, which do not allow for a detection of bias. As a consequence, standpoint epistemology conducts more objective research (ibid.: 51f.), in the sense that it is less relative to singular standpoints (ibid.: 50).

Based on the view that every knowledge claim is inevitably socially situated, Harding argues that the acknowledgement and systematic inclusion of plural standpoints is a necessary condition for objective knowledge (ibid.: 57). The plurality of epistemic subjects is thus the central aspect of objective knowledge (Harding 2006: 258). However, Harding, such as other standpoint epistemologists, support the view that marginalised perspectives are epistemically privileged, as they enable criticism of dominant theories and concepts, given that they are opposed to them. Therefore, they are more likely to uncover blind spots and thus "are better than others as starting points for knowledge projects" (Harding 1993: 56). The idea is hence to prioritise marginalised standpoints in epistemological endeavour to reduce bias, and, consequently, increase objectivity: "[A] maximally critical study of scientists and their communities can be done only from the perspective of those whose lives have been marginalised by such communities" (ibid.: 69). Such a "strong standpoint project" (Harding 2006: 255) aims at changing the epistemic structures that support discrimination, and, thereby, improve epistemic practice. It addresses the ethical and epistemological issues outlined so far.

3. The Problem of Reflection

In summary, standpoint epistemology combines three theses: the situated knowledge thesis which highlights the partiality of knowledge claims, the achievement thesis which emphasises critical reflection as central means to

acquire knowledge, and the epistemic privilege thesis which focusses on the benefits of marginalised perspectives (Toole 2021: 340ff.).

Standpoint theory has often been accused of theoretical inconsistency, mostly arguing for the incompatibility of the situated knowledge thesis and the epistemic privilege thesis. The assumption that every knowledge is situated is apparently not compatible with the idea that marginalised perspectives are epistemically privileged due to their social location that opposes the dominant one. Put differently: How come some of the limited perspectives are apparently less limited? Kristina Rolin gives an analysis of why this “bias paradox”, as she calls it, “is not a paradox after all” (Rolin 2006: 130). She contends that the thesis of epistemic privilege does not entail any form of epistemic absolutism. Rather, it remains compatible with the concept of situated knowledge, insofar as examples of epistemic privilege are to be understood as default assumptions that function as foundational premises within a given epistemic context. While such assumptions enjoy a provisional epistemic authority, they remain, in principle, subject to debate, with the burden of justification falling upon those who seek to contest them. Accordingly, the notions of epistemic privilege and situated knowledge are compatible, and standpoint epistemology remains theoretically robust and coherent (ibid.: 126ff.).

The compatibility of the achievement thesis and the situated knowledge thesis is, however, less discussed, although it is the crucial aspect for the argument against epistemic relativism. The claim that knowers are not determined by their social location is mainly based on the achievement of a standpoint through critical reflection (Harding 1993: 64; Toole 2021: 343). In this view, the knower is in somewhat distanced position to the social situation which enables knowledge rather than mere perspective.

However, as I argue, standpoint epistemology does not acknowledge that the basic social constructionism regarding the epistemic subject, which states that “one's social situation enables and limits what one can know” (Harding 1993: 55), cannot account for the epistemic subject's critical reflection on their own and others' situations, because it entails that the capacity of reflection is as well limited by one's social situation.

In other words, the social constructionist framework does affect the capacity of reflection as well, which undermines the achievement thesis, giving rise to the following tension: If reflection is itself a product of social processes, it is no obvious means of critically investigating and evaluating social structures. Conversely, if reflection is not conceived as produced by social structures, it is effectively reinstated as an epistemic universal, which not only undermines the feminist commitment to the plurality and situatedness of epistemic subjects, but also trivialises the very epistemic concerns feminist theorists aim to address, by presupposing a solution that was presumably available all along.

The questions remain: How does one come to know one's epistemic practices and their circumstances if they are the enabler of one's knowledge of them? And how does one come to know the epistemic practices of others when one can only rely on one's own to do so?

Harding's own account falls short on giving a proper answer to these questions. The mere reference to a plurality of local situations, i.e., a heterogenous discursive field (Harding 2006: 257f.), does not explain how any subject comes to reflect. Another perspective does not automatically enable a critical stance towards one's own or the recognition of the other. Rather, the accurate inclusion of perspectives does not guarantee but exceeds itself 'critical reflection' (Intemann 2010: 789; Arnold et al. 2019: 4f.).

Harding might not care for the theoretical accuracy of standpoint epistemology, as she says herself that any theoretical inquiry is judged by its practicability: "The adequacy of standpoint projects is to be judged by the success of the practices they legitimate rather than the truth or verisimilitude of representations of nature and social relations. (Recollect that the success of practices has always been the ultimate test of the adequacy of scientific claims.)" (Harding 2009: 195).

Nonetheless, I explore two approaches that give an account of the possibility and usefulness of critical reflection within social constructionism about epistemic subjects. Paige Sweet's analysis focusses more on self-reflection, while Briana Toole centres on the reflection on other's perspectives.

4. Possible Answers

Sweet argues that reflexivity is to be understood as self-awareness of social positions, which is only productive when rooted in a marginalised social location (Sweet 2020: 925f.). Reflection, she argues, is enabled only within social relations, and those that are marginalised are in a position to “see[] the social world” (ibid.: 926). This epistemic privilege of a marginalised perspective can be seen as a somewhat natural reflective stance towards social structures. In this context, Sweet speaks of incorporation, and of social product, when characterising the marginalised perspective’s knowledge. Ultimately, she considers the marginalised standpoint a kind of practical knowledge derived from the social position (ibid.: 930f.). In this framework, the kind of productive reflexivity that standpoint epistemology is aiming for, derives from a socially produced marginalised position, which inherently grants knowledge of the social world. Resultingly, reflexivity and epistemic privilege are to some extent the same.

This is problematic, because it does not explain how reflection can be used as an active tool. Instead, it renders it a passive, “contingent outcome” that turns any epistemic inquiry into a matter of chance (ibid.: 928). Productive reflection derives from social positioning that dominates the epistemic subject, wherefore the epistemic subject is unable to intentionally engage in critical reflection. Rather, critical reflection comes naturally from a marginalised position. Furthermore, Sweet provides no additional information about the transition from marginalised experience to objective knowledge (ibid.: 930), so it remains unclear how social positions can be somewhat inherently reflexive and produce objective knowledge, while others are restricted to a limited perspective. Ultimately, Sweet’s framework leaves open the question of how reflexivity, understood as a productive, knowledge-generating process, can be actively cultivated, as standpoint epistemology aims to do, rather than passively inherited through social position.

Briana Toole’s account focusses more on reflection of the other’s standpoint. Drawing on Laurie Paul’s argument of “de-se knowledge”, Toole argues that the sharing of standpoints, i.e. the reflection on other perspectives, requires more than an imagination of having the other one’s experience. Instead, it requires the imagination of ‘being’ the other person by adopting their first-personal perspective. Toole identifies our capacity for empathy as a legitimate

means of accessing another's first-personal perspective. In her view, empathy, which gives way to becoming the other, thus functions as a kind of reflection (Toole 2022: 59f.).

Yet, this is not a sufficient explanation of productive reflection on another's perspective for two reasons. First, to adopt another's first-personal perspective is to abandon one's own perspective, from which alone one knows about *the other*. Becoming the other results in giving up the distinction between oneself and the other, which undermines the plurality of standpoints that is crucial in standpoint epistemology. Second, Toole's account ties reflection to singular perspectives, but attempts to counter this relativism by positing the possibility of inhabiting another's first-personal perspective. This description of the attempt to imaginatively *become* the other does not offer a coherent explanation of how we come to *know* the other's perspective. There is a difference between being and knowing, that is mirrored by the standpoint epistemology's distinction between experienced perspectives and reflexively achieved standpoints. This is not sufficiently acknowledged in Toole's account.

Distinguishing between perspectives and standpoints, on the one hand, and between different perspectives and different standpoints, on the other, is central to standpoint epistemology, in which the plurality of reflexively achieved standpoints is a foundational element of objectivity. In sum, Toole's framework still lacks an account of how standpoints, rather than perspectives, can be achieved and translated across different first-personal perspectives, which is the central function of critical reflection.

5. Conclusionary Remarks

I have argued that the social constructionist understanding of the epistemic subject of standpoint epistemology cannot account for critical reflection as a tool for accessing and integrating plural standpoints. This lack of a proper concept of reflexivity weakens the achievement thesis central to standpoint epistemology which introduces the distinction between perspectives and standpoints achieved through critical reflection on those perspectives, as an argument against epistemic relativism.

One possible way to resolve the tension around reflexivity within standpoint epistemology is to differentiate between external factors that undermine subjective capacities, such as the capacity for reflection, and the capacities themselves. If one asserts that every epistemic capacity is socially dependent in such a way that the social structures dominate them, one faces the challenge of explaining how, firstly, to become reliably aware of this domination, and, secondly, how to change this. Harding herself highlights the tension between socially situated knowledge claims and the need for critical distance from those social contexts in order to achieve more objective knowledge (Harding 2015: 35ff.).

References

- Arnold, C., Stetten, M. v., & Schmitz, A. (2019) "Wie können sozial situierte Subjekte objektive Standards bereitstellen? Überlegungen in Anschluss an die Methodologie Sandra Hardings", N. Burzan (ed.) *Komplexe Dynamiken globaler und lokaler Entwicklungen: Verhandlungen des 39. Kongresses der Deutschen Gesellschaft für Soziologie in Göttingen 2018*, Deutsche Gesellschaft für Soziologie (DGS).
- Code, Lorraine (1993) "Taking Subjectivity into Account", Linda Alcoff & Elizabeth Potter (ed.) *Feminist Epistemologies*, New York: Routledge, 15-48.
- Fricker, Miranda (2017) "Evolving Concepts of Epistemic Injustice", Ian James Kidd, José Medina & Gaile Pohlhaus, Jr. (ed.) *The Routledge Handbook of Epistemic Injustice*, New York: Routledge, 53-60.
- Harraway, Donna (1988) "Situated Knowledges: The Science Question in Feminism and the Privilege of Partial Perspective", *Feminist Studies*, vol. 14, no. 3, 575-599.
- Harding, Sandra (1993) "Rethinking Standpoint Epistemology: "What Is Strong Objectivity"?" Linda Alcoff & Elizabeth Potter (ed.) *Feminist Epistemologies*, New York: Routledge, 49-82.
- Harding, Sandra (2006) "Transformation vs. Resistance Identity Projects: Epistemological Resources for Social Justice Movements", Linda Alcoff et al. (ed.) *Identity politics reconsidered*, New York: Palgrave Macmillan, 246-263.
- Harding, Sandra (2009) "Standpoint theories: Productively controversial", *Hypatia*, vol. 24, no. 4, 192-200.
- Intemann, Kristen (2010) "25 Years of Feminist Empiricism and Standpoint Theory: Where Are We Now?", *Hypatia*, vol. 25, no. 4, 778-796.

Harding, Sandra (2015) *Objectivity and Diversity. Another Logic of Scientific Research*. Chicago: University of Chicago Press.

Rolin, Kristina (2006) "The Bias Paradox in Feminist Standpoint Epistemology", *Episteme: A Journal of Social Epistemology*, vol. 3, no. 1-2, 125-136.

Sweet, Paige L. (2020) "Who Knows? Reflexivity in Feminist Standpoint Theory and Bourdieu", *Gender & Society*, vol. 34, no. 6, 922-950.

Toole, Briana (2022) "Demarginalizing Standpoint Epistemology", *Episteme*, vol. 19, 47-65.

Toole, Briana (2021) "Recent Work in Standpoint Epistemology", *Analysis*, vol. 81, no. 2, 338-350.

Towards a Feminist Philosophical Inquiry on Motherhood

Krissah Marga Taganas (Quezon City, Philippines)

Abstract

Motherhood remains a marginalized area in feminist philosophy. While there were early accounts of philosophical exploration of motherhood (De Beauvoir 1949/2012, Ruddick 1989, hooks 1984/2014, Collins 1994) and recent discourse on philosophy of pregnancy and childbirth (Kingma & Woollard 2025), the present literature often only tackles issues surrounding motherhood, such as adoption, parenting, bodily autonomy, childrearing, and reproduction in areas like ethics, metaphysics, epistemology, and socio-political philosophy.

In this paper, I build on Sara Ruddick's initial notes on feminist maternal standpoint in *Maternal Thinking* (1989), arguing that a sustained feminist philosophical inquiry into motherhood is pressing and crucial. I claim that feminist philosophy brings new understandings and meanings of motherhood, maternal experiences and emotions, and empowerment, more specifically by: 1) expanding the meaning of "mother", 2) challenging the material and ideological conditions of motherhood, 3) rethinking maternal emotions and affect, such as guilt, blame, anger, and ambivalence, and 4) committing to this inquiry as resistance. To do this, I build my arguments by examining the status of motherhood studies as a distinct academic field and identify reasons for justifications why feminist philosophy of motherhood is philosophically pressing and salient. I conclude by emphasizing the relevance and importance of this issue in the fields of feminism and philosophy.

Over the years, motherhood has occupied diverse positions and meanings. Motherhood in the 21st century has profoundly changed society's understanding of mothering, bodily autonomy, family, identity, and even the global economy. The growing area of motherhood studies in the Global North takes an interdisciplinary approach, drawing from various disciplines such as sociology, philosophy, psychology, gender and sexuality studies, history, international relations, and literary studies. This space enabled mothers to formally theorize and critique their experiences and practices of motherhood. Discourses on motherhood emerged from what Adrienne Rich (1976) identified as patriarchal motherhood as an oppressive institution and mothering as an empowering concept. Feminism's relationship with motherhood is a common and recurring theme in motherhood scholarship. More recently, the term "matricentric feminism" became a crucial concept in the motherhood literature (O'Reilly, 2016). Motherscholars found this emergent feminism to be an insightful perspective in motherhood studies. According to Andrea O'Reilly: matricentric feminism is "explicitly matrifocal in its perspective and emphasis, it begins with the mother and takes seriously the work of mothering, and that

it is multidisciplinary and multi-theoretical in its perspective” (2016:18). Matricentric feminism, then, regards motherhood studies as a sustained scholarly inquiry about mothers. Although this is considered a new form of feminism, matricentric feminism aims to create a feminism for and about mothers with the recognition that mothers’ desires, experiences, and needs continue to be marginalized.

In philosophy, accounts of motherhood were sporadic in literature (Wollstonecraft, 1792; De Beauvoir 1949/2015; Firestone 1970; Daly 1978; Ruddick 1989; hooks 2014; Collins 1994). Writings on related motherhood concerns, such as family, body, and reproduction, have been tackled by philosophers in metaphysics, epistemology, socio-political philosophy, and feminist philosophy. However, philosophical inquiries from and about a subjective maternal point of view are lacking (Lintott & Staudt, 2012). In comparison with other universals in human experience, such as death or beauty within the canon of philosophy, the subject of motherhood and maternal perspectives, like birth, was not given particular philosophical attention over the years. According to Lintott and Staudt (2012: 3): “Motherhood and related maternal experiences have been alternatively ignored, derided, exploited, or naturalized to the point of utmost reduction of their philosophical implications by most of the famed philosophers of the West and East.” Motherhood is replete with philosophical dilemmas, and philosophical and critical reflection in formal contexts remains desirable.

The recent decades have demonstrated the potential of feminist philosophy in changing society. With this in mind, this paper builds on Sara Ruddick’s initial arguments on developing a feminist maternal standpoint. Ruddick notes:

“Like a women’s politics of resistance, feminism shifts the balance within maternal practice from denial to lucid knowledge, from parochialism to awareness of others’ suffering, and from compliance to stubborn, decisive capacities to act. This transformation begins, paradoxically, in a tense relationship between feminists and women who are mothers. Mothers and feminists cannot leave each other alone.” (1989: 236)

Ruddick emphasized the complex yet significant role of feminist politics in mothering, yet up to this day, motherhood remains a marginalized area in

feminist thought and feminist philosophy. In this paper, I identify four (4) reasons why a feminist philosophical inquiry of motherhood is an important and crucial work in feminist and philosophical discourse. I argue that feminist philosophy can bring new understandings to motherhood as practice and experience by 1) expanding the meaning of “mother”, 2) challenging the material and ideological conditions of motherhood, 3) rethinking maternal emotions and affect, and 4) committing to this inquiry as resistance. In the next sections, I discuss my reasons for justification.

EXPANDING THE MEANING OF “MOTHER”

Adrienne Rich (1976) identified motherhood as a patriarchal institution that controls how women do mothering and become mothers. The recognition of the oppressive forces of this institution also enabled ways for empowerment. Motherhood, both a site of oppression and empowerment, is not monolithic. Every mother’s experience is unique from one’s context and intersectional identities. Ruddick recognized the problem of identifying mother as a fixed biological or legal relationship with children. This definition undermines some social and cultural practices of motherhood, particularly those who do mothering even without the genetic or biological requirement. For Ruddick, “anyone who commits her or himself to responding to children’s demands, and makes the work of response a considerable part of her or his life, is a mother.” (1989:xii). Likewise, Gil Anidjar mentions: “I write of mothers, not of women; of mothering, not of caring; obviously, women have been mothers, yet, equally clear and certain is the fact that not all women mother, nor are all mothers women.” This is an important distinction—women should not be reduced to motherhood, and motherhood should not be equated with womanhood (Anidjar, 2024:xxiv).

The definition of “mother” as an individual who engages in motherwork, opens possibilities to philosophically challenge the traditional roles of women as mothers. A feminist inquiry into the concept of “mother” provides an expansion of its meaning, as reflected in non-traditional practices of mothering, and sheds light on society’s understanding of the responsibilities of childrearing, caring, and mothering. Feminist Philosophy has been changing conceptions and practices through the constant reevaluation of concepts like woman, gender, and sexuality. Further extending this work in motherhood brings new meanings and perspectives to mothering. Philosophizing on the

term ‘mother’ not only embraces diverse experiences of those who engage in mother work, but also illuminates the gendered dimension of motherhood. Collins’ (1994) *Othermothering*, which is rooted in Black feminist activism, is a good example of this case. It is the concept of accepting responsibility for a child that is not one’s own in an arrangement that may or may not be formal” (Story, 2014:11). Migration is also changing the dynamics of motherhood, as in the case of Filipina migrant workers (Parreñas, 2010). In these Filipino families, every part of the extended family uniquely takes part in mothering in the absence of the mother. As Parreñas notes, the feminization of migration destabilizes families and contributes to the emergence of transnational families.

Hence, this expansion enables us to rethink about motherhood in the cases of non-traditional family settings, such as adoptive, trans, queer, and non-binary mothers. It poses challenges to dichotomies in sex and gender, and emphasizes the complexity this brings to motherhood, marriage, and the family. Finally, this not only counters gender essentialism in motherhood but also promotes inclusivity in the experiences of mothers.

CHALLENGING THE MATERIAL AND IDEOLOGICAL CONDITIONS OF MOTHERHOOD

Feminism, as a movement, has long contested ideologies and stereotypes about women. This includes criticizing patriarchal-defined motherhood and essentialist thinking on women and mothers. A sustained feminist philosophical inquiry into the material and ideological conditions of motherhood creates space to resist these imposed values and dismantle oppressive systems. As Ruddick describes, “In becoming feminists, mothers acquire a ‘feminist consciousness’, a confusing, often painful, but irresistible recognition that the stories they have told themselves about “being a woman” are self-deceptive and do not serve their interests.” (1989: 237) To name and resist oppression, for mothers, means challenging the romantic and idealized accounts of motherhood.

Further inquiry on these motherhood conditions, the following concepts of motherwork, intensive mothering, normative motherhood, and care ethics—all of which are crucial in motherhood inquiry, emerged. *Motherwork* (Collins, 1994) is defined as “nurturance, love, or mothering in general”. This term encapsulates the problem of disregarding mothering as labor. Motherhood as

labor engages with the issues of the economic realities of mothers and families. This concept can be extended philosophically through illuminating how motherhood as labor is also an embodied experience, generative of moral and political questions. On the other hand, Sharon Hays (1996) argued that *intensive mothering* is a kind of mothering centered around children's needs, with methods informed by experts, and is labor-intensive. Basically, mothers should prioritize their children above all else and spend all their time and energy on their children. This concept criticizes the unrealistic expectations imposed on mothers regarding duties and responsibilities surrounding motherhood. In philosophically probing into this issue, the assumption of these "innate" characteristics of mothers raises the lack of focus on the roles and contributions of fathers.

Normative motherhood emphasizes "good and bad motherhood" as a regulatory institution (O'Reilly, 2016). These expectations compound to form the patriarchal institution of motherhood—persistent imposition of high and impossible standards of motherhood and womanhood. Opportunities to challenge mothers' oppressive conditions introduce new questions and redefine the motherhood experience and identity. In philosophy, the issues raised by feminist philosophers Rich (1976), Daly (1978), and Firestone (1970) on how patriarchal expectations of motherhood continue to demean women have been addressed by *care ethics* in the works of Nel Noddings (1984) and Carol Gilligan (1982). While care ethics has been initially associated with motherhood, this notion remains a rich area for further philosophical exploration.

Feminist Philosophy is not only limited to the identification of oppressive barriers, but also commits to offering alternative ways of approaching these issues. Criticizing these ideologies makes us rethink motherhood as a socially and historically constructed practice, rather than a defining dimension of a woman's life. This also breaks and challenges the notion of intrinsic motherhood and the associated maternal qualities that reify women's roles and identities as mothers. Technological advancements and the social media age have drastically changed contemporary motherhood. The new motherhood image meant new responsibilities and expectations for mothers.

The phenomenon of celebrity moms, Instagram mommies, and the return of the traditional wife (tradwife) values warrants a more comprehensive analysis of its impact on motherhood and women empowerment.

RETHINKING MATERNAL EMOTIONS AND AFFECT

While mothers have unique and diverse experiences in mothering, one common feature is the recognition that motherhood is an emotional experience. Ruddick's emphasis on feminist consciousness includes:

“For a mother, ‘coming to know the truth’ includes looking at the real feelings and conflicts of mothering. It is a feminist project to describe realistically the angers and ambivalences of maternal love.” (1989: 237)

A cognizance of the conflicting maternal emotions, such as love, ambivalence, anger, guilt, and shame, contributes to conscious awareness of the role of emotions in mothering. Philosophical probing of maternal emotions enables conversations on how emotions empower mothers to actively engage and construct their mothering experiences.

The culture of the good-bad mother takes a disciplinary force in mothering. This is evident in how patriarchal motherhood has prevented mothers from tackling emotions that deviate from the maternal feelings and values associated with motherhood. Mothers who experience nonmaternal feelings or don't seem to possess intrinsic maternal qualities pose a threat and raise suspicion about their capability to do mother work. Cultural ideologies continue to dictate the emotional framework involved in motherhood. Feelings of guilt, regret, blame, anger, and ambivalence stand in opposition to the maternal love and sacrifice of a good mother, and consequently, must be rejected. This view on maternal emotions further cements patriarchal motherhood in women's roles and identities. Studies on maternal emotions appear marginal to mainstream discussions of motherhood.

In the Western philosophical canon, emotions, mostly associated with women, have been considered subversive of knowledge (Jaggar 2007). The reason/emotion dichotomy identifies emotion as linked to feminine appetites. One of the main projects of feminist philosophy is to challenge this dichotomy and rethink the value of emotions to human experiences (Jaggar 2007, Nussbaum 2001). In challenging the traditional binary of reason and emotion, feminist

philosophers argue for the importance of emotions and how they are shaped by socio-political contexts, gender norms, and power structures. Emotions are regarded as crucial and intelligent judgments that inform knowledge and action (Nussbaum, 2001).

A feminist philosophical inquiry into maternal emotions of blame, rage, anger, ambivalence, and guilt can provide an epistemological, feminist, and political resource to motherhood discourse. Questions on how to respond and regulate maternal emotions bring attention to the multifaceted and often complex motherhood experiences. Moreover, feminist epistemology emphasizes the role of emotion in understanding the world. By providing perspectives on how societal expectations on mothers shape how they receive and experience these emotions, we also understand the phenomenological and epistemological aspects of mothers' experiences, complementing the science and psychology of maternal health. This also makes empowerment possible, where tensions and conflicts are negotiated through making sense of one's emotions and experiences in motherhood. It further supports the idea that conflicted emotions can coexist with positive maternal emotions of happiness, fulfillment, joy, and love. As Jaggar notes, "just as appropriate emotions may contribute to the development of knowledge, so the growth of knowledge may contribute to the development of appropriate emotions" (2007: 700). The emotional insights from the mother's subjectivity may generate new questions and directions in political and emotional practice in motherhood.

With these lines of inquiry, even the relationships of mothers and children can be transformed when mothers are finally permitted to experience, reeducate, and regulate all these emotions and are equipped with tools on how to respond to them. Finally, the lived experience of a mother undergoing major changes in their life through motherhood takes center stage in this discourse. As we redefine concepts of care, devotion, suffering, and agency, we also find the ground on which mothers can still feel empowered and liberated with all these complexities.

COMMITTING TO THIS INQUIRY AS RESISTANCE

Feminist politics, as underscored by Ruddick, play a key role in the motherhood discourse. She highlights, "To be a feminist mother is to recognize that many dominant values—including, but not limited to, the subordination

of women—are unacceptable and need not be accepted.” (1989: 238) A feminist perspective on motherhood is rooted in the goal of challenging and changing the system that oppresses mothers. Feminist and intersectional lenses provide a broader framework to interrogate how systems of gender, power, and social institutions continue to impact and shape motherhood experiences.

This inquiry is aligned with the goals of feminist philosophy. The commitment to social change and justice means an inquiry rooted in resistance to the most damaging forms of patriarchy. Staying true to its commitment to theory and practice, a feminist philosophical inquiry into motherhood positions motherhood as a social and political act. This stresses the importance of this inquiry, as further analyses on motherhood reveal how these struggles, while distinct from other experiences of women, are not in opposition and in conflict with feminist goals.

Motherhood does not only concern mothers, but everyone. Since we are all born of mothers, and we take part in aspects of mothering and being mothered, these philosophical inquiries raise important moral and political questions about human life. The continuous interrogation of patriarchal standards of motherhood, reevaluation of emotions, and emphasis on the primacy of mothers’ subjective experiences constitute the resistance to oppressive ideologies in motherhood. These discourses provide responsibilities, both individually and collectively, to contemplate social justice and the good life (Lintott & Staudt, 2012).

Motherhood discourse in both feminism and philosophy carries huge potential for transformation. For one, a feminist philosophical perspective supplies a deeper understanding of motherhood-related issues, like reproduction, teenage pregnancy, and development and health-focused dialogues. On the other hand, this inquiry aims to further expound on issues involving justice in the family, such as domestic violence, child abuse, power dynamics, and violence against women and children. As a resistance, this hopes to forge communities and institutions to support mothers, provide space to redefine family, and pave the way for a rich scholarship inquiry and public conversation on motherhood. Feminist Philosophy holds great value in

reflexivity (Alcoff & Kittay, 2007). This shall warrant strong conversations on the dynamic nature of motherhood, further aiming for solidarity, inclusion, diversity, and positionality.

CONCLUSION

In this paper, I have argued for the importance of a feminist philosophical inquiry into motherhood and have shown how this inquiry project is aligned with the broader goals of the feminist movement and feminist philosophy projects. Motherhood studies is fundamentally intersectional and multidisciplinary; thus, feminist philosophical perspectives on motherhood play a crucial role in reevaluating its concepts and issues, and changing and shaping the public discourse. I have identified four (4) reasons why a feminist philosophical inquiry into motherhood is desirable. First, in a changing and dynamic society, the goal of expanding the meaning of “mother” proves to be a crucial aim for inclusion, such as to accommodate and embrace non-traditional identities and experiences. The term “mother” is interrogated and provides a space to conceptualize motherhood beyond just an identity set for women in society. Second, alongside the aim to address this essentialist thinking in motherhood discourse, further inquiry into motherhood necessitates the examination of the cultural and ideological stereotypes in motherhood. Motherhood, as a practice, is also a rich area for philosophical scrutiny. Feminist probing into the various practices that harm, demean, and demonize mothers is crucial to the goal of providing mothers with tools to resist this condition. Third, maternal emotions can be reassessed and revalued to reflect the many ways they can affect, influence, and even shape one’s experience of motherhood. Finally, on a wider scale, commitment to a feminist inquiry on motherhood emphasizes the potential of these discourses to be liberatory, not only on an individual aspect, but more so on the levels of policies, political, and societal changes.

This inquiry hopes to enrich the discourse, raise awareness, and contribute to a lucid understanding of gender and power structures in inspiring social and political change. Motherhood is rethought as a site of empowerment and resistance, further defying patriarchal ideologies and expectations, and making space for power and liberation.

References

- Alcoff, Linda & Kittay, Eva. (2007). *The Blackwell Guide to Feminist Philosophy*. Blackwell Publishing.
- Anidjar, Gil. (2024). *On the Sovereignty of Mothers*. Columbia University Press.
- Collins, Patricia Hill. (1994). "Shifting the Center: Race, Class, and Feminist Theorizing About Motherhood." *Mothering: Ideology, Experience, and Agency*, edited by Evelyn Nakano Glenn et al., Routledge
- Daly, Mary. (1978). *Gyn/Ecology: The Meta-Ethics of Radical Feminism*. Boston: Beacon.
- De Beauvoir, Simone. (1949/2015). *The Second Sex*. Vintage Classics.
- Firestone, Shulamith. (1970). *The Dialectic of Sex: The Case for Feminist Revolution*. New York: Bantam.
- Gilligan, Carol. (1982). *In a Different Voice*. Boston: Harvard University Press.
- Hays, Sharon. (1996). *Intensive Mothering: The Cultural Contradictions of Motherhood*. Yale UP
- Hooks, Bell. (1984/2012) *Feminist Theory from Margin to Center*. South End.
- Jaggar, Alison. (2007). "Love's Knowledge: Emotion in Feminist Epistemology", in Alison Bailey and Chris Cuomo (eds.) *The Feminist Philosophy Reader*, NY: McGraw Hill, 687-705.
- Kingma, Elslijnn & Woollard, Fiona. (2025). "Can You Do Harm to your Fetus? Pregnancy, Barriers, and the Doing/Allowing Distinction". *Ethics*. 135 (2) The University of Chicago Press.
- Lintott, Sheila & Sander-Staudt, Maureen (2012). *Philosophical Inquiries into Pregnancy, Childbirth, and Mothering*. NY: Routledge.
- Noddings, Nel. (1984). *Caring: A Feminine Approach to Ethics and Moral Education*. Berkeley: University of California Press.
- Nussbaum, Martha. (2001). *Upheavals of Thought: The Intelligence of Emotions*. Cambridge: Cambridge University Press.
- O'Reilly, Andrea (2016). *Matricentric Feminism: Theory, Activism and Practice*. Demeter Press
- Parreñas, Rhacel. (2010) "Transnational Mothering: A Source of Gender Conflicts in the Family", 88 *N.C.L.Rev.* 1825
- Rich, Adrienne. (1976). *Of Woman Born: Motherhood as Experience and Institution*. 2nd ed., W.W. Norton.
- Ruddick, Sara. (1989). *Maternal Thinking: Toward a Politics of Peace*. Boston: Beacon Press

Story, Kaila. (2014). *Patricia Hill Collins: Reconceiving Motherhood*. Demeter Press.

Wollstonecraft, Mary. (1792). *A Vindication of the Rights of Woman*. Penguin Books.

On Wittgenstein on Translation

Tue Trinh (Berlin, Germany)

Abstract

Proposition 4.025 of the *Tractatus* describes translation as replacing each word in the original language text with its target language counterpart. This requires a one-one correspondence between vocabularies of the two languages, which never transpires in real life. I argue that Wittgenstein's remark should be read as articulating the theory of language which he promoted in the *Tractatus*: it tells us what translation looks like when that theory applies under ideal conditions. I then propose an account for the difference between translation in real life and translation as described by Wittgenstein. The account draws on the fact that language has a communicative function in addition to its representative function.

Translation is one of the topics Wittgenstein touched on in the *Tractatus*. Here is what he said. (I will use the English translation by Michael Beaney (Wittgenstein 2023), and will present the English version of the propositions quoted from the *Tractatus* in square brackets.)

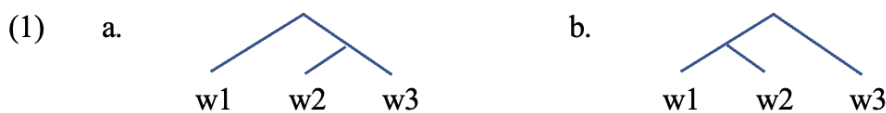
4.025

Die Übersetzung einer Sprache in eine andere geht nicht so vor sich, dass man jeden Satz der einen in einen Satz der anderen übersetzt, sondern nur die Satzbestandteile werden übersetzt. (Und das Wörterbuch übersetzt nicht nur Substantiva, sondern auch Zeit-, Eigenschafts- und Bindewörter etc., und es behandelt sie alle gleich.) [The translation of one language into another does not proceed by translating every proposition of one into a proposition of another, but only the constituents of the proposition are. (And the dictionary translates not only substantives but also verbs, adjectives, and conjunctions, etc.; and it treats them all alike.)]

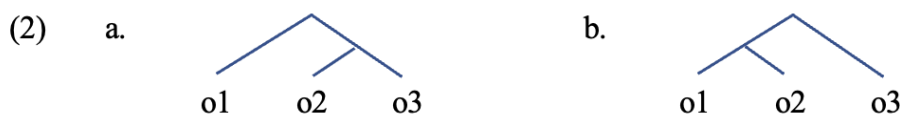
This description makes translation out to be a very dull and trivial exercise: substitute each word in the original language with its counterpart in the target language. Of course, we know this is not reality. The question is then why Wittgenstein said what he said. One possible answer is that he had no idea about translation. This I find hard to believe. Wittgenstein spoke German and English, and it is known that he translated some of his texts from German to English himself. He must have seen, immediately, that he could not just replace one word with another. Another possibility is that 4.025 is a joke which is part of a bigger joke that is the whole book. While I am not ruling out this 'resolute reading' of Wittgenstein's text (cf. Conant 1989, Diamond 1991), mostly because

I don't think I know enough about it, I do want to explore a third possibility. I propose we read 4.025 as describing not what translation is but what it is supposed to be if language works the way it is supposed to work. The *Tractatus* promotes a specific view on language: the so-called 'Picture Theory of Language', henceforth PTL (Anscombe 1959, Hintikka 2000). One way to articulate a theory is to say how things look if it is true. I submit that 4.025 articulates PTL in this sense: it tells us how translation looks if PTL is true. I will now turn to a short exposition of PTL, using a set of terms which largely but not completely overlaps with Wittgenstein's nomenclature in the *Tractatus*.

The basic idea of PTL is that language can represent reality in essentially the same way musical notation can represent sonatas and symphonies. There is a one-one correspondence between the basic building blocks of language, i.e. the 'words', and the basic building blocks of reality, i.e. the 'objects'. The 'form' of a word determines how it can combine with other words, just as the 'form' of an object determines how it can combine with other objects. Importantly, a word and the object it represents share the same form, which means that words and objects are embedded in the same space of combinatorial possibilities. To each combination of words there exists a structurally identical combination of corresponding objects. Suppose, for example, that the words w_1 , w_2 , w_3 represent the objects o_1 , o_2 , o_3 , respectively, and furthermore, that the forms of these words allow for the combination (1a) but do not allow for the combination (1b).



Then it will be the case that the combination of objects in (2a) is possible but not that in (2b).



A legitimate combination of words, i.e. one in which the words 'fit together' by virtue of their forms, is a 'sentence'. A combination of words which do not fit

together is 'gibberish'. A combination of objects is a 'situation'. There is, of course, no extra-linguistic counterpart of 'gibberish', as there can be no combination of objects which do not fit together. A sentence depicts a situation by the words being structurally related to each other in the sentence the same way the objects represented by these words are structurally related to each other in the situation. Thus, sentence (1a) depicts situation (2a). The words, each of which comes with its form, determine the set of legitimate word combinations, i.e. the set of sentences. Call this set the 'language'. The objects, each of which comes with its form, determine the set of all situations. Call this set 'logical space'. Since words and objects share the same form, language and logical space are isomorphic in the same way two combinatorial systems, i.e. two algebras, are. A 'maximal situation', i.e. one in which every object is present, is a 'possible world', or more simply, a 'world'. There is one distinguished world in logical space which is not merely possible but also 'actual'. If a sentence depicts a situation which is part of the actual world, the sentence is 'true'. Otherwise it is 'false'. We use language to 'provide information', i.e. to locate the actual world in logical space. This means when we 'say that φ ', we perform at least two acts: (i) presenting the situation depicted by φ and (ii) claiming that that situation is part of the actual world. Thus, 'saying that φ ' is the same as 'saying that [φ] is true'. Note, incidentally, that under the view of language as a combinatorial system isomorphic to logical space, 'inference rules' become superfluous, in the sense that they cannot be violated: [Mary walks] is guaranteed to be true if [John sleeps and Mary walks] is true in the same way the situation in which Mary walks is guaranteed to be part of the actual world if the situation in which John sleeps and Mary walks is part of the actual world.

5.473

Die Logik muss für sich selber sorgen [...]. Wir können uns, in gewissem Sinne, nicht in der Logik irren. [Logic must take care of itself [...]. In a certain sense we cannot make mistakes in logic.]

Let us now turn to the concept of 'translation'. I think we can agree on the following characterization of it: translating means saying the same thing in another language. Thus, translation requires there be exactly one logical space and at least two languages. Saying the same thing means claiming of the same situation that it is part of the same actual world. That would be impossible if

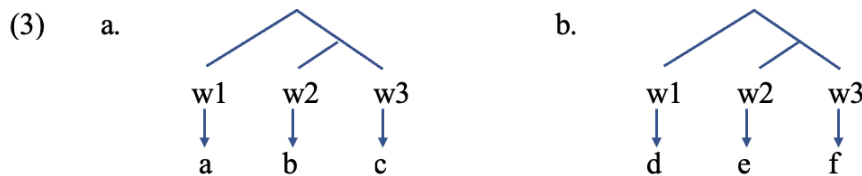
speaker and hearer dwell in different logical spaces. And there would, trivially, be no translation if there is only one language. But I have presented PTL as a theory about 'language', not one about 'languages'. How do we square PTL with Wittgenstein's remark in 4.025? To resolve this question, I propose we assign two readings to the term 'language' used by Wittgenstein in the *Tractatus*: it can either mean any symbolic system which is isomorphic to logical space, or it can mean a symbolic system which is isomorphic to logical space. Understood in the first sense, there is only one language: it is what is common to all symbolic systems isomorphic to logical space. Understood in the second sense, there can be more than one, perhaps infinitely many, languages. Now, isomorphy is an equivalence relation. That means that if two symbolic systems are isomorphic to logical form, they are isomorphic to each other. Thus, different languages are isomorphic to each other. The question then arises as to what the difference between them consists in. What distinguishes one language from another? Suppose L and L' are two different languages. Since they are isomorphic to each other and to logical space, a word w in L corresponds to a word w' in L' which is of the same form, and both w and w' correspond to an object o in logical space which is also of the same form. If a word is exhaustively identified by its form and denotation (i.e. the object it represents), w and w' would be identical, and since we have chosen w and w' arbitrarily, this holds for all words in L and L' , which means L and L' are identical, in contradiction to our supposition.

The way out of this dilemma, as I see it, is to say that form and denotation do not exhaustively identify a word. There is more to a word than how it combines and which object it represents. And it is this extra something that varies from language to language. Recall one crucial fact about language mentioned above: it is used to provide information. Our characterization of 'saying that φ ' as 'saying that $[\varphi]$ is true' explicates what 'information' is. But we have not considered the 'provide' part. To provide is to provide someone with something. A more accurate description of 'saying that φ ' is thus 'telling someone that $[\varphi]$ is true'. This means sentences must be 'externalized', i.e. made perceptible by the senses.

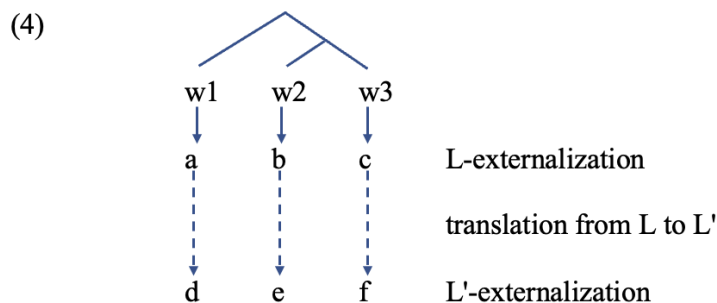
3.1.

Im Satz drückt sich der Gedanke sinnlich wahrnehmbar aus. [In a proposition the thought is expressed perceptibly.]

It is externalization, I will argue, which gives rise to cross-linguistic variation. Take sentence (1a), reproduced in (3a) and (3b), for example. It would be possible to externalize w1, w2, w3 as a, b, c, as in (3a), or as d, e, f, as in (3b). If we add this layer of analysis to the sentence, the sentence turns into two sentences in two different languages.



Thus, languages are identical up to externalization. As far as structure is concerned, there is only one language, because there is only one logical space. Words are structurally arranged in the same way in all languages, because they have the same forms with the objects and hence with their counterparts on all other languages. It is only when words are assigned auditory or visual shapes that cross-linguistic differences emerge. We now have a straightforward way to understand Wittgenstein's remark. Translation from a language L to another language L', under this perspective, would just be replacing the L-externalization of each word with its L'-externalization.



The dictionary, then, would in fact be a list of pairs $\langle x, y \rangle$, but x and y are two different ways to externalize the same word. Saying that the word x in L is translated as the word y in L' would mean saying that the word which is externalized as x in L is externalized as y in L'.

But that, of course, is not how translation actually works. Translation in real life is very different from its description in 4.025. This is a puzzle, because that description follows from assumptions that seem obvious: (i) words stand for

things and they are put together in a sentence to represent how the things they stand for are arranged in a situation; (ii) sentences are used to convey information and must therefore be externalized, as we cannot read each other's minds; and (iii) externalization may vary across different speech communities, giving rise to different languages. The question, then, is what makes translation in real life so different from translation as described in 4.025.

The answer I want to defend is this: externalization in natural language is much more complicated and chaotic than how it is presented in (3) above. The scenario in (3) is one where each language maps each basic building block to exactly one perceptible sign. But that is completely unrealistic. Consider the English sentence in (5), for example.

(5) he thought his mom would call him

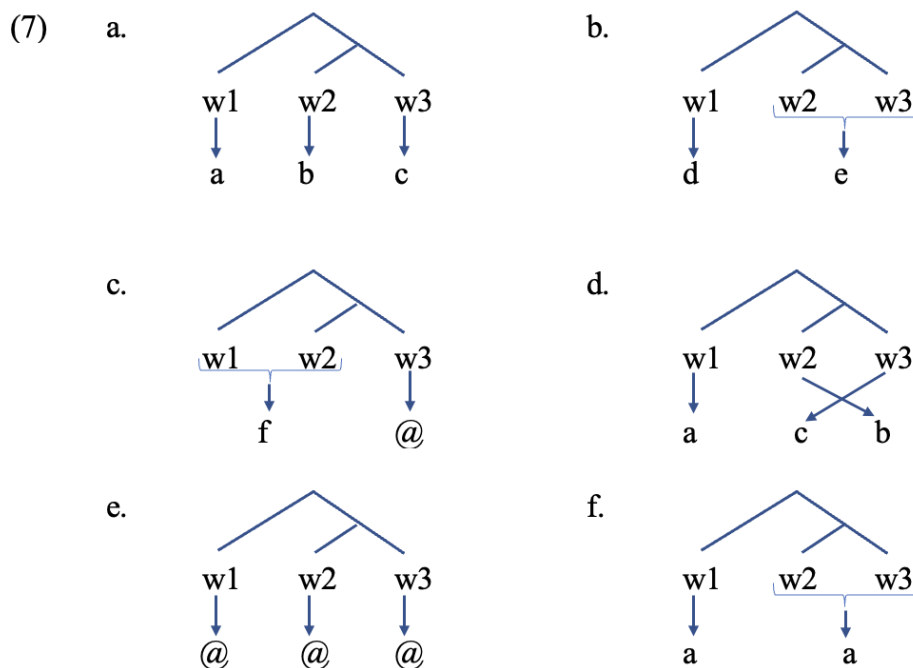
We would agree, I hope, that this sentence consists of seven words. But note that the word 'word' in ordinary language is not to be identified with the term 'word' as it was introduced in the last section. A 'word' in ordinary language is almost never a basic building block of language. We know, for example, that the words [he], [his], and [him] have some but not all things in common. That means that none of the three is basic, because basic entities do not have subparts and hence cannot have some but not all things in common. Linguists will say that [he], [his] and [him] are all [pronominal], [singular], [third person], while [he] is [nominative], [his] is [genitive], [him] is [accusative]. The same holds for the word [thought]: it is really the fusion of [PAST] and [think]. Now, it turns out that in Vietnamese, [singular] and [third person] are also externalized as one word, but the distinction between [nominative], [genitive] and [accusative] is not realized phonologically. We may say that these building blocks are all mapped to @ in Vietnamese, where @ is the 'null formative', i.e. one that has no phonological consequence at all. Moreover, Vietnamese differs from English in that it spells out [PAST] as an independent word.

(6) nó đã nghĩ mẹ nó sẽ gọi nó (he PAST think mother his would call him)

There is also, as the attentive reader may have noticed, the issue of 'linearization': mapping a hierarchical structure into a linear structure. Semantics cares about constituency: which elements make up one unit and

which ones do not. It does not care about linear order. A thought does not take up space or time. However, a thought made perceptible to the senses does take up space or time. It so happens that the primary modality of language is sound, or more specifically, temporally extended sequences of phonemes. Suppose we have a unit with two elements, a and b. One language might map this unit into the string [a b], while another might map it to [b a]. This is the case with English [his mom] and Vietnamese [mẹ nó]. English places the noun after the possessive, while Vietnamese does the opposite.

We now see that a structure of basic building blocks such as (1a) may undergo externalizations which differ in more ways than indicated by (3). Some of these other ways are presented schematically in (7).



Thus, one and the same proposition may be expressed differently across various languages as [a b c], [d e f], [d e], [f], [a c b], etc. Some language may not externalize w1, w2, or w3 at all, making (1a) inexpressible in it. This is the case of (7e). It is also possible that different words, or combination of words, receive the same externalization, leading to 'ambiguity', i.e. different meanings being associated with the same sound. This possibility is illustrated by (7f). Wittgenstein discussed ambiguity in 3.323.

3.323

In der Umgangssprache kommt es ungemein häufig vor, dass dasselbe Wort auf verschiedene Art und Weise bezeichnet – also verschiedenen Symbolen angehört –, oder, dass zwei Wörter, die auf verschiedene Art und Weise bezeichnen, äußerlich in der gleichen Weise im Satz angewandt werden. So erscheint das Wort "ist" als Kopula, als Gleichheitszeichen und als Ausdruck der Existenz [...]. [In everyday language it occurs extremely often that the same word signifies in different ways – that is, belongs to different symbols – or that two words, which signify in different ways, are applied in a proposition in ostensibly the same way. Thus the word "is" appears as a copula, as an identity sign, and as an expression of existence [...].]

There are, in principle, many more ways of externalizing (1a). I should also note here that the 'building blocks' which we have mentioned above, e.g. [PAST] or [singular], are not really basic. We can easily entertain analyses of these concepts which break them down into more basic ones. Suppose we succeeded in arriving at the final analysis of (6), i.e. its maximally articulated logical form, we would probably be looking at an enormously complex structure of ultimate basic building blocks. I submit that it is these ultimate basic building blocks, the 'elementary particles' of the language of thought, which Wittgenstein call 'simple signs' or 'words'. These 'names' will be atomic and hence have nothing in common with each other. They are far removed from the syntactic and semantic elements of natural language with which we are familiar. Natural language is an instrument that emerges from the need for communication among social beings which know that they are extremely similar in their conception of reality. When I speak, I am conveying the logical forms of the propositions which I claim to be true. And since I know that my hearer, and know that he knows, that we share the same logical space and the same strategy of externalization, I can compress the gigantic logical form to be conveyed into a few bits of sound and rely on his ability to reverse engineer this radically impoverished output back to what I am trying to get across.

4.002

[...] Die stillschweigenden Abmachungen zum Verständnis der Umgangssprache sind enorm kompliziert [...]. [The tacit agreements underlying the understanding of everyday language are enormously complicated.]

Given the wide range and numerous dimensions of variation among speech communities with respect to how configurations of basic building blocks of the language of thought are made perceptible to the senses, it is clear why translation cannot work as Wittgenstein described in 4.025. Even if the translator has a perfect grasp on the meaning of the original sentence, it is rarely, most likely never, the case that the relevant logical form has an externalization in the target language which conveys it as well as the externalization it receives in the original language. How a speech community decides which (combinations of) basic building blocks should be externalized in what way depends so much on historical accidents and chance events as to practically preclude such scenarios as (4). We should keep in mind that the *Tractatus* seeks to reveal the conditions for the possibility of perfect symbolic representation and communication. The description in 4.025 should therefore be considered one of an 'ideal' case, where logical forms are completely explicit and no reliance on 'stillschweigende Abmachungen' is presupposed.

References

- Anscombe, Gertrude E. M (1959) *An Introduction to Wittgenstein's Tractatus*, Philadelphia: University of Pennsylvania Press.
- Conant, James (1989) "Must we show what we cannot say?", In: Richard Flemming and Michael Payne (eds.) *The Sense of Stanley Cavell*, Lewisburg: Bucknell University Press, 242–283.
- Diamond, Cora (1991) *The Realistic Spirit*, Cambridge: MIT Press.
- Hintikka, Jaakko (2000) *On Wittgenstein*, Belmont: Wadsworth.
- Wittgenstein, Ludwig (2023) *Tractatus Logico-Philosophicus. Translated by Michael Beaney*, Oxford: Oxford University Press.

Acknowledgement

This work has been financially supported by the Slovenian Research Agency (ARIS) project no. J6-4615.

SSK and Feminist Epistemology as Adverse Allies - Meaning Finitism as a Resource for Debates on Hermeneutical Oppression

Sophie Juliane Veigl, Sonja Riegler (Vienna, Austria)

Abstract

This paper explores the often-contentious relationship between feminist epistemology and the sociology of scientific knowledge (SSK), two fields that are united in their rejection of the supposed neutrality of knowers present in traditional epistemologies. Despite these commonalities, feminist epistemology has largely dismissed SSK due to its perceived alignment with epistemic relativism and its historical neglect of marginalized knowers. We argue, however, that SSK, particularly meaning finitism (Barnes, Bloor and Henry, 1996) can offer valuable insights into processes of epistemic oppression, particularly through its Wittgensteinian approach to rule-following. Our focus will be on instances of hermeneutical oppression, where marginalized individuals face barriers in communicating their experiences due to gaps and power imbalances in collective interpretive resources. By integrating SSK's concept of meaning-finitism, we propose a framework that bridges structural and individual dimensions of hermeneutical oppression. Ultimately, we suggest that feminist epistemology and SSK, rather than being adversaries, have the potential to enrich each other.

1. Introduction

The relationship between feminist epistemology and the sociology of scientific knowledge (SSK) is difficult. Both schools of thought have emerged at a similar time and are characterized by a rejection of epistemological individualism and the identification of traditional epistemologies as ideological (Rouse, 1996). However, there are no major strands of influence between both programs. For instance, gender played almost no role in key SSK case studies and methodological reflections (354). Also, the type of detachment required for sociological approaches toward knowledge claims is absent from engaged and transformative feminist epistemologies and philosophies of science (364). Furthermore, most feminist epistemologies explicitly reject the sociology of scientific knowledge because of its explicit allegiance with epistemic relativism.

But is the relationship between SSK and feminist epistemology necessarily adverse or is the opposition instead a product of contingent emphases and interpretations of both programmes? In this paper, we will argue that there is hope for a more constructive relationship between feminist epistemology and SSK. While SSK scholars have historically not focused on marginalized knowers, we submit that SSK approaches can help to elucidate central

processes at play in discriminatory injustices against marginalized knowers, a central area of interest for feminist epistemologists. In particular, we will argue that SSK's Wittgensteinian approach towards rule-following helps elucidate the dynamics of certain instances of epistemic oppression. We shall illustrate our claim by providing an SSK-interpretation of the phenomenon of hermeneutical oppression (Fricker, 2007, Dotson, 2012, 2014a).

In Section Two, we briefly introduce the concept of hermeneutical oppression and situate it within a tension between structures and individuals. In Section Three, we propose meaning finitism, a central tenet of SSK, to alleviate this tension. In Section Four, we reconsider feminist epistemology and SSK as "adverse allies" and argue for the potential of both approaches for each other.

2. Nothing but Structure? Problems of Hermeneutical Oppression

Debates in feminist epistemology have long focused on power imbalances in knowledge acquisition and dissemination. In recent decades, these discussions have gained renewed momentum and attention through the concepts of "epistemic injustice" (Fricker, 2007) and "epistemic oppression" (Dotson, 2012, 2014). Discussions on "hermeneutical oppression" are central to contemporary theories of epistemic injustice. Marginalized individuals frequently face gaps in collective epistemic resources or discover that these resources inadequately represent their lived experiences. Miranda Fricker's concept of "hermeneutical injustice" (2007) illustrates this issue, describing situations where the absence of necessary concepts keeps individuals or groups from recognizing and articulating the injustices they endure (Fricker, 2007, 6). In essence, hermeneutical injustice occurs when "a significant area of one's social experience is obscured from collective understanding due to persistent and widespread hermeneutical marginalization" (154). Consider terms like "white privilege," "date rape," or "sexual harassment." Before these concepts entered public discourse, those affected by these injustices struggled to communicate their experiences to others, including friends, family, or institutional representatives.

Hermeneutical oppression exemplifies a primarily structural epistemic exclusion. Fricker argues that this type of injustice arises from gaps in collective hermeneutical resources rather than from deliberate actions by specific individuals (Fricker, 2016, 172). And yet, can we so easily dismiss the

role of individuals in perpetuating hermeneutical oppression? Several scholars have critiqued Fricker's interpretation, asserting that she too readily assumes that marginalized individuals lack the hermeneutical resources to comprehend their social realities — a view that risks undermining the agency of those experiencing epistemic oppression (Mason, 2011, 2021; Pohlhaus, 2012; Medina, 2013, 2017; Dotson, 2014). In response to these critiques, Fricker revised her concept in 2016, introducing a distinction between “minimal” and “maximal” cases of hermeneutical oppression. The former describes situations where individuals lack the means to fully understand their own experiences, while the latter refers to cases where individuals grasp their experiences but struggle to communicate them due to a lack of uptake and prevailing forms of ignorance in others (Fricker, 2016, 165). Dotson describes a similar dynamic via her concept of “contributory injustice” (Dotson, 2012). Contributory injustice arises when individuals are denied the opportunity to contribute to the pool of shared knowledge due to structural prejudices and ignorance.

The distinction between minimal and maximal cases of hermeneutical oppression as well as Dotson's concept of “contributory injustice” marks a significant conceptual and political refinement in understanding the active role dominantly situated individuals play in perpetrating hermeneutical marginalization. Individuals actively participate in epistemic systems by choosing which concepts to adopt, reject, or overlook. This agency suggests that hermeneutical oppression is not merely a passive, structural issue but also involves active contributions from individuals within these systems. It is thus a misconception that hermeneutical and contributory injustices lack identifiable perpetrators. Forms of hermeneutical oppression are both structural and agential. They stem not only from flawed or entrenched epistemic systems (Dotson, 2014a), but also from power imbalances among agents, which shape the use and interpretation of concepts. This analysis broadens the framework of hermeneutical oppression by emphasizing the role of power dynamics in determining which meanings prevail and whose experiences and conceptual resources are acknowledged.

3. Concept Use and Hermeneutical Oppression

Nevertheless, more needs to be said about how agents actively shape (dominant) epistemic systems and wield the power to determine whose experiences gain uptake and become part of the shared pool of conceptual

resources. We propose that insights from SSK, particularly meaning-finitist social theory grounded in Wittgenstein's argument on rule-following, can offer valuable perspectives on this issue.

Barry Barnes, David Bloor, and John Henry describe "finitism" as a social theory that views classification as a human activity in science as well as in every-day life (1996). This theory aligns closely with Wittgenstein's argument about rule-following, which posits that rules emerge from their applications rather than determining them (ÜG, 1969). Furthermore, rules continuously extend to new cases, making this extension an open-ended and ongoing process.

What is more, processes of classification always involve social elements. First, classification is a human practice in which relationships between words and the world are constantly reestablished and renegotiated among individuals, leaving room for variation. In essence, each person is exposed to a different set of examples, shaping their classificatory understanding differently. Second, classification entails a degree of authority, as others influence and dictate what is considered an appropriate classification.

Finitist social theory thus highlights both the provisional, revisable nature of classifications and their social and ideological dimensions. Any classification system can be deemed incorrect and subsequently revised (Barnes, Bloor & Henry, 1996, 57). (Epistemic) systems are fluid, evolving with the individuals who use and shape them, always situated within a particular social context. It is, however, also crucial to recognize that classification is inherently tied to ideology and authority. Ideology and authority determine which classifications persist over time. To wit, epistemic power not only governs the rules of knowledge production but also shapes the meanings of concepts and the dominant epistemic frameworks available. Terms and concepts do not exist in isolation from their social applications; rather, they are deeply intertwined with the individuals who use them. In other words, there is a dynamic relationship between language and its usage, shaped by both social practices and authoritative structures. We submit that this perspective is particularly illuminating when examining the tension between structural and individual agency in discussions of hermeneutical injustice.

Let us illustrate this dynamic by means of an example - the concept of “family.” For many years, the concept of “family” has been closely linked to the traditional nuclear model—a heteropatriarchal unit consisting of a man, a woman, and their children. However, various other forms of kinship, such as queer relationships, single-parent households, friendships, blended families, non-monogamous partnerships, and other intimate connections, have always been present. Groups that don't fit the dominant model have long used the term “family” to define their relationships. Still, mainstream interpretations of “family” often overlook or marginalize these alternative definitions.

In recent years, feminist and queer activists, along with scholars, have advocated for a broader, more inclusive definition of “family”—one that acknowledges a variety of household structures, kinship ties, and personal connections. As Sally Haslanger and Jennifer Saul (2006) argue, the concept of “family” should extend beyond biological kinship ties and instead focus on the broader social context in which terms like “parenthood” and “family” operate (101). Instead of limiting the definition of “parents” to biological connections, they advocate for recognizing a wider range of caregiving relationships as equally valid (*ibid.*).

Despite these efforts, alternative, queer, and non-dominant conceptions of “family” and “parenthood” continue to be regarded as inferior or illegitimate. Instead, the heteronormative two-parent model by and large still shapes societal norms, mainstream language-use and legal frameworks. Moreover, power dynamics do not just influence how oppressive structures—such as heteronormativity—affect the interpretation of concepts; they also regulate material interactions between different communities. For instance, individuals whose lived experiences align with alternative, resistant meanings of “family” are often denied legal or administrative benefits associated with the dominant interpretation of the term. For example, Vienna, the capital of Austria, is internationally renowned for its inclusive social housing system, which has been successful in creating affordable living conditions for all. However, the administrative framework behind this system still operates with conservative notions of family ties. Under Vienna's municipal housing policies, the right to inherit a flat is granted to biological family members only, excluding e.g., step-siblings, partners or foster children. This illustrates how the dominant definition of the concept of “family” can limit its application, hindering the

recognition of diverse experiences and relationships. Ultimately, rigid interpretations of “family” reinforce exclusionary norms and narrow our understanding of kinship in ways that fail to capture diverse lived realities.

The different meanings of the concept of “family” illustrate why Fricker’s initial model of hermeneutical injustice is insufficient. “Resistant” interpretations of concepts used within (marginalized) communities are fairly common, suggesting that structures are not as constraining as Fricker originally proposed. This calls for a more nuanced distinction between “maximal” and “minimal” hermeneutical injustice. We argue that many instances of hermeneutical injustice do not stem from a hermeneutical “gap”—a situation in which no concept exists to articulate a particular experience. Instead, as in the case of “family,” the issue arises when an existing concept is dominated by a singular, authoritative interpretation that overshadows alternative, community-generated meanings and appropriations of specific terms. The problem, therefore, is not the absence of a particular concept but a discontinuity between competing interpretations and meanings of the concept.

To analyze this situation, we turn, again, to meaning finitism. The dominant understanding of “family” is shaped not only by prior examples, analogical reasoning, and historical applications but also by deeply embedded social intuitions about related concepts. For example, mainstream interpretations of “family” are tied to heteropatriarchal gender roles about family structures and reproductive labor as well as to binary notions of “male” and “female.” These ideas shape societal structures, reinforcing existing hierarchies of political, social, and epistemic power.

From a meaning finitist perspective, different communities engage with distinct sets of exemplars and prior applications of the concept. However, power relations complicate this dynamic. The marginalized understanding of “family” does not exist in isolation from the dominant one. Those who adopt a non-mainstream interpretation constantly encounter the dominant meaning—through media, institutional interactions, and public discourse. They thus must be “biconceptual,” proficient in both their community’s understanding and the dominant one, to navigate various aspects of their public and private lives. Drawing on feminist standpoint theory, the proficiency in navigating diverse conceptual frameworks can be seen as an epistemic advantage. Marginalized

groups, through “double consciousness” —an awareness of both the dominant system and their own lived realities— gain a deeper understanding of social structures and develop concepts to describe and critique them (Du Bois, 1903, Hill Collins, 2000). However, this epistemic advantage remains limited as long as alternative, community-generated meanings, such as those of “family,” are not recognized by the broader public. For instance, individuals who adhere to the dominant definitions of concepts face no such necessity of being proficient in diverse concept use; they can remain unaware of alternative interpretations. Epistemic power, therefore, determines which communities must be conceptually flexible and which can remain insulated within their conceptual framework. The marginalized meaning of “family” is simply not part of the dominant group's accepted set of exemplars. Those with less power must continuously engage with mainstream interpretations to function within society. While patterns of concept use are deeply conditioned by structural oppression, they are also shaped by individual agency, that is, by the ways in which people negotiate meaning in real-world interactions. However, structural conditions ultimately dictate which concept users must adapt and which can remain rigid in their understanding.

It is crucial to recognize that variability within a community of concept users is shaped not only by meaning finitist parameters—such as exposure to exemplars, past applications, and situational contingencies—but also by broader social and political power dynamics. The existence of “enclaves,” where different patterns of concept use develop, further deepens variations in meaning while also reinforcing power relations. These communities of users do not exist in isolation; communication persists between them, largely because marginalized groups remain dependent on dominant institutions and structures. It is here, where we see much promise in thinking SSK and feminist epistemology together.

4. Adverse Allies?

We now turn to reconsider the relationship between feminist epistemology and SSK, arguing that as “adverse allies,” these research programs have the potential to offer valuable insights to one another. Our goal was to especially establish meaning-finitist social theory as one potential resource for feminist-epistemological theorizing, particularly regarding meaning dominance, power, and hermeneutical oppression.

Though SSK and feminist epistemology share a critique of traditional epistemology and philosophy of science, they have rarely been brought together (Rouse, 1996). Our aim was to show how SSK can enrich feminist epistemology, particularly when it comes to refining the concept of “hermeneutical oppression.” Feminist epistemologies often rely on implicit assumptions about the social world. For example, Dotson highlights the structural dimensions of epistemic oppression, arguing that flawed epistemic systems — not just individual biases — lead to contributory injustices. In essence, structural prejudices and ignorance prevent individuals and groups from contributing to the collective pool of knowledge (Dotson, 2012, 2014a). Dotson emphasizes the need for a deeper examination of epistemic systems to understand their role in sustaining structural epistemic oppression. While we in principle align with this perspective, Dotson’s notion of “epistemic systems” remains broad and ambiguous, encompassing everything from communicative practices and conceptual tools to entire cognitive and epistemic environments. Moreover, we submitted that more must be said about the relation between agential and structural levels of hermeneutical oppression. For example, hermeneutical oppression represents a primarily structural form of epistemic exclusion, stemming from gaps in collective hermeneutical resources (Fricker, 2016, 172). Yet, can we so readily overlook the role of individuals in sustaining this sort of oppression?

We suggested meaning-finitist social theory as a valuable tool for making the implicit assumptions about the social world in feminist-epistemological theorizing more explicit. Finitism emphasizes classification as a human activity, highlighting that epistemic systems are not merely abstract entities but are crucially shaped by how agents deploy (or ignore) concepts. As such, epistemic systems are fluid and always embedded in specific social contexts, evolving with the individuals who shape these systems through concept use. We argued that this perspective better explains instances of hermeneutical oppression, especially when they do not result from a lack of concepts, but come about through dominant meanings that overshadow alternative or resistant uses of concepts. Adopting such a perspective alleviates a tension between individual and structural forms of hermeneutical oppression, pinpointing how individual concept users shape collectively shared epistemic resources and systems.

Why, for example, has the dominant definition of “family” endured for so long, despite numerous activist and theoretical efforts to expand its meaning to include queer and alternative kinship relations? Finitist social theory can be illuminating in this regard. It helps us understand that our focus should not only be on conceptual change—such as the introduction of new concepts that reflect diverse experiences—but also on why dominant concepts persist. In other words, explaining the stability of epistemic systems is as crucial as explaining their capacity for change. Hence, a key insight of our paper is that epistemic power not only dictates the rules of knowledge production but also influences the meanings of concepts and the dominant epistemic frameworks that are available.

Needless to say, in this paper, we have only explored one potential way in which SSK and social theory can enhance feminist-epistemological discussions on epistemic power and oppression. Additionally, considering productive connections between these two research programs must also address how feminist epistemology can enrich SSK by explicitly focusing on power relations and marginalized knowers. Future research in this area thus holds great potential for discovery. Our goal, for now, was to demonstrate how SSK and feminist epistemology can be productively integrated through meaning-finitist social theory, shedding light on their shared focus on de-idealizing epistemic processes.

References

- Barnes, Barry, Bloor, David, & Henry, John (1996) "Scientific Knowledge" University of Chicago Press.
- Dotson, Kristie (2012) "A Cautionary Tale: On Limiting Epistemic Oppression" *Frontiers: A Journal of Women Studies* 33(1), 24–74.
- Dotson, Kristie (2014) "Conceptualizing Epistemic Oppression" *Social Epistemology* 28(2), 115–138.
- Du Bois, William (1903) "The Souls of Black Folk" Chicago: AC McClurg & Co.
- Fricker, Miranda (2007) "Epistemic injustice: Power and the ethics of knowing" Oxford University
- Fricker, Miranda (2016) "Epistemic Injustice and the Preservation of Ignorance." In R. Peels & M. Blaauw (Eds.), *The Epistemic Dimension of Ignorance* (pp. 144–159). Cambridge University Press.

Haslanger, Sally & Saul, Jennifer (2006) "Philosophical Analysis and Social Kinds." *Proceedings of the Aristotelian Society* 80, 89–143.

Hill-Collins, Patricia (2000) "Black Feminist Thought" Routledge.

Mason, Rebecca (2011) "Two Kinds of Unknowing" *Hypatia*, 26(2), 294–307.

Medina, Jose (2013) "The Epistemology of Resistance" Oxford University Press.

Pohlhaus, Gaile (2012) "Relational Knowing and Epistemic Injustice—Towards a Theory of Wilful Hermeneutical Ignorance" *Hypatia* 27(4), 715–735.

Rouse, Joseph (1996) "Feminism and the Social Construction of Scientific Knowledge" In S. Harding (Ed.), *The Feminist Standpoint Reader* (353-375). Routledge.

Wittgenstein, Ludwig (1969) "On Certainty" Oxford: Blackwell

The Present State of Hermeneutic Injustice Regarding Sexual Harassment and Sexual Violence

Marcia Villanueva (Mexico City, Mexico)

Abstract

This paper aims to offer a contemporary evaluation of hermeneutic injustices associated with the concepts of sexual harassment and sexual violence by revising the past fifty years of history within feminist discussions on these topics. The analysis begins by describing the story of how the term "sexual harassment" was coined in the early 1970s, based on the story of Carmita Wood, this being the central case that Fricker uses to characterize hermeneutic injustice. The subsequent section provides a concise historical account of the conceptual shifts in sexual harassment and sexual violence from their initial formulation to the present. The central argument of the paper is then developed by examining select revisions of Fricker's theory, which illuminate contemporary hermeneutic injustices related to sexual harassment and sexual violence. The paper's conclusion posits that Fricker's paradigm of hermeneutic injustice has become obsolete, underscoring the imperative to take into account the historical evolution of feminism when discussing sexual violence and epistemic injustice in the contemporary context.

In the past five decades, the collective social understanding of sexual violence has significantly changed. Miranda Fricker's (2007) seminal account of hermeneutical injustice highlighted an initial gap or lacuna in shared hermeneutic resources that constrained women's ability to understand their own experiences of sexual violence in the workplace during the early 1970s. Following the conceptualization of sexual harassment, a notable shift occurred from this lack of hermeneutic resources to a pervasive denunciation of cases of sexual violence, as evidenced by digital movements like #MeToo.

The heightened awareness of the diverse manifestations of sexual violence experienced by women globally signifies a substantial achievement in contemporary feminism. Concurrently, this evolution in the collective social understanding of sexual violence gives rise to novel inquiries concerning the epistemic dimensions of the oppression of women and other subordinated groups. Hence, the present paper aims to offer a contemporary evaluation of hermeneutic injustices associated with the concepts of sexual harassment and sexual violence by revisiting the past fifty years of history within feminist discussions on these topics.

1. Sexual harassment as the central case of hermeneutic injustice

Fricker's key example of hermeneutic injustice comes from Susan Brownmiller's account of how the phenomenon of sexual harassment appeared on the feminist agenda in the early 1970s. In her memoir, Brownmiller (1999) describes the case of Carmita Wood, a woman who had worked in the Nuclear Physics Department at Cornell University for eight years and quit her job after experiencing persistent and unwanted sexual advances from a professor. These encounters caused Wood increased stress and anxiety, as well as a variety of physical ailments, such as chronic back pain. When Wood applied for unemployment benefits, she was asked to explain why she had resigned. She answered that her reasons had been "personal" because she found herself at a loss for words to articulate her experience. She was denied insurance.

Fricker (2007) recounts this story to show how a gap in the available stock of hermeneutical resources engenders a particular kind of epistemic injustice: hermeneutic injustice. This stands for the injustice of having some significant area of one's social experience obscured from collective understanding owing to marginalization in the social methods used to determine hermeneutical resources. In this central case, Fricker underscores that Wood lacked the interpretive resources necessary to identify her experience as sexual harassment because this concept didn't exist at that time, since women had been marginalized from determining hermeneutical resources for workplace regulations. Therefore, Wood suffered hermeneutic injustice.

Fricker contends that hermeneutic injustice results in primary and secondary harms. Firstly, it renders unintelligible a distinctive social experience for the victim and, secondly, this unintelligibility results in downstream negative consequences such as her being denied unemployment insurance. Therefore, filling the hermeneutical gap could prevent unintelligibility and its consequences. Wood's case also serves to illustrate this last point in Fricker's argument. As Brownmiller's account details, Wood's disclosure of her history to a group of women led to the realization that each woman present in the meeting had experienced a comparable event, despite the absence of prior dialogue on the subject. This revelatory moment served as the impetus for the conceptualization of sexual harassment, remedying the hermeneutical injustice suffered by Wood and her contemporaries.

2. From Carmita Wood to Alysa Milano: Fifty Years of Feminist History

Hermeneutical resources are not constructed at once, but rather over time. It is therefore worthwhile to review the feminist struggle against sexual violence since the term *sexual harassment* was coined in the 1970s based on the case of Carmita Wood to the present day. A foundational moment in this history was the publication of two books: *Sexual Harassment of Working Women* by Catharine MacKinnon (1979) and *Woman Hating: A Radical Look at Sexuality* by Andrea Dworkin (1974). These seminal works initiated the paradigm of sex-negative feminism (also known as dominance feminism), which maintains that sexuality is the cause of women's oppression and that male domination rests on the power of men to treat women as sexual objects.

Both MacKinnon and Dworkin focused on the ways that sexuality could be used to harm and concluded that unequal power dynamics are the backdrop to heterosexual sex, so true consent is impossible to achieve. Initially, MacKinnon problematized the ability of working women to refuse the sexual advances of men in positions of power. Soon after, however, she expanded her argument about sexual harassment as an institutional phenomenon to a broader thesis that all heterosexual sexual relations occur on the same terms: men dominate, and women are dominated. For Judith Butler, this evolution in MacKinnon's position constituted "a tragic mistake" because it equated heterosexual intercourse with violence and placed women in the position of victim forever. In the same vein, Dworkin characterized men as if domination were their sole objective and their only object of sexual desire. She went as far as to argue for the significance of abstaining from penetrative sex, positing that heterosexuality is a fundamental aspect of male domination over female partners (cf. Lamas, 2018; Cheng, 2020; Serra, 2024).

Sex-negative feminism has had a significant impact on the collective social understanding of sexual violence. In 1993, it was documented that Princeton students wrote signs on the walls of library bathrooms that read "sex is death" and "sex is rape" (Lamas, 2018, p. 62). This position has led to the extremism of considering any sexual request as sexual harassment and making sexual assault indistinguishable from other non-criminal behaviors.

In contrast to this stance, the sex-positive paradigm focuses on the pleasure of sex and defends women's autonomy and agency to consent and conquer

spaces within sexuality, even in the contexts of misogyny and sexism in which we live. It challenges the prevailing assumption that sexual harassment constitutes an infringement on women's dignity. Proponents of this paradigm contend that the sex-negative approach fosters a form of sexual conservatism that is detrimental to women. This perspective criticizes dominance feminism for being rooted in the puritanical belief that a woman is dishonored if her sexual purity is tainted, but a man is not. In contrast, this approach asserts that victimhood is not an inherent aspect of the female condition, sexual purity is not of paramount importance for women, and sexual pleasure is possible for us even under patriarchy. Therefore, sex-positive feminism defines sexual violence based on consent, while acknowledging that in some contexts this is not possible to achieve (cf. Lamas, 2018; Cheng, 2020; Chaparro, 2021; Serra, 2024).

The conflict between sex-negative and sex-positive approaches led to the emergence of the so-called Sex Wars in the 1980s. In this context, Gale Rubin (1984) published her essay, *"Thinking Sex: Notes for a Radical Theory of the Politics of Sexuality"*, in which she argued that sex is inherently political. Rubin posited that sexuality is a product of human activity and, as such, is imbued with conflicts of interest and political maneuver, both deliberate and incidental. She acknowledged that "heterosexual encounters may be sublime or disgusting, free or forced, healing or destructive, romantic or mercenary" (p. 152) and defended that we "should judge sexual acts by the way partners treat one another, the level of mutual consideration, the presence or absence of coercion, and quantity and quality of the pleasures they provide" (p. 153). Moreover, she advanced a pluralistic sexual ethics that acknowledges sexual variation (heterosexual, gay, lesbian, monogamic, polygamic, transsexual, fetish, sadomasochistic, commercial, and all variations) without resorting to criminalization or pathologization.

Sex Wars appeared to reach their conclusion with the advent of postfeminism and the ascendance of the sex-positive approach during the 1990s and early 2000s. Rosalind Gill (2017) defines postfeminism as a late 20th- and early 21st-century current that sought to dismantle feminism under the assumption that women could now enjoy the achievements of this movement, namely, voting, studying, and working. During that period, the sex-positive approach became mediated, and shortly after, postfeminism took a turn towards neoliberal

feminism, which deems women as self-sufficient, autonomous, with free choice, and responsible for their lives (Favaro, 2017). Television series such as *Sex and the City*, women's magazines like *Cosmopolitan*, and song lyrics promoted a stereotype of femininity based on an assumed sexual liberation of self-confident women, transforming the adjective "prude" into a gendered pejorative (Chen, 2020).

In the 2010s, however, the trends began to reverse again with the rise of the fourth wave of feminism. In 2015, *The New York Times Magazine* published an article titled "The Return of the Sex Wars," which discussed how to address sexual assault on campus. Two years later, Jodi Kantor and Megan Twohey published a detailed report on how Hollywood director Harvey Weinstein harassed and sexually abused women in the industry for over three decades. Subsequently, actress Alyssa Milano tweeted #MeToo, inviting women to share their stories and propelling the movement on social media. Milano used this hashtag on Twitter at noon on October 15, 2018, and within twelve hours, it had been used 200,000 times. By the following day, it had been retweeted 500,000 times. On Facebook, the hashtag garnered a significant number of users, reaching 4.7 million in the first 24 hours (Chaparro, 2021).

The viralization of the #MeToo movement precipitated a significant shift in global public discourse regarding the practices of silencing and oppressing women who came forward with allegations. This movement proved that sexual harassment and sexual abuse were commonplace around the world, and that women no longer wanted to cover it up. In addition, this movement led to a redistribution in the epistemic credibility of those who report having suffered gender-based sexual violence (Guerrero MacManus, 2023).

Nevertheless, the #MeToo movement also exposed a lack of clarity in the conceptual understanding of sexual harassment, as well as the nuances of its connection to systemic forms of gender-based violence. Digital accusations included experiences such as groping, exposure of genitals, catastrophic dates, stolen kisses, compliments, glances longer than five seconds, and innuendoes that occur in the workplace, at school, or on the street between acquaintances—especially colleagues and friends—and strangers. All these different

experiences were lumped into the same box (Chaparro, 2021). In this context, the sex-negative approach regained strength, while defenders of the sex-positive approach were “canceled” or silenced.

This brief historical account shows that, over the past fifty years, both sex-negative and sex-positive approaches have gained strength at different times. As a result, the collective social understanding of sexual violence has fluctuated between these two positions, demonstrating the dynamic relationship between hermeneutic resources and the broader social milieu in which they are embedded.

3. Sexual Harassment and Hermeneutic Injustice Revisited

As with the concept of sexual harassment, the meaning of hermeneutic injustice has changed over time. Fricker's original definition refers to a lack of collective hermeneutic resources that renders social experiences unintelligible because of hermeneutic marginalization. Many authors have pointed out that a conceptual gap or lacuna is not necessary for hermeneutical injustice.

Mason (2011) and Medina (2012) assert that frequently, marginalized groups possess the conceptual resources necessary to make their experiences intelligible; however, groups with greater hermeneutic power do not take them up. In a similar manner, Goetze (2018) poses hermeneutical dissent as the hermeneutical injustice that occurs when a subject is prevented from sharing the knowledge produced within her marginalized group because others find her testimony unintelligible. Conversely, Dular (2023) proposes the concept of hermeneutic excess to explain that groups with epistemic power generate concepts as a form of backlash toward social progress (e.g., “reverse racism”), hindering the understanding of the social experiences of the oppressed through hermeneutical domination.

On the other hand, Crerar (2016), Jackson (2019), and Falbo (2022) present cases of hermeneutical injustice due to competing interpretive resources. Crerar's argument posits that societal taboos have the potential to hinder an individual's conceptual framework, thereby impeding the ability to understand their experiences. In a similar vein, Jackson describes the case of rape myths competing with the formal concept of rape to show that frequently the former serves to obfuscate the latter. This dynamic can give rise to instances of hermeneutical injustice, even when an agent's conceptual

repertoire is perfectly adequate. Likewise, Falbo uses the term *hermeneutical clash* to describe cases in which a pair of existing concepts compete to offer understanding of a social experience. By juxtaposing the notions of “golden boy” and “rapist,” she elucidates that when the concept that aligns with backward oppressive conditions legitimizes a particular interpretation, epistemic injustice ensues through hermeneutical defeat and preemption.

Furthermore, Elzinga (2018) and Mason (2021) distinguish between failures of concept possession and failures of concept application. Elzinga contends that possessing interpretive resources entails not only having conceptual resources but also knowing how to utilize them. Mason's argument further asserts that in certain instances of hermeneutic injustice, individuals possess the necessary concepts yet fail to apply them in a manner that illuminates the significance of their social experiences. This failure, she contends, stems from a distorted collective hermeneutical resource. This phenomenon aligns with another instance of hermeneutic injustice outlined by Fricker (2007) and exemplified by the reluctance of the protagonist in Edmund White's autobiographical narrative to self-identify as homosexual, given its association with pathology and heteronormative conventions. In this context, Mason emphasizes that the collective hermeneutical resource is not merely a chaotic accumulation of words and concepts, but rather, it manifests as a structured network where concepts are interconnected by inferential relationships.

While none of these revisions to Fricker's theory directly address the hermeneutic injustices that emerged during the Sex Wars, they offer valuable insights that contribute to a more comprehensive understanding of this issue. In the remainder of this text, four of these insights will be highlighted.

First, the cited revisions theorize about forms of hermeneutic injustice that are not characterized by a conceptual gap. Such is the case of the hermeneutic injustices we encounter today in relation to sexual harassment and sexual violence. In this context, it seems that rather than lacking concepts, what we have is a controversy over the meaning of existing concepts.

Secondly, the perspective on competing interpretive resources can explain the hermeneutical clash (Falbo, 2022) between sex-negative and sex-positive approaches. Societal taboos (Crerar, 2016) about sex have the potential to

hinder our conceptual framework, thereby impeding the ability to understand our sexual experiences beyond puritanical and conservative traditions. Furthermore, the propagation of myths (Jackson, 2019) concerning sexuality, stemming from neoliberal consumerist ideologies (e.g., "if you know what you like, you'll get what you want"), serves to obfuscate the intricacies of sexual encounters and the experience of sexual pleasure. Both puritanical and neoliberal perspectives on sex align with backward oppressive conditions, so they might ensure hermeneutical defeat and preemption (Falbo, 2022), hindering the development of a more nuanced and comprehensive understanding of sexual behavior such as the pluralistic sexual ethics advanced by Rubin in the 1980s.

Thirdly, the perspective on the failure of concept application elucidates that, despite the existence of the concepts of sexual harassment and sexual violence, their pervasive application in the context of the fourth feminist wave and the #MeToo movement was, in some cases, inaccurate. The conceptual know-how (Elzinga, 2018) of sexual harassment and sexual violence was evidenced by accusations that lumped into the same box behaviors such as groping, exposure of genitals, catastrophic dates, stolen kisses, compliments, glances longer than five seconds, and innuendoes (Chaparro, 2021). This phenomenon can be attributed to the distorted structure of the collective hermeneutical resource provided by sex-negative feminism, which poses inferential relationships (Mason, 2021) between all sorts of expressions of heterosexual desire and domination and violence. This phenomenon elicited disapproval and frustration among feminists who subscribe to a sex-positive approach, leading in many cases to their silencing. Such was the case of Mexican feminist scholar Marta Lamas, who was "cancelled" in 2018 after the publication of her book *Harassment: Legitimate Complaint or Victimization?* In that essay, Lamas claims that new concepts must be created to distinguish some annoying behaviors and certain unpleasant sexual experiences from sexual crimes. This is exemplified by the recent coining of the term "compulsory sexuality," which constitutes the fourth and final insight regarding how the cited revisions to Fricker's theory illuminate the hermeneutic injustices concerning sexual violence that have emerged during the past fifty years.

The concept of compulsory sexuality was recently drawn from Adrienne Rich's (1980) conceptualization of compulsory heterosexuality as an institution that

codifies heterosexuality as natural and therefore morally superior to other sexual orientations. Likewise, compulsory sexuality codifies sexual desire, sexual expression, and sexual identification as natural for all people. It rests on assumptions such as that sex is needed for a person to be healthy and whole, as well as for intimacy and healthy romantic relationships, so it demands that people identify themselves in terms of sexuality to be legible to others (Przybylo, 2022).

In her book *ACE: What Asexuality Reveals About Desire, Society, and the Meaning of Sex*, Angela Chen (2020) presents several cases of asexual people who consented to sexual experiences until they learned about compulsory sexuality. James, an asexual programmer residing in Seattle, recounts his initial reluctance to engage in sexual activity and the contentious debates that ultimately culminated in consenting sex with his ex-girlfriend. When James gained awareness of compulsory sexuality, he reevaluated his past experiences and recognized a form of sexual violence for which his former partner cannot be accused. James attributes this violence to compulsory sexuality. Chen contends that James's experience exemplifies a hermeneutical injustice prompted by a conceptual gap. In addition, the analysis presented in this paper enables us to identify in James's ex-girlfriend a hermeneutic injustice without a conceptual gap that corresponds to Goetze's (2018) proposal of hermeneutical dissent: she was unable to comprehend James's lack of sexual desire because she belonged to a group with greater hermeneutic power that has not embraced the concept of compulsory sexuality.

In conclusion, this paper shows that Fricker's paradigmatic example of hermeneutic injustice regarding sexual harassment is outdated. To talk about sexual violence and epistemic injustice today, it is necessary to consider the history of feminism and the recent resurgence of sex-negative feminism.

References

- Brownmiller, Susan (1990) *In Our Time: Memoir of a Revolution*, New York: Dial Press
- Chaparro, Amneris (2021) "Acoso y hostigamiento sexual: una revisión conceptual a partir de#MeToo". *Géneroos*, 28(29), 243-268.
- Cheng, Angela (2020) *Ace: What asexuality reveals about desire, society, and the meaning of sex*. Beacon Press.

Crerar, Charlie (2016) "Taboo, Hermeneutical Injustice, and Expressively Free Environments". *Episteme*, 13(2), 195–207.

Dular, Nicole (2023) "One too many: Hermeneutical excess as hermeneutical injustice". *Hypatia*, 38(2), 423-438.

Dworkin, Andrea (1974) *Woman Hating: A Radical Look at Sexuality*. EP Dutton, Nueva York.

Elzinga, Benjamin (2018) "Hermeneutical Injustice and Liberatory Education". *The Southern Journal of Philosophy*, 56(1), 59-82

Falbo, Arianna (2022) "Hermeneutical Injustice: Distortion and Conceptual Aptness". *Hypatia*, 37, 343–363

Favaro, (2017) "'Just be confident girls!': Confidence Chic as Neoliberal Governmentality." In *Aesthetic Labour: Rethinking Beauty Politics in Neoliberalism*, edited by A. S. Elias, R. Gill, & C. M. Scharff, 283–99. London, UK: Palgrave Macmillan.

Fricker, Miranda (2007) *Epistemic Injustice. Power and the Ethics of Knowing*. Oxford University Press

Gill, Rosalind (2017) "The affective, cultural and psychic life of postfeminism: A postfeminist sensibility 10 years on." *European Journal of Cultural Studies* 20(6): 606–26.

Goetze, Trystan (2018) "Hermeneutical Dissent and the Species of Hermeneutical Injustice." *Hypatia*, 33(1), 73-90

Guerrero MacManus, Siobhan (2023) "Prejuicios que silencian: Injusticia Testimonial y Muerte Hermenéutica". *Revista interdisciplinaria de estudios de género de El Colegio de México*, 9, 37p.

Jackson, Debra (2018) "Me Too: Epistemic Injustice and the Struggle for Recognition". *Feminist Philosophy Quarterly*, 4(4), 20p.

Lamas, Marta (2018) *Acoso: ¿ Denuncia legítima o victimización?*. Fondo de Cultura Económica.

MacKinnon, Catharine (1979) *Sexual Harassment of Working Women*. New Haven: Yale University Press

Mason, Rebecca (2011) "Two kinds of unknowing". *Hypatia*, 26(2), 294-307.

Mason, Rebecca (2021) "Hermeneutical injustice". In *The Routledge Handbook of social and political philosophy of language* (pp. 247-258). Routledge.

Medina, Jose (2012) "Hermeneutical injustice and polyphonic contextualism: Social silences and shared hermeneutical resources". *Social Epistemology* 26 (2): 201–20.

Przybylo, Ela (2022) "Asexuality and Compulsory Sexuality", in *The Palgrave Encyclopedia of Sexuality Education*, 10p.

Rich, Adrienne (1980) "Compulsory heterosexuality and lesbian existence". *Signs: Journal of women in culture and society*, 5(4), 631-660.

Rubin, Gale (1984) "Thinking Sex: Notes for a Radical Theory of the Politics of Sexuality", available at: <https://bpb-us-e2.wpmucdn.com/sites.middlebury.edu/dist/2/3378/files/2015/01/Rubin-Thinking-Sex.pdf>

Serra, Clara (2024) *El sentido de consentir*. Anagrama.

The 'Arbitrariness' of Ethics: Pursuing Wittgenstein's Suggested Analogy to Logic and Mathematics

Vincent Vincke (Brussels, Belgium)

Abstract

Wittgenstein wrote relatively little on ethics, and his remarks on the subject are often sparse and conceptually dense. As a result, scholarship on his (meta)ethical thought faces challenges of continuity and interpretation. One of the most detailed glimpses into his later ethical thinking comes from Rush Rhees's notes, particularly a conversation dated September 12, 1945.

In this discussion, Wittgenstein reiterated his rejection of ethical enquiry as a search for “the essential nature of goodness or duty” or “the right ethics,” viewing these as misguided pursuits of moral philosophy. Yet he also maintained that rejecting this picture need not entail abandoning the idea of a moral imperative—raising important questions about how these views might coexist.

Although he did not address this tension explicitly, Wittgenstein suggested that the “arbitrariness” of adopting ethical systems could be analogized to the “arbitrariness” involved in adopting mathematical or grammatical systems. This paper explores that analogy by reconstructing what Wittgenstein may have understood as the *arbitrariness of ethics*. The account draws on his extensive writings on mathematics and logic—including *The Big Typescript* and later lectures—as well as Rhees's transcriptions of their conversations. The paper concludes by examining how, for Wittgenstein, ethics like mathematics also involves a non-arbitrary dimension rooted in human agency and commitment.

1. Introduction

Among the few available resources, Rush Rhees's (1965) article is of seminal importance for any attempt at accounting for Wittgenstein's later thoughts on ethics, as it includes numerous recollections of their exchanges on the topic. The same holds, by extension, for his (2015) relatively recent published personal notes and transcriptions of the conversations both had from 1939 to 1950. This is not solely due to Rhees's citation of these sources in the article, but also because they contain “significantly more” material than what was included (Citron 2015: 26). Additionally, their consideration is increasingly relevant, as they have received relatively little attention in the literature so far.

In these recordings we find, for instance, notes of a discussion both had on September 12, 1945 where Wittgenstein reiterated his lifelong rejection of ethical enquiry as a search for “the right ethics,” viewing this as a flawed pursuit of moral philosophy. However, contrary to his earlier remarks, Rhees's

notes *also* reveal the seeds of a more positive account of ethics and of the role and function of declarative ethical statements. Indeed, according to Wittgenstein the disavowment of philosophical prejudices need not entail abandoning the notion of a moral imperative —raising important questions about how these views might coexist. (Wittgenstein et al. 2015: 26)

Although he did not address this tension explicitly, Wittgenstein suggested that the “arbitrariness” of adopting ethical systems could be understood analogously to the “arbitrariness” of adopting mathematical and grammatical systems:

This may raise questions as to whether the adoption or recognition of a particular ethical system is arbitrary. And presumably this would have to be treated on lines analogous to those on which one would treat the question *whether mathematics or whether logic is arbitrary. {Whether grammar is arbitrary [...] }.* (my emphasis) (Wittgenstein et al. 2015: 28)

Reflecting on the sense and extent to which mathematics and logic (grammar) can be called ‘arbitrary’ was a central theme in Wittgenstein’s middle-period works and continued to influence his later thought on mathematics. His use of the term should, however, by no means be understood as an offhand suggestion that, ultimately, the sentences of mathematics and the laws of logic might be random or dispensable.

Wittgenstein, rather, introduced it (1) to discuss the sense in which mathematical and grammatical propositions, concepts and systems, despite their necessity and reliability, are *not* answerable to anything external to grammar or mathematics, and, accordingly, (2) to underline how the propositions of mathematics and grammar, serving as rules, differ fundamentally from empirical propositions.

Hence, the aim of this paper is to examine Wittgenstein's suggested, yet hitherto unexplored, analogy by reconstructing what he may have understood as the *arbitrariness of ethics*. The account draws on his writings on mathematics and logic as well as Rhees’s transcriptions of their conversations. The paper concludes by examining how, for Wittgenstein, ethics like mathematics also involves a non-arbitrary dimension rooted in human agency and commitment.

2. Ethical Systems Are Not Accountable to An External Authority

Insofar as ethical statements and systems are indeed 'arbitrary' like those of mathematics, this would first and foremost, amount to the notion that they are not directly established by or accountable to an external authority, such as an empirical reality, a putative essence or an overarching standard. Rather, they hold themselves as 'autonomous' (networks of) norms; they reflect and delimit what is grammatically possible, permissible and meaningful within a given *moral* framework in the first place. (cf. BT 2005: 185, §133; Ms 110: 133, Ms 124, 180; Ms 122, 53v; Ms 124, 11)

Characterizing ethical systems in this way not only echoes Wittgenstein's earlier depiction of ethics as "groundless" (FWN 1965: 15; WWK 1979: 115; MS 183: 75), but also highlights the grammatical or conceptual dimension of adopting—or following—such an 'autonomous' ethical system, a notion that clearly surfaces in his response to an ethical dilemma raised by Rhees during a 1942 conversation.

The problem was described as a situation in which someone realizes that they must either leave their wife or abandon their cancer research. As Wittgenstein saw it, a different attitude will be taken to this dilemma by a man who has and one who has not adopted a particular system of ethics. The man who adheres to such a system, "takes up an *ethical* attitude" according to Wittgenstein. (Rhees 1965: 22) That is, to him "[t]he question 'Should I leave her or not?' is not a problem here". (Rhees 1965: 23) For that man, who judges within a system of ethics, say 'Christian Ethics' as Wittgenstein suggests, Rhees's problem is not a problem at all: "it is absolutely clear: he has got to stick to her come what may." (Rhees 1965: 22) The man who has not committed to an ethical system, on the other hand, has an attitude that "will vary at different times" and could thus, in principle, be open to a vaster array of arguments, depending on all sorts of variables, ranging from his taste and mood to the specifics of the case at hand. (Rhees 1965: 22)

Whether we recognize a moral dilemma or puzzle as such, will depend on whether we take up an ethical attitude, i.e., on whether we consider the ethical implications of a situation, action, or decision. Something which, according to Wittgenstein, does not happen in a vacuum but within a particular moral framework and in terms of its conceptual apparatus. In the case presented by

Rhees, the ethical dilemma was never one for the committed Christian, because it was 'absolutely' clear what he had to do.

In other words, *this* moral problem only arose and only held significance *beyond* or *outside* that person's particular moral framework. That is, either in moral philosophy, from the perspective of an uncommitted individual with a flexible (ethical) stance, or within the confines of a different ethical system. However, this does not imply that his Christian friend might not have to face moral challenges or conundrums of his own: "[H]is problem is different. It is: how to make the best of this situation, what he should do in order to be a decent husband in these greatly altered circumstances, and so forth." (Rhees 1965: 23) Hence, ethical systems might be open-ended or might reveal themselves to be incomplete in light of hitherto unimagined scenarios.

To adopt new ethical rules or to transition to a different ethical system, however, is then not simply a matter of altering one's behaviour or decisions; it involves, to varying degrees, a reconfiguration of one's concepts, redefining the words, statements, and actions that both express and constitute our moral judgments. Thus, much like Wittgenstein's famous wood sellers, someone operating under different ethical rules is not necessarily making statements or performing actions that can be indiscriminately judged as false or wrong, nor does this imply that they are irrational. Instead, they may simply be participating in a differing or different kind of ethical discourse.

Hence, such potential conceptual incommensurability does not inherently prevent (alternative) moral frameworks from being understood or compared. To say that ethical systems are incommensurable simply means that they lack a *shared* standard of measurement; it serves to indicate that we should remain vigilant of assuming an overlap between their respective conceptual apparatuses that guarantees the possibility of their translatability and comparative evaluation. That is, either by believing that various ethical systems merely express "different ways of fulfilling the same imperative or of realising the same good". (Wittgenstein et al. 2015: 27) Or by succumbing to a certain conceptual chauvinism and approaching "what adherents of a different ethics are doing and saying in terms of some conception of good that we ourselves hold, and to say that this interpretation is what they really mean". (Wittgenstein et al. 2015: 27, 28)

3. Ethical Systems Are Not Right or Wrong

That ethical statements and systems, like mathematical ones, are not directly accountable to an external authority, entails that they cannot be determined to be true or false, correct or incorrect, right or wrong by virtue of any extra-systemic factors or considerations, such as the obtainment of a particular (empirical) end or result. Rather, as standards of correctness, they establish the background against which we can speak of right and wrong to begin with. (cf. BT 2005: 186; LFM 1976: 95, 97-98, 101)

As Wittgenstein indicated, the 'uncommitted' cancer researcher might still be uncertain about how he should address his moral conundrum despite the advice of his Christian friend. He might then consider approaching the situation through the prism of different ethical systems. Yet, in stark contrast with his Christian friend ("it is absolutely clear, he has got to stick to her come what may") he might now become even more perplexed ("Now it is not clear that he must stick with her. On the contrary, etc."). (Rhees 2001: 410)

In such circumstances, there might be a strong inclination to ask what 'the right ethical system' is, and to think that, perhaps, after careful examination, we might discover the correct solution to the problem or establish the superior moral framework, as "it *must* be possible to decide which is right and which is wrong." (my emphasis) (Rhees 2001: 410) Hence, one seems to be under the impression that the resolution of such a moral dilemma would be akin to determining which of two sticks is the longest, and that even though we can only perceive them "through the wavering air of a hot pavement [,] surely one of them *must* be longer." (my emphasis) (Rhees 2001: 411)

If anything, however, Wittgenstein considered the request to determine *the* correct ethical system comparable to asking, "which of two different standards of accuracy is *the* right one." (my emphasis) (Rhees 2001: 410) Establishing what *the* correct ethical system is would then be like settling what *the* correct unit of measurement is, or deciding whether a ruler marked in inches is *more accurate* than one marked in centimetres, rather than specifying which of two objects is the longest.

Whereas the latter would be perfectly sensible, the former would be nonsensical, he argued. A standard of accuracy is not correct or incorrect by virtue of an *external* relation (e.g., an empirical result). Rather, *it* stipulates

what is to count as such *by definition*, i.e., independently from experience. (cf. BT 2005: 186, LFM 1976: 42, 44, 251)

When measuring the height of a table in either centimetres or inches both results are valid, but only within their respective systems. Asking whether centimetres or inches provide the 'true' height, or which is 'more accurate' doesn't make sense because these are two different standards of measurement that do not rely on external factors to determine their correctness. Instead, each system stipulates *by definition* what counts as accurate within its own terms, independent of any empirical outcome.

Similarly, different ethical systems provide distinct frameworks for determining what is morally right or wrong. Just as it's meaningless to compare centimetres and inches as 'better' or 'truer' measures of length, it is similarly meaningless to ask which ethical system is '*the* right one' when each defines its own standards for moral judgments.

Yet, the fact that the question is posed at all highlights the difference in attitude between the man committed to Christian ethics, for whom it is 'absolutely clear' what must be done in the face of said moral dilemma, and his uncommitted friend, who considers various ethical systems in search of the right one. For the Christian, the correct solution to the problem is predetermined—it is the right solution *by definition*, stipulated 'independently of experience'; he must "stick to her *come what may*." (Rhees 2001: 410) In contrast, for the uncertain cancer researcher, what qualifies as the right solution to the ethical problem seems largely dependent on how events unfold, i.e., on an external relation:

Whatever he finally does, the way things turn out may affect his attitude. He may say, 'Well, thank God I left her; it was better all around'; or 'Thank God I stuck to her'. Or he may not be able to say 'thank God' at all, but just the opposite. I want to say that this is the solution of an ethical problem. Or rather: it is so with regard to the man who does not have an ethics. (Rhees 2001: 410)

4. Ethical Systems Are Not Accountable To But Made Important By Reality

Ethical statements and systems, like mathematical ones, are arbitrary in the sense that they cannot be justified or refuted—over any alternatives—by being

matched with certain facts in the way empirical propositions can. (cf. BT 2005: 187-188, 232; Ms 124: 77; RFM 1978: VII, §18) This is because, like grammatical propositions, they provide the structure that enables us to make ethical judgments about events in the first place. Hence, even though an ethical statement or system “seems to point to a reality outside itself, it is merely the expression of the acceptance of a new measure (of reality)” (MS 122: 52v; RFM 1978: III, §27).

However, as with other “measures of reality,” there is also a sense in which mathematical and ethical statements and systems are not arbitrary.

Mathematical and logical statements are in a sense exactly, as arbitrary as the choice of the unit of length is arbitrary [i.e.] because of the original decisions involved. But of course these decisions weren't made without regard to anything. (WCL 2017: 74)

That is to say, although ethical systems are not directly imposed on us by or accountable to “a reality”, there is a sense in which they are still “made important and justified by reality” (LFM 1976: 246). They do not emerge in a vacuum but are shaped by external factors—such as the dispositions that characterize us as human beings, the objects we interact with, and our education, culture, experiences, goals, and interests. Hence, although various alternatives mathematical and moral frameworks exist and are *logically* possible, “all sorts of facts, psychological and otherwise,” and “all sorts of reasons,” as Wittgenstein phrases it, *practically* affect and limit those suitable and available to us and the ones we choose or have come to adopt. (LFM 1976: 245, 153) Moreover, nothing about our mathematical or ethical concepts, rules, or systems determines their role, status, or application. They neither emerge in a vacuum, nor are they valuable, interesting, or ‘practical’ within one.

As such, our recognition of something as the answer to a moral quandary, or our adoption of a certain system of ethics over any alternatives will then, importantly, also be informed by our *personal* goals, interests, sensibilities. Our adherence is intimately tied to the intricacies of the lives we lead, the particularity of the situations we find ourselves in, the worries we have, and all the other commitments we make. (cf. Rhees 1965: 22, 23)

The importance of 'the personal' leads us back to the question of determining the right ethics as, from the vantage point of the uncommitted person, looking for '*the* right ethics' or '*the* right solution to an ethical problem', this might amount to saying that even though not all imaginable ethical systems are possible, in the end it would be a matter of personal 'choice' or 'preference'.

Yet, according to Wittgenstein, that need not be so:

If one says that there are various systems of ethics, one is not saying that they are all equally right. That would have no meaning. Just as it would have no meaning to say each was right from its own standpoint. That could only mean that each judges as he does. (Wittgenstein et al. 2015: 30)

In this the uncommitted person or moral philosopher might then still sense a need for external criteria. Insofar as for someone who adheres to a specific ethical system there is a 'right answer' and a 'right ethical system', the question still remains how they determined that system to be the right one as opposed to any alternatives, prior to choosing or adopting it. Making such a choice would then, in principle, require *external* or *extra-systemic* standards, and yet these would have to be *ethical* standards as well.

Wittgenstein would agree with this, as you "seem to be adopting certain ethical criteria if you can talk about the right one or the wrong one at all." (Wittgenstein et al. 2015: 29) Yet he contends that one's act of labelling a system of ethics as 'the right one' and one's act of adopting that system are intrinsically linked.

"If *I* decide that a certain ethical judgement is true — or that a certain system of ethics is the right one — then *I* am also making a judgement of value. In other words, *I* should be adopting that system of ethics, or making the same ethical judgement." (my emphasis) (Wittgenstein et al., 2015, p.29)

Validating an ethical system, thus, entails that one embraces it or that one 'should' do so. Yet the criteria one seems to adopt when making this judgement of value appear to be the same that would appeal to and hold for the person who already adheres to said system of ethics.

So what does it come to say that 'one of them must be the right one'? One for which conclusive reasons can be given? They would be conclusive reasons for

those who recognised a certain system of ethics, no doubt. You seem to be adopting certain ethical criteria if you can talk about the right one or the wrong one at all. (Wittgenstein et al. 2015: 29)

Hence, Wittgenstein does not seem to distinguish between internal and external standards or criteria when it comes to ethics; to deem certain ethical reasons, standards, or criteria to be decisive is to commit oneself to these. The main differentiator thus appears to be one's *personal* commitment to that ethical system.

The only meaningful way in which we can speak of 'the right ethics' is by *personally* committing ourselves to a particular system of ethics and it is this *personal* commitment that Wittgenstein presented to Rhees as the main (extra-systemic) and ultimate standard that makes ethics non-arbitrary, or that grounds it. That is to say, ethical systems are only arbitrary insofar as we do not take up a particular moral attitude, do not recognize certain ethical principles, criteria and reasons, problems, and solutions as being 'the right ones', and do not speak, act and judge in accordance with them.

Even though an ethical system is not accountable to or imposed on us by an *external* authority, and *even though* nothing *about* makes it intrinsically right or more right than any alternative

This does not mean that I must be indifferent, or that I must therefore think that there is no decisive reason for or against the action. I needn't cease to adhere to one system of ethics—and in this sense be indifferent—and if I do adhere then like as not I will recognise reasons which are decisive. (Wittgenstein et al. 2015: 28)

In other words, only from the vantage point of the uncommitted person, will ethics not only be arbitrary in principle, but also *in practice*.

This echoes what Wittgenstein said about the arbitrariness of mathematics in LFM:

If you compare mathematics to a game, one reason is that you want to show that in some sense it is arbitrary – which is certainly misleading and very dangerous in a way. (LFM 1976: 142-143)

However:

[M]athematics is not arbitrary, only in this sense, that it has an obvious application. Whereas chess hasn't got an obvious application in that way. That's why it is a game. (LFM 1976: 150)

A system of ethics isn't inherently correct or supreme, nor does it determine or prescribe the way and the extent to which it should be implemented in the first place; all this depends on what is 'natural' for us as human beings, i.e., our dispositions, those of the entities we engage with, our culture, our education, goals and interests, but above all, the place and function we actively and personally assign to it. Without the latter an ethical system is indeed as arbitrary as a game of chess.

References

Citron, Gabriel (Ed.) (2015) "Wittgenstein's Philosophical Conversations with Rush Rhees (1939-50): From the Notes of Rush Rhees", *Mind* 124, 1-71. DOI: doi:10.1093/mind/fzu200.

Rhees, Rush (2001) "On Religion: Notes on Four Conversations With Wittgenstein", *Faith and Philosophy* 18, 1-17.

Rhees, Rush: III (1965) "Some Developments in Wittgenstein's View of Ethics", *The Philosophical Review* 74, 23-45. DOI: 10.2307/2183528.

Waismann, Friedrich (1965) "II: Notes on Talks with Wittgenstein", *The Philosophical Review*, 74, 12-16. DOI: <https://doi.org/10.2307/2183527>.

Wittgenstein, Ludwig; Rhees, Rhees & Citron, Gabriel (2015) "Wittgenstein's Philosophical Conversations with Rush Rhees (1939-50): From the Notes of Rush Rhees", *Mind* 124 (2015), 1-71. DOI: 10.1093/mind/fzu200.

Wittgenstein, Ludwig (2005): *The Big Typescript: TS 213*. Hg. und übersetzt von C. Grant Luckhardt und Maximilian A. E. Aue. Malden (Mass.), Oxford, Carlton (Vic): Blackwell.

Wittgenstein, Ludwig (1965): "A Lecture on Ethics". In: "Wittgenstein's Lecture on Ethics". In: *The Philosophical Review* 74, S. 3-12 (Ithaca, USA, 1965).

Wittgenstein, Ludwig (1976): *Wittgenstein's Lectures on the Foundations of Mathematics: Cambridge, 1939*. Hg. von Cora Diamond. Ithaca, USA: Cornell University Press.

Wittgenstein, Ludwig (1978): *Remarks on the Foundations of Mathematics*. Hg. von G. H. von Wright, R. Rhees und G. E. M. Anscombe, übersetzt von G. E. M. Anscombe. Third edition, revised and reset. Oxford: Basil Blackwell.

Wittgenstein, Ludwig (2017): *Wittgenstein's Whewell's Court Lectures. Cambridge 1938–1941. From the Notes by Yorick Smythies*. Hg. und eingeleitet von Volker A. Munz und Bernhard Ritter. Oxford: Wiley Blackwell.

Wittgenstein, Ludwig (1979): *Ludwig Wittgenstein and the Vienna Circle. Conversations recorded by Friedrich Waismann*. Hg. von B. F. McGuinness, übersetzt von Joachim Schulte und B. F. McGuinness. Oxford: Basil Blackwell.

Acknowledgement

This research was made possible with the support of the Research Foundation – Flanders (FWO) through a PhD Fellowship in Fundamental Research (grant number 1132221N).

Power – A Surveyable Representation

Robert Vinten (Lisbon, Portugal)

Abstract

Although Wittgenstein did not concern himself much with politics his work is useful in getting clearer about the notion of power in a number of ways. In the first place, although Wittgenstein did not have much to say about power he *did* have quite a lot to say about ways of getting clearer about concepts and he made clarity or understanding a central aim in his philosophy. One of the ways in which he helped us to get clearer about concepts was to guide us away from the assumption that all of the items subsumed under a term must have something in common. Wittgenstein highlighted that some terms, such as 'game' and 'number' were family resemblance terms and it is plausible that many terms used in discussing political matters are family resemblance terms. He also reminded us of the diversity of ways in which we use language and of the fact that speech takes place within a rich weave of activities. He called these different uses of language 'language games'. In analysing power it is useful to remember that discussions of power take place in many different spheres, in religion, in politics, in meteorology, chemistry, etc., and that we speak about power in many different language games. Moreover, Wittgenstein warned us against thinking that the methods of the natural sciences can be transferred straightforwardly to the social sciences and to philosophy. One of the problems with studies of power conducted in the past is that they have tried to be 'scientific' by focussing on observable instances of the exercise of power. Finally, this paper will briefly look at remarks from *Culture & Value* where Wittgenstein does talk about power: where he distinguishes power from its bases and brings our attention to the fact that human powers are two-way powers.

Power – A Synoptic Representation

Power plays a central role in many recent philosophical discussions of injustice and resistance to it. Miranda Fricker begins her hugely influential book, *Epistemic Injustice*, with a discussion of gender power, social power, and identity power (2007, 9-17) and the role they play in epistemic injustices. Stephen D'Arcy has discussed how power imbalances warp democracy and how protest can help correct those imbalances (2013, 101-2). In José Medina's recent book, *The Epistemology of Protest*, he discusses the performative power of protest that empowers those who suffer injustices (2023, 133-5; see also Moody-Adams 2022; Delmas, 2018).

Issues of power are part of our everyday lives. We adjust our behaviour due to being aware of those who have power over us. Other people have the power to deprive us of our work, our livelihood, our liberty, and even our lives and so we tend to do what we are contractually obliged to do, follow laws, avoid violence, and pay close attention to the activities of politicians, bosses, and

armed forces (power is obviously implicated in activities like ‘doomscrolling’). We are wise to be alert to our own powers and to develop them if we want to make better lives for ourselves.

So, power is clearly an important issue but it is also a difficult one to achieve clarity about. We speak about the power of an acid to dissolve metal, the power of a storm to destroy a village, the power of politicians and CEOs to make decisions which affect the interests of millions of people, identity power, formal power, and resistant power. We speak about being *in* someone’s power, of people having power *over* us and of having the power *to* do certain things (see Pitkin 1972, 276 and Lukes 2021, 78 on this). Is there something that all of these varieties of power have in common that makes them all instances of power? How do we determine whether one person or group of people is more powerful than another? Is there any way of measuring power?

It is tempting to think that we already have sufficient command of the concept of power and that achieving clarity about a concept is less important than the empirical work involved in discovering where power lies and the empirical-philosophical work involved in thinking about its proper distribution as well as the legwork involved in bringing that about. However, work by sociologists, political theorists, and philosophers on these latter issues has been hotly disputed and these disputes are to some extent down to the different understandings of power the different theorists have. As Steven Lukes notes in his classic study of power, conceptions of power and the methods used to study its distribution are interrelated (2021, 8-10). Even if the work of studying the distribution of power, thinking about its proper distribution, and bringing about a better distribution is more important than the conceptual work it is clearly important that we first get clear about the notion of power before we go out and study it in society.

Wittgenstein and Power

Wittgenstein’s focus was not on political theory or political concepts and he had very little to say about power in his philosophical work. However, there are a number of ways in which Wittgenstein’s work can be useful. (1) Although Wittgenstein had little to say about power he did have quite a lot to say about clarifying concepts. The aim of clarity was central to Wittgenstein’s philosophizing (“For me...clarity, transparency, is an end in itself” (1998, 9e),

“the clarity we are aiming at is indeed *complete* clarity” (2009, §133). Here I will take up his suggestion that what we lack is a survey of the grammatical terrain (the grammar of ‘power’ and its relationships with other similar notions such as ‘potential’, ‘force’, ‘influence’, ‘control’ and so on) and that what we need is a surveyable representation of that grammar. (2) One of the ways in which Wittgenstein helps us to achieve clarity is by alerting us to the fact that philosophers have been led astray by a “craving for generality” and one of the manifestations of this craving for generality is “the tendency to look for something in common to all the entities which we commonly subsume under a general term” (1958, 17). Being alert to this possibility we might be open to recognizing ‘power’ as a family resemblance concept (2009, §§66-7). Indeed, Steven Lukes has suggested that this is so (2021, 66; see also Gakis, 2025, 185). (3) Lukes also suggests another possible way in which a Wittgensteinian framework might be useful in getting to grips with the concept of ‘power’. He suggests that, “different concepts of power have their place in different local ‘language games’ and this entails that the search for a single concept of power is illusory” (2021, 66-7). (4) Wittgenstein’s criticisms of scientism are relevant to criticising particular (behaviourist) conceptions of power. (5) Wittgenstein did at least have a few things to say about power that are useful. It is worth looking at his scattered remarks where he uses the word ‘power’ and reflects upon the concept and one can also look at Wittgenstein’s remarks about abilities and about dispositions in order to get clearer about the notion of power.

Power – a surveyable representation

Power is linked, etymologically, to words like ‘potency’, ‘potent’, and ‘potential’ (from the latin verb *posse*, to be able to). This close linguistic association between being able to do something and power is retained in many of the Latin languages. For example, in Portuguese the word for power is *poder* and *poder* is also the word for the verb ‘to be able to’. This means that we should be careful to separate power from its exercise. Somebody might have a power (ability or capacity) to do something even if they never exercise the power to do that thing. Power very often operates in the absence of its exercise. Miranda Fricker makes this point in terms of the *active* and *passive* operation of power and uses the example of the traffic warden, whose power to issue fines affects the behaviour of motorists (they generally do not park on

double yellow lines etc.) even when she is not issuing tickets (2007, 9-10). Her power is operative, even if not exercised. This might seem like a straightforward point but it is one that has often been ignored by philosophers. For example, Fricker criticizes Foucault for claiming that “power exists only when it is put into action” (Foucault 1982, 219). She notes that this cannot be correct “because it is incompatible with power’s being a capacity” (2007, 10). Anthony Kenny criticizes Hume for saying that “the distinction...betwixt power and the exercise of it is... without foundation” (Hume, 2006, Book 1, Part 3, Section 14, p. 115; Kenny 1975, 10). Steven Lukes criticizes Robert Dahl and Nelson Polsby for thinking that they can study power simply by examining moments when powerful people make decisions (i.e. exercises of power) (2021, 21-33; see also Morriss and Pansardi 2025, 588). In this last case we can see a connection between the methodology used in studying power and a theorist’s conception of power.

Recognising this connection between power and potential, ability or capacity, many theorists have said that power is a ‘dispositional concept’. In their recent entry on power in the *Routledge Companion to Social and Political Philosophy* Peter Morriss and Pamela Pansardi note that disagreement over this point (identification of power with its exercise) has largely dissipated and that “there is now some agreement...that power should be thought of as a *dispositional* concept” (2025, 587). This is fine if all that we mean by describing power as a ‘dispositional concept’ is that, like dispositions, powers might be said to exist or operate in the absence of being actualised, put into action, or exercised at some particular time. However, we should be wary of identifying powers with dispositions. As Peter Hacker says, “Human dispositions...are possessed only if what they are dispositions to do is, from time to time, done” whereas human powers or abilities “can be possessed but never exercised. For an ability and an opportunity do not imply action. We all have the ability to kill other humans, but, mercifully, few of us choose to exercise it” (2007, 95) (note, however, that the powers of inanimate objects *are* dispositions in some sense (Hacker 2007, 94-5)).

Is ‘power’ a family resemblance concept?

As noted above we attribute powers to acids, storms, people, and groups of people, amongst many other things (perhaps also structural operations of power (Fricker 2007, 10-11) . We also make discriminations amongst the

powers that things have, describing them as natural, formal, charismatic, resistant, passive, dyadic and so on. The powers of inanimate things are a sort of disposition but the powers of human agents are not. Sometimes we use words like ‘power’ and ‘force’ interchangeably (as Wittgenstein does when he suggests that speaking of the “force of law” is roughly equivalent to speaking of “the power of custom” (1980, §343)). On other occasions ‘force’ and ‘power’ are not interchangeable (as in the case Wittgenstein describes where it is in someone’s power to lift a weight (1998, 87e)). When we talk about the powers of inanimate and animate things to do something we can very often exchange the words ‘power’ and ‘ability’ without loss of meaning. However, in other cases we cannot do this. In speaking of ‘the power of custom’ or ‘the power of love’ we cannot exchange the word ‘power’ for ‘ability’ and retain the same meaning. One can turn on the power in one’s house but it is not clear what ‘turning on the ability’ might mean. Human agents can choose to exercise their power when the opportunity presents itself but they can also refrain from exercising their power in appropriate circumstances. Human powers are very often *two-way* powers whereas the powers of inanimate things are not (Kenny 1975, 53; Hacker 2007, 95; Lukes 2021, 76). Although, as Anthony Kenny notes, the passive powers of human beings are not two-way powers: if “someone speaks a language I know in my hearing it isn’t in my power not to understand it” (1975, 53). Polsby thought that ‘power’ and ‘influence’ were “serviceable synonyms” (1963, 3-4) but it is easy to see how the two notions could come apart. Van Gogh has been very influential but that does not mean that Van Gogh has been, is, or was, very powerful.

There is some unclarity about what it means when we ask whether somebody has a certain power. If we ask whether an American President has the power to overrule the judiciary someone might respond that formally he does not but that President Trump has nonetheless acted in defiance of court orders (as he did when he ignored a court order to turn around planes carrying deportees (Levine 2025)). The fact that he was able to deport people without due process (and that it was not a fluke that he did so) tells us that it is in Trump’s power to do so but the Constitution of the United States and its laws nonetheless tell us that it is not within Trump’s (formal) powers.

These distinctions that we make between uses of the word ‘power’ suggest that power is a family resemblance concept, like the concepts ‘game’ and

‘number’ (2009, §§66-7). In the *Philosophical Investigations* Wittgenstein noted that there is no one thing in common between the activities that we call games (“board-games, card-games, ball-games, athletic games, and so on” (2009, §66). There are “similarities overlapping and criss-crossing” (Ibid.) between the various things we call games, much like there are similarities between members of a family. When it comes to the concept of ‘power’, it may be that we want to distinguish different concepts of power (the meaning of ‘power’ in ‘turning on the power’ is quite different to the meaning of ‘power’ in ‘the power to overrule the judiciary’) but it seems that when we talk about the powers of inanimate things and animate creatures to do certain things we are using the same concept. Although the concept is the same when we talk about the power of an acid to dissolve metal and the power of the King to dissolve parliament the concept will behave differently in the different cases (in some cases we will be talking about one-way powers, in some cases two-way powers, in some cases active powers, in others passive powers, in some cases formal powers, etc.).

Power and language games

In response to logicians and philosophers of language that particularly focused on assertions (Frege and the early Wittgenstein), the later Wittgenstein pointed out that there are “countless different kinds of use of all the things we call ‘signs’, ‘words’, ‘sentences’” and Wittgenstein tries to capture this diversity in the uses of language by speaking about different “language games” (2009, §23). Wittgenstein uses the word ‘language game’ he says “to emphasize the fact that the *speaking* of language is part of an activity, or of a form of life” (Ibid.). Now, if we look at the different spheres within which we speak about power and the different kinds of utterances that the word ‘power’ appears in it is certainly plausible that ‘power’ makes an appearance within a number of different language games. A meteorologist might report that a powerful storm is approaching the mainland and that flood defences should be raised. A religious believer might enquire about and contemplate the power of God. A political commentator might discuss whether Trump has the power to ignore the rulings of judges.

As we have seen, Steven Lukes has suggested that “different concepts of power have their place in different local ‘language games’ and this entails that the search for a single concept of power is illusory” (2021, 66-7). Lukes is surely

right that there are different concepts of power (the Oxford English dictionary lists twelve different uses of the term ‘power’) and that the search for a single concept is illusory. However, the mere fact that the term ‘power’ appears in different language games is not enough to conclude that there are different concepts involved. If we follow Wittgenstein’s suggestion and take reports of events and speculation about events to be distinct language games (2009, §23) we can see that it is clear that we might report that Trump does not have the formal power to defy court orders and that we might also speculate about whether Trump has the power to defy court orders but it seems implausible that different concepts of power are in the two cases.

Wittgenstein and Scientism

Lukes’ criticism of Dahl and Polsby, criticising their focus on incidences of decision making where one party wins out over another, was briefly mentioned earlier. Dahl and Polsby were driven by a kind of scientism. Since potentialities cannot be observed they thought that the focus of studies of power must be on observable behaviour in order to be scientifically valid. Polsby asks,

“How can one tell...whether or not an actor is powerful unless some sequence of events competently observed, attests to his power? If these events take place, then the power of the actor is not ‘potential’ but actual. If these events do not occur, then what grounds have we to suppose that the actor is powerful? There appear to be no significant grounds for such a supposition.” (1963, 60).

Dahl and Polsby subscribe to a kind of behaviourism, in thinking that in order to study the distribution of power one must study the behaviour which is the actualisation (or exercise) of that power because that behaviour is observable whereas the potentiality is not.

Wittgenstein warned against this kind of scientism in *The Blue Book*, where he said that one of the sources of the ‘craving for generality’ was “our preoccupation with the method of science...the method of reducing the explanation of natural phenomena to the smallest possible number of primitive natural laws” (1958, 18). We have already seen that we have reason to be wary of reducing power to its observable exercise, because we can observe the operation of power in the altered behaviour of people even when

power is not being exercised. Moreover, we have reason to be wary of thinking that the methods of natural science will be applicable to power because examinations of political power are inevitably going to involve value judgements. We judge some people to be more powerful than others based on whether they make a significant impact on other people's interests but "the question of where people's interests lie, of what is basic or central to their lives and what is superficial, is inherently controversial" (Lukes 2021, 85-6; see also Hacker 2007, 21 and Morriss and Pansardi 2025, 588 for criticisms of reduction of power to its exercise).

Wittgenstein on power

I'll finish here with some quick remarks on what Wittgenstein has to say about power. In *Culture & Value* Wittgenstein remarks that, "Power & possession are not the *same* thing. Even though possession also gives us power" (1998, 18e). That is, it is useful to distinguish between power and the bases of that power (strength, wealth, popularity, and so on (see Lovett 2007, 710 on this)).

Elsewhere in *Culture and Value* Wittgenstein says,

"He refuses to..." means: it is in his power. And *who* wants to say that? Well, of what do we say 'it is in his power'? – We say it in cases where we want to draw a distinction. I can lift *this* weight, but I will not lift it; that weight I cannot lift" (87e)

This remark again highlights that human powers are often two-way powers. We can choose whether to exercise them when appropriate circumstances obtain. It also highlights that we use talk of power to distinguish cases where someone does not do something because they cannot and cases where someone does not do something but they could do it (they refuse or choose not to).

References

- D'Arcy, Stephen (2013) *Languages of the Unheard: Why Militant Protest is Good for Democracy*, London: Zed Books.
- Delmas, Candice (2018) *A Duty to Resist: When Disobedience Should Be Uncivil*, Oxford: Oxford University Press.

Foucault, Michel (1982) 'How is Power Exercised?', trans. Leslie Sawyer from Afterword in H. L. Dreyfus and P. Rabinow (eds.) *Michel Foucault: Beyond Structuralism and Hermeneutics*, Hemel Hempstead: Harvester Press.

Fricker, Miranda (2007) *Epistemic Injustice: Power & the Ethics of Knowing*, Oxford: Oxford University Press.

Gakis, Dimitris (2025) "Wittgenstein's 'Family Resemblances' and Their Political Potential", in Lotar Rasiński, Anat Biletzki, Leszek Koczanowicz, Alois Pichler, and Thomas Wallgren (eds) *Wittgenstein and Democratic Politics: Language, Dialogue and Political Forms of Life*, London : Routledge.

Hacker, Peter M. S. (2007) *Human Nature: The Categorical Framework*, Oxford: Blackwell.

Hume, David (2006) *A Treatise of Human Nature*, David Fate Norton and Mary J. Norton (eds.), Oxford: Oxford University Press.

Kenny, Anthony (1975) *Will, Freedom and Power*, Oxford: Basil Blackwell.

Levine, Sam (2025) "Trump's defiance of court orders is 'testing the fences' of the rule of law", *The Guardian*, 23rd March.

Lovett, Frank (2007) "Power" in Robert E. Goodin, Philip Pettit, and Thomas Pogge (eds.) *A Companion to Contemporary Political Philosophy*, 2nd ed. Vol. 1, Oxford: Blackwell, 709-718.

Lukes, Steven (2021) *Power: A Radical View*, 3rd. edition. London: Red Globe Press.

Medina, José (2023) *The Epistemology of Protest: Silencing, Epistemic Activism, and the Communicative Life of Resistance*, Oxford: Oxford University Press.

Moody-Adams, Michele (2022) *Making Space for Justice: Social Movements, Collective Imagination, and Political Hope*, New York: Columbia University Press.

Morriss, Peter and Pansardi, Pamela (2025) 'Power' in Gerald F. Gaus, Fred D'Agostino, and Ryan Muldoon (eds.) *The Routledge Companion to Social and Political Philosophy*, London: Routledge.

Pitkin, Hanna F. (1972) *Wittgenstein and Justice: On the Significance of Ludwig Wittgenstein for Social and Political Thought*, Berkeley: University of California Press.

Polsby, N. W. (1963) *Community Power and Political Theory*, New Haven CT: Yale University Press.

Wittgenstein, Ludwig (1958) *The Blue and Brown Books: Preliminary Studies for the 'Philosophical Investigations'*, Oxford: Basil Blackwell.

Wittgenstein, Ludwig (1980) *Remarks on the Philosophy of Psychology, Volume 1*, Oxford: Basil Blackwell.

Wittgenstein, Ludwig (1998) *Culture and Value: Revised Edition*, edited by G. H. von Wright, Oxford: Blackwell.

Wittgenstein, Ludwig (2009) *Philosophical Investigations* Revised 4th edition by P. M. S. Hacker and Joachim Schulte. G. E. M. Anscombe, P. M. S. Hacker and Joachim Schulte (trans.), Oxford: Wiley-Blackwell.

Coerced Illocutions: 'Don't Put Words in My Mouth'

Eirini Vryza (Cambridge, United Kingdom)

Abstract

Members of socially marginalised groups have their discursive agency stifled in different ways. It is by now a familiar idea that sometimes, they fail to speak, while other times, they speak but fail to do the things that they wanted to do with those words; they fail to perform their intended illocutions, in J.L. Austin's terminology. But there is a flipside to this. Sometimes, marginalised individuals are made to speak when they do not want to, as Rachel A. McKinney (2016) argues. And, as I argue in this paper, sometimes, they are made to illocute when they do not want to, even though they have made no intentional speech contribution at all. This is what I call a 'coerced illocution', an overlooked and distinctive way in which the discursive agency of the marginalised is curtailed. For example, a woman sitting silently in a professional meeting because she is fed up with having her contributions dismissed may be mistakenly taken to silently assent to her interlocutors' ideas. Because of her marginalised position, she may coercively play along and act as if she were, indeed, expressing agreement with her interlocutors in remaining silent, although originally she had no intention to illocute at all. She performs a coerced illocution, or at least so I will argue in this paper. Some would be worried that allowing for the possibility of coerced illocutions has morally unpalatable implications. I discuss this worry and, more broadly, the role of moral intuitions in theorising about our communicative practices.

1. Introduction

Philosophical discussions of silencing focus on what members of marginalised groups *cannot do* as speakers. Sometimes, they are prevented from speaking (Langton 1993) or gag themselves into silence (Dotson 2011). Other times, they speak but cannot perform the actions they want with their words (Langton 1993; Hornsby and Langton 1998; Kukla 2014; McGowan 2017, and others). McKinney (2016) invites us to shift our attention to what marginalised speakers are *made to do*. McKinney argues that dominant speakers sometimes unjustly elicit speech from marginalised speakers, offering as a central example coerced false confessions to crimes. She calls this phenomenon 'extracted speech'.

My project is similar in spirit to McKinney's. I identify a neglected, distinctive way in which marginalised speakers are *made to do* something. I call this phenomenon 'coerced illocutions'. Roughly, speakers are coerced into illocuting in my sense when they end up performing an illocution in fear of unjust consequences, although they intended to perform none and made no intentional speech contribution. This is importantly different from extracted

speech, where speakers are coerced, manipulated, or tricked, into intentionally making a speech contribution.

The paper proceeds as follows. After laying the necessary theoretical groundwork (§2), I argue that speakers can illocute without originally intending to do so and without intentionally making any speech contribution (§3), and sometimes they may do so coercively (§4). Finally, I consider what I call the 'moral objection' (§5).

2. Theoretical Groundwork

Austin (1962) argued that when we speak, we do several things at once. We perform a locution, a perlocution, and an illocution. Suppose I say, 'Bring me some coffee!'. I make an utterance, using meaningful words with a certain sense and reference, thereby performing a locution (1962: 64). By making this utterance, I may cause you to think that I want coffee, to feel annoyed with me, and to bring me coffee. I produce "certain consequential effects upon the feelings, thoughts, or actions of the audience" (101), thereby performing a perlocution. Importantly, in saying 'Bring me coffee', I perform an illocution; I *order* you to bring me coffee. Otherwise put, my utterance has the *illocutionary force* of an order, at least when certain conditions that are necessary for illocutionary success, felicity conditions, obtain (e.g., to issue an order, the speaker needs authority).

My discussion centers on examples of non-verbal illocutions. These, I contend, can occur by means of non-linguistic bodily movements and/or silences. For Austin, I can protest in throwing tomatoes at someone (1962: 118). For Alston, I can agree with someone in nodding (2000: 27). For Tanesini (2018: 112) and Klieber (2024: 2), I can tell someone that I am mad at them in remaining silent when asked 'Are you mad at me?'.

What fixes illocutionary force? What determines what kind of illocution one's utterance constitutes? Against the backdrop of theorists taking either the speaker (e.g., Bach and Harnish 1979; Bianchi 2020) or the hearer (Kukla 2014; Johnson 2023) to be responsible for fixing force, some argue that force is fixed by both speaker and hearer through a process of negotiation (e.g., Sbisà 2023; Caponetto and Cepollaro 2023; Witek 2023). According to one such theory of illocutionary force, Lucy McDonald's Collaboration Theory,

an utterance has illocutionary force Φ iff both (a) and (b) occur:

- (a) The hearer communicates to the speaker that she interprets the utterance as having force Φ .
- (b) The speaker then communicates to the hearer that she accepts the hearer's interpretation of the utterance as having force Φ (2022: 924).

That is, what matters for an utterance to have force Φ is, largely, that interlocutors reach an agreement; they both go off thinking (that they both think) that S performed Φ (932).

To illustrate, suppose I say to you, 'Get me some coffee', intending to issue an order. You manifestly interpret my utterance as an order, saying 'Yes ma'am/sir' (927). I respond, 'Chop chop!', thereby accepting your interpretation. Since (a) and (b) obtain, I perform an order (provided that additional felicity conditions obtain) (927). Now suppose you manifestly interpret my utterance as a request, rather than an order, saying 'OK, but you'll owe me'. If I reject your interpretation, saying 'This was an order!', my utterance is not a request (927), but perhaps it can be an order. If I accept your interpretation, saying 'Thank you!', I perform a *request* (927). Thus, a speaker can end up performing a different illocution than the one they originally intended.

3. Partially Accidental Illocutions

Speakers, I argue, can also end up illocuting although they originally had no intention to perform *any* illocution and without originally making any (intentional) speech contribution. Consider an example.

Offer: George and Lina are having dinner, sitting across each other. Lina pushes her plate of food forward, towards George, to make space to rest her arms on the table. George says 'Thank you. I might have some later', interpreting Lina as offering him her food. Lina thinks it would be awkward if she said that she did not intend to offer her food, so she plays along saying, 'No worries'.

Lina, I claim, performs the illocution of offering her food. Suppose that a little later George eats some of Lina's food. If no offer was performed, then Lina could get angry and complain to George. However, this would be unreasonable after the exchange in *Offer*, because in accepting George's interpretation, Lina

commits herself to giving George her food if he wants it. This is the normative job of offers, which indicates that an offer was performed. This seems aligned with McDonald's theory, since (a) George manifestly interprets Lina's gesture as an offer, and (b) Lina accepts his interpretation.

Illocutions like these are *partially accidental*. They are *accidental* because S originally has no intention to illocute, and *partially accidental* because intention comes in later, when S accepts H's interpretation. Provided that additional felicity conditions obtain, partially accidental illocutions occur iff

1. S makes an utterance without intending to perform any illocution.
2. H manifestly interprets S's utterance as illocution Φ .
3. S accepts H's manifest interpretation.

I understand 'making an utterance' in a broad sense to include bodily movements, silences, and words. Not any bodily movement, silence, or word will do, though. We take our communicative practices to be governed by reasonableness, and not *anything* can be reasonably interpreted as some illocution Φ . If I scratch my elbow while sitting in a café, a customer sitting nearby cannot reasonably interpret me as ordering them to get me coffee (unless we priorly agreed on a code). We should thus restrict 'making an utterance' to moves that can be reasonably interpreted as constituting illocution Φ . In PAIs, these are moves that are *conventionally understood* as Φ ing.

To precisify this, we can appeal to the notion of a 'conventional procedure', proposed by Austin (1962: 14) and later developed by Sbisà (2023). The conventional procedure for any illocution Φ is "a complex script, which may have a higher or lower number of options at each step, such that whoever executes its steps completely enough in the appropriate circumstances" (2023: 18) can be reasonably recognised as Φ ing. Sbisà suggests that illocutionary procedure is conventional in Millikan's sense of 'convention' (250): it is a pattern repeated partly due to weight of precedent (Millikan 2005: 2). For example, the first step of the conventional procedure for offering would be: S says something like, 'Would you like x?', or extends something to H (e.g.,

pushes their food towards H). If S makes one of these moves in the appropriate circumstances (when S has something they think H would want), they can be reasonably interpreted as making an offer.

Therefore, 'makes an utterance' in condition 1 means 'makes a move (speaks words, stays silent, makes bodily movements) that is a step in the conventional procedure for illocution Φ in appropriate circumstances'.

4. Coerced Illocutions

Sometimes, partially accidental illocutions can be coerced. Sometimes, we make an utterance, as defined above, without intending to illocute, and when our hearer manifestly interprets us as illocuting, we are coerced to accept this interpretation. I understand 'coerced' in the following way (inspired by but weaker than Nozick's (1969) understanding of coercion):

1. S would not want to accept H's interpretation.
2. S fears that H or some other person will bring about some unjust consequence C if S rejects H's interpretation, i.e., a consequence that it would be morally wrong to bring about because S rejected H's interpretation
3. S has been given reasons for the fear expressed in (ii), be that by H, a person other than H, or some cultural pattern (e.g., a woman may be given reason to fear that stranger men in certain contexts may harm her, even if no specific person has given her this reason).
4. S wants to avoid C, so they accept H's interpretation.

To illustrate coerced illocutions, consider a revised version of *Offer*.

Coerced Offer: Lina is dining with George. Lina knows that if George thinks someone impolite or disrespectful, he gets angry and swears at them. Lina pushes her plate of food forward, towards George, to make room to rest her arms on the table. George interprets Lina as offering him her food. He says 'Thank you! I might have some later!'. Lina would like to tell George that she did not intend to offer her food. However, fearing that George will start swearing at her, Lina plays along saying 'No worries'.

This case illustrates illocutionary coercion. Although Lina would want to reject George's interpretation, she fears that, if she does so, George will verbally abuse her, something that George has given her reason to fear given his past behaviour, and something that would be morally wrong for George to do to Lina for rejecting his interpretation. To avoid verbal abuse, Lina accepts George's interpretation. While here Lina performs a coerced illocution, in *Offer* this may not be the case, since she avoids rejecting George's interpretation because she fears awkwardness, which may not be an unjust consequence.

When we are coerced into illocuting, we are wronged as speakers. The wrong of illocutionary coercion can be characterised in terms of McDonald's notion of 'negative speaker autonomy' (2021: 19). Coerced illocutions unjustly undermine a speaker's capacity not to perform illocutions that they do not want to perform. Unlike cases where speakers want to perform illocution Φ but unjustly end up performing Ψ instead (see Langton (1993), Kukla (2014)), where speaker autonomy is undermined mid-way through the process of illocuting, in coerced illocutions it is undermined from start to finish. Coerced speakers do not choose to start illocuting and do not freely choose to accept H's interpretation.

This injustice can be *local* or *systematic*. Some speakers may be coerced to illocute in fear of unjust consequences that arise from local, idiosyncratic circumstances, as Lina does in *Coerced Offer*, whose fear that she may be sworn at arises from George's idiosyncratically poor moral character. Other times, though, a speaker may be coerced into illocuting in fear of unjust consequences arising from prejudices against a social group that S belongs to, and that apply to S *qua* member of the group. In those cases, coerced illocutions are *systematic*. Consider an example.

Aggressive Woman: Aliko, a female employee, is silent in a professional meeting because she is frustrated with colleagues not taking her contributions seriously. A male colleague says, 'I'm glad you agreed with everything we said', manifestly misinterpreting her silence as assent. Aliko would like to say that she did not agree; that she was just too fed up to speak. However, unlike men, in many work environments women are expected to make positive and supportive contributions; when they flout this expectation, they are often labeled

‘aggressive’ and ‘uncooperative’, and risk seeing their chances to advance their careers curtailed (Howell et al., 2002: 118). Fearing such repercussions, Aliko plays along and says, ‘Yes your ideas were good’.

Aliko performs a *systematic* coerced illocution. Although she does not intend to illocute, her colleague misinterprets her as assenting because of a convention in professional meetings that silence means assent. She accepts his interpretation reasonably fearing unjust consequences that apply to her *qua* woman and that arise from a prejudice against women, namely that women are aggressive or uncooperative if they speak their mind or disagree with men.

There are many prejudices against marginalised groups making it risky for marginalised speakers to correct dominant speakers. They may be labeled in stigmatizing ways (e.g., ‘hysterical’, ‘aggressive’) and fear repercussions that dominant speakers have the power to enforce (e.g., getting fired, losing a promotion, physical violence). Thus, marginalised speakers are particularly vulnerable to illocutionary coercion.

5. The Moral Objection

One may object that the claim that speakers can be coerced into illocuting has morally unpalatable consequences.

For instance, in certain cultural environments, there is a pattern of interpreting women who wear short skirts as inviting men to sexual comments on their appearance, and of men responding ‘cooperatively’ to this ‘invitation’ by making such comments. Men report that they interpret women this way, and women report that they are aware that they are interpreted this way (Kissling 1991: 453; Kissling and Kramarae 1991: 81). Sometimes people are taught that this is what wearing short skirts means and, as a result, reproduce the pattern (Kissling 1991: 453). Therefore, one can plausibly suggest that this is a convention (a pattern repeated partly because others have instantiated it before), such that if a woman wears a short skirt, she is conventionally interpreted as inviting men to make sexual comments. Consider the following scenario in such a cultural environment.

Short Skirt: A man misinterprets a woman wearing a short skirt as inviting him to sexually comment on her appearance. He ‘responds’ by making such a comment. The woman realises that he interpreted her as

inviting him. She would want to tell him to leave her alone, but reasonably fearing that he may harm her if she responds confrontationally, she plays along by smiling and saying, 'Thank you!'. They both think (that they both think) that the woman invited the sexual comment.

If we allow that speakers can be coerced into illocuting, we must concede that the woman invited the man to make a sexual comment. This is morally unpalatable, for at least two reasons. First, the man arguably did something morally wrong, but if the woman invited him, then perhaps he has not done anything wrong. Second, our conclusion reinforces the sexist narrative that women invite their harassers.

An objection like this has been raised by theorists like Jacobson (1995), Bird (2002), and, in fact, McDonald (2021, 2022) herself. For this reason, McDonald suggests (tentatively) that "agreements about illocutionary force only have constitutive power when the speaker and the hearer can participate freely in the interaction" (2022: 936). McDonald effectively proposes a universal 'freedom' felicity condition, thereby foreclosing the possibility of illocutionary coercion.

One line of defense against the moral objection is to argue that the conclusion that the woman invited the man is not as morally unpalatable as the objector suggests. First, the man (morally) *should not* have interpreted her as inviting to begin with. He follows a warped, sexist convention when he could easily access information suggesting that women do not intend to invite men in wearing short skirts (e.g., by asking women in his community, via online search). He is *culpably ignorant*, as he has done what Smith (1983: 544) calls *deficient investigation*. Further, he is blameworthy for his utterance's perlocutionary effects, such as the woman feeling scared, which again, he should have been aware of. His blameworthiness for following the sexist convention and scaring the woman holds regardless of whether the woman played along with him.

Second, although both the sexist narrative and our conclusion have it that the woman invited the man's comment, the two are radically different. The sexist narrative implies that, in wearing a short skirt, the woman intended to invite

the man to make sexual comments because she wants such comments. By contrast, our conclusion says that the woman did *not* intend to invite the man; she was *misinterpreted* as inviting due to a warped convention and was *coerced* into accepting that she invited him, even though she did *not* want to do so. On *this* picture, the woman inviting the man is an injustice that she suffers, not a defense for perpetrators.

A second line of defense is to question the methodological claim implicit in the objection (and rather explicit in McDonald (2021: 11)), namely, that our language theories should capture our moral intuitions. Clearly, we need to capture everyday language intuitions because we aim to understand our communicative practices, and those intuitions offer a ‘window’ to those practices. But why capture moral intuitions? They mostly show us what our communicative practices *should* look like rather than what they *do* look like in our morally non-ideal world.

Further, the two desiderata (capturing moral *and* everyday language intuitions) often clash. Setting a high bar for illocuting driven by our moral commitments often conflicts with our everyday language intuitions, which suggest that we *do* routinely illocute, even when we are not ‘sufficiently free’. McDonald acknowledges this when she says that the ‘freedom’ condition “will have quite radical consequences, because situations in which all interlocutors participate freely seem few and far between... The result is that a significant number of utterances will lack illocutionary force.” (2022: 936). Setting clashing desiderata is methodologically problematic, at least without a further story about which should be prioritized and when.

A methodology that avoids these problems is consulting our everyday language intuitions to cartograph our communicative practices, as morally flawed as they are, and then using our moral intuitions to critique and revise those practices. This is the methodology I have been assuming. On this methodology, the morally unpalatable consequences of my proposal are, to borrow Kukla’s phrase, “not a flaw in a speech act theory that acknowledges [that speakers can be coerced into illocuting], but a problem with the world as we find it” (2023: 25).

References

- Alston, William (2000) *Illocutionary Acts and Sentence Meaning*, Ithaca, NY: Cornell University Press.
- Austin, John Longshaw (1962) *How To Do Things With Words*, London: Oxford University Press.
- Bach, Kent; Harnish, Robert M. (1979) *Linguistic Communication and Speech Acts*, Cambridge, MA: M.I.T. Press.
- Bianchi, Claudia (2020) "Discursive Injustice: The Role of Uptake", *Topoi* 40(1), 181–190.
- Bird, Alexander (2002) "Illocutionary Silencing", *Pacific Philosophical Quarterly* 83, 1-15.
- Caponetto, Laura; Cepollaro, Bianca (2023) "Bending as Counterspeech", *Ethical Theory and Moral Practice* 26(4), 577–593.
- Dotson, Kristie (2011) "Tracking Epistemic Violence, Tracking Practices of Silencing", *Hypatia* 26(2), 236–257.
- Hornsby, Jennifer; Langton, Rae (1998) "Free Speech and Illocution", *Legal Theory* 4(1), 21–37.
- Howell, Sharon L.; Carter, Vicky K.; Schied, Fred M. (2002) "Gender and Women's Experience at Work: A critical and feminist perspective on Human Resource Development", *Adult Education Quarterly* 52(2), 112–127.
- Jacobson, Daniel (1995) "Freedom of Speech Acts? A Response to Langton", *Philosophy & Public Affairs* 24(1), 64–78.
- Johnson, Cassie Rebecca (2023) "Illocutionary Relativism", *Synthese* 202(3), 1-18.
- Kissling, Elizabeth Arveda (1991). Street Harassment: The Language of Sexual Terrorism. *Discourse & Society*, 2(4), 451–460.
- Kissling, Elizabeth Arveda; Kramarae, Cheris (1991) "Stranger Compliments: The interpretation of Street Remarks", *Women's Studies in Communication* 14(1), 75–93.
- Klieber, Anna (2024) "Silencing Conversational Silences", *Hypatia* 39(4), 934-952.
- Kukla, Rebecca (2014) "Performative Force, Convention, and Discursive Injustice", *Hypatia* 29(2), 440–457.
- Kukla, Quill Rebecca (2023) "Uptake and Refusal", *Inquiry*, 1–27.
- Langton, Rae (1993) "Speech Acts and Unspeakable Acts", *Philosophy & Public Affairs* 22(4), 293–330.
- McDonald, Lucy (2021) "Your word against mine: The Power of Uptake", *Synthese* 199(1–2), 3505–3526.

- McDonald, Lucy (2022) "Reimagining illocutionary force", *The Philosophical Quarterly* 72(4), 918–939.
- McGowan, Mary Kate (2017) "On Multiple Types of Silencing", in: Mari Mikkola (ed.) *Beyond Speech: Pornography and Analytic Feminist Philosophy*, NY: Oxford University Press, 39–58.
- McKinney, Rachel Ann (2016) "Extracted speech", *Social Theory and Practice* 42(2), 258–284.
- Millikan, Ruth Garrett (2005) "Language Conventions Made Simple", in: *Language: A Biological Model*, Oxford: Oxford University Press, 1–23.
- Nozick, Robert (1969) "Coercion", in: Sidney Morgenbener, Patrick Suppes, and Morton White (eds.) *Philosophy, Science and Method: Essays in Honor of Ernest Nagel*, New York: St Martin's Press, 440–472.
- Sbisà, Marina (2023) *Essays on speech acts and other topics in pragmatics*, Oxford: Oxford University Press.
- Smith, Holly (1983) "Culpable Ignorance", *The Philosophical Review* 92(4), 543–571.
- Tanesini, Alessandra (2018) "Eloquent Silences", in: Cassie Rebecca Johnson (ed.) *Voicing dissent: The Ethics and Epistemology of Making Disagreement Public*, New York: Routledge, 109–128.
- Witek, Maciej (2023) "Interactional Negotiation", in: Laura Caponetto and Pablo Labinaz (eds.) *Sbisà on Speech as Action*, Cham, Switzerland: Palgrave Macmillan, 97–12.

Acknowledgement

I want to thank Laura Caponetto for her immensely helpful feedback on several drafts of versions of this paper. I also want to thank Philosophy postgraduate students at the University of Cambridge for helpful feedback when I presented this paper to them.

AI Models as Conceptual Agents

Lena Wang (Cambridge, United Kingdom)

Abstract

Conceptual engineering is taken to be synonymous with semantic *amelioration*—that is, improving a concept. However, I propose that we think of conceptual engineering as not merely a methodology, but as a process, one that includes semantic *deterioration*. Doing so allows us to identify systems that make our concepts worse, namely, AI tools that predict social kinds. That AI conceptually engineers might seem surprising, as these models are typically thought of as instruments that reflect their programmers' beliefs. I argue, however, that when put to the task of classification, the social kind they pick out is distinct from whatever concept their programmers have. Attending to how AI performs conceptual engineering clarifies that AI does not merely reflect existing problematic concepts, but creates them.

Problems emerge if we believe that AI models inherit our concepts. Firstly, we may come to wrongly believe that these classifications are fixable, for example, by training the model on less defective concepts. Secondly, it is by thinking of AI as engineering its own distinct defective concepts that clarifies how it does not merely perform quicker classifications, but *changes* existing classifications, often in unpredictable and worse ways. We can then come to understand that AI can result in a kind of *epistemic hijacking*, where we allow its defective concepts to override our own.

1 Introduction

The raced and gendered harms of Artificial Intelligence ('AI') models and other algorithms have been discussed widely in technology ethics (Zuboff, 2019; Birhane, 2023; Noble, 2018; Benjamin, 2019). This paper identifies a distinct kind of risk that emerges from the use of AI tools—specifically, from their conceptual agency. By drawing on the methodologies of conceptual engineering, we come to understand that AI tools do not merely reflect existing problematic concepts, but create them.

§2 provides an overview of conceptual engineering, and how the field seeks to assess and improve the representational devices that we use to categorise and understand the world. However, there are also systems that make our concepts worse. AI models, as I will argue in §3, are examples of these latter kinds of systems. §4 defends the upshot of this framing. Namely, it demonstrates that two widespread beliefs in AI ethics literature can be harmful misunderstandings: that biased AI classifications can be improved by

diversifying the representation of oppressed groups in training sets, and that AI predictions are accelerating the rate at which problematic decisions are currently being made.

2 Conceptual Engineering

Concepts are representational devices that pick out a set of things, and groups them as a kind. These kinds may be natural, such as WATER, or social, such as FAMILY. When our concepts undermine our ability to reason properly, or impede political justice, they are described as defective. Conceptual engineering seeks to assess and revise concepts to address defects (Cappelen & Plunkett, 2020). This process of redefinition is described as *semantic amelioration* (Haslanger, 2020). Crucially, for Haslanger, this does not transform the concept into a new one, so long as the function of the concept remains the same. So, the concept of FAMILY historically had a set F as its content, representing a social kind consisting of all and only heterosexual couples with their biological offspring. However, this concept has undergone semantic amelioration such that it now has set F^* as its content, which includes domestic arrangements of queer couples, adoptees, and single-parent families. We need not form a new concept to pick out this changed content, because the revised FAMILY continues to play the same functional role in social coordination—for example, in economic partnership and child-rearing. Hence, the process of semantic amelioration is thought of as a methodology that philosophers can pursue to improve the concepts we use.

However, consider that the concept of ABORTION in United States law has undergone a change in contents. Before the repeal of Roe v Wade, its content included the termination of pregnancy up until foetal viability. Now in some states, a subset of these terminations has been recast as murders and ABORTION picks out the smaller, remaining set—including terminations only up until 6 weeks of pregnancy. However, this conceptual change is for the worse. I posit that we should think of conceptual engineering not merely as a methodology performed by philosophers, but as a process that can be enacted by any system, institution, or community. This allows us to identify systems that enact semantic deterioration. It is my contention that AI tools are one such system.

3 AI as a Conceptual Agent

Artificial Intelligence (AI) technologies represent, in theory, intelligent machines, and in practice, a set of statistical tools used to automate and supplant human prediction and problem-solving methods (Russell and Norvig, 2021). These include machine learning (ML) systems that analyse patterns in existing data to build models that can perform tasks like classification and prediction (Ibid.). This paper focuses specifically on AI models that make predictions that categorise based on social kinds, for example, those that classify people by protected categories like gender (Antipov et al., 2016) and sexuality (Wang & Kosinski, 2018).

I will specifically focus on the gender classification model described in Antipov et al. (2016), which is a deep learning model that uses facial image inputs, trained using supervised learning. Supervised learning requires a dataset of annotated inputs. Antipov et al. (2016) use a dataset of approximately 505,000 facial images with gender annotations. An annotated dataset is partitioned into two exclusive sets: the *training set*, used for building the model, and the *testing set*, used after training to evaluate the model's accuracy.

Because this model uses deep learning, it eliminates the need for researchers to manually select visual features that are relevant for classification. Instead, the model will establish the importance of visual features based on how determinative they are in enabling predictions that correctly match the gender annotations in the training set. For example, during the training phase, the model may isolate brow and lip thickness as the most relevant features distinguishing women and men in the training set. It finds that thin brows and thick lips are good predictors of a woman, and vice versa for a man. The model prioritises these features in performing predictions, and will categorise images with thinner brows and thicker lips as women. After training, the model is evaluated for overall accuracy based on how many predictions in the testing set accord with the annotations. The model is thus configured to categorise inputs—images representing people—as WOMAN or MAN. This training process, I contend, is sufficient for enabling the AI to conceptually engineer the concept of WOMAN.

That AI may perform conceptual engineering might be surprising to some. The surprise may be motivated by the following premises: the process of

conceptual engineering can only be undertaken by things with agency—in particular, a kind of conceptual agency. This can be understood, on the standard view of action, as some capacity to perform intentional actions, where the intentional action in this case is the creation and amendment of concepts. The capacity to perform intentional actions, in turn, requires the possession of intentional mental states. Similarly, the manipulation of concepts requires the formation of concepts as mental representations, which requires mental states. As AI tools do not possess mental states, they cannot be conceptual agents, and thus, cannot conceptually engineer.

Wading into debates around agency and representational mental states is unnecessary for the claims I am making in this paper. I have a minimal definition of conceptual agency in mind. Specifically, I mean that AI models can independently or autonomously amend the extension of concepts.

Let me clarify that I am using concepts as a mind-independent description, analogous to a “category” that performs ontological classifications picking out particular social kinds. Hence, even systems without mental states can create concepts. For example, the state determines and applies concepts like ELIGIBLE VOTER, CITIZEN, and so on.

Secondly, I have posited that AI models can amend the extension of concepts to contents that they have autonomously selected. Supposing that AI is indeed successful in amending the extension of concepts, one could argue that it does not do so independently. Human intervention is required at two stages: during model development, and in applying the predictions of the AI. For the latter, it is thought that since the outcomes of AI predictions go on to inform human decision-making, the AI model itself cannot be described as an independent concept engineer. However, what matters for conceptual agency is not whether humans are causally dependent on the outcome of some categorisation, but whether they constitutively determine the definition of that concept, and the selection of its extension.

The remaining contention, then, rests on whether human intervention at the model development stage entails that AI models are autonomously performing semantic deterioration. Specifically, the thought is that if AI models do not independently create or change concepts, but instead, merely make their

predictions using the categories that their programmers have embedded in the model, this deprives the model of conceptual agency. The typical view is something as follows. Machine learning models inherit the categorisations that exist in the training data. For example, machine learning models are thought to reinforce existing biases when the data they are trained on is biased (Buolamwini & Gebru, 2018). If these models were put to the task of categorising, for example, men and women, they would misclassify darker-skinned women at a much higher rate. So, the concept of WOMAN that they are employing is defective. These conceptual defects, it is thought, are inherited from the defective concept of WOMAN intrinsic in the training data, and thus, intrinsic in the perspective of the creators of that training set.

I argue that this is not precisely the case. AI models do not simply ‘reuse’ the concepts they are trained on. To see this, imagine that the programmers in Antipov et al. have some concept of WOMAN, which has the contents represented by the set W . Imagine that all samples in the training set are labelled according to this concept. The definition of WOMAN may be something like the following: S is a woman in context C iff S has observed or imagined bodily features presumed to be evidence in C of female sex characteristics (Haslanger, 2000).

The AI model also performs classifications, and in determining based on input who can be classified as WOMAN, thereby picks out its own extension of WOMAN. The extension would contain the images annotated as ‘Woman’ in the training set. Let us define this set of people (represented by their images), $W1$. When the model is applied to the testing set and the real world, it will identify new sets of images (and people) that it takes to belong to the extension of WOMAN, $W2$ and $W3$ respectively. The model has therefore defined a set of people, $W^* = W1 \wedge W2 \wedge W3$, which belongs to its extension of WOMAN.

Suppose that the model has been applied to all possible people in the world, such that W^* encompasses all people that, according to the model, are in the extension of WOMAN. Whether the model has changed the contents of the concept WOMAN, then, depends on whether W is the same as W^* . While it is the case that the subset $W1$ would also be contained in the extension of W , it is not the case that W is identical to W^* . Given that AI models are, inherently, predictive, and that no prediction methodology is universally and flawlessly

accurate, the social kind picked out by the model will inevitably be distinct from the social kind picked out by its creators. Its functional role, however, remains the same. For example, when the model is used for targeted advertising, these gender concepts have the same function as non-AI engineered concepts—they elucidate the intended market for a product. As such, the model has performed conceptual engineering, in changing the extension of the concept of WOMAN. Specifically, it has defined an *algorithmic concept*.

4 Upshot

First, it should be noted that conceptual agency does not equate to moral agency. Recognising that AI models are conceptual agents does not mean we also need to recognise them as the targets of blame when morally bad outcomes ensue. Furthermore, AI systems are not unique in creating defective concepts, nor are they the only kind of technology that engineers defective concepts.

That AI is not a uniquely bad kind of conceptual agent, then, raises the question of why they are the focus of this paper. Indeed, one may think that these systems are merely as concerning as individuals having different contents for concepts, which they inevitably do, as it is highly unlikely that two individuals share the exact same notion of the members of a social kind. Moreover, while it is not the case that the algorithmic concept is identical to the programmers', provided that the model is sufficiently trained, it is likely that they share a significant subset of their extensions.

Despite these facts, there are two upshots for recognising AI as a conceptual agent. Specifically, there are harmful misunderstandings that emerge if we were not to think of AI as a conceptual agent. Such harms are compounded by three contingent ways of treating these kinds of models. The first is that we allow these models to supersede human decision-making. The second is that we believe and act as if these models were entirely objective, true predictors. The last is that we allow these models to proliferate, and operate at enormous scale, thus performing conceptual engineering systemically.

The first harmful misunderstanding is the thought that these kinds of models are fixable by, for example, improving the representation of the oppressed groups in the training set. This misunderstanding arises from thinking that AI

merely ‘reuses’ our concepts—as Geburu (2020: 260) writes, our “biases are propagated through AI.” The thought, then, is that we can ‘feed’ it less biased training data that encompass less defective social concepts. On this view, doing so will allow us to create AI models that do not engineer defective social concepts. Yet, we can see why this will not work. For one, some concepts like WOMAN or INDIGENOUS are highly specific to a certain individual and their community, and are not conducive to external prediction (Keyes, 2018; Scheuerman et al., 2019). AI models, unlike people, cannot simply ask an individual or group for the correct categorisation, because that would betray its fundamental purpose—to predict, rather than to merely report. Attempts to diversify the training set inevitably fail—intrinsically, the more diverse a training set becomes, the less shared features exist, thus reducing the presence of patterns required for prediction, and reducing predictive capabilities of the model. We can conclude, then, that attempts to diversify the training set of some AI prediction models will not allow them to output diverse predictions. Mistaking AI as lacking conceptual agency lulls us into the false belief that AI can adopt some non-defective concepts, once we determine their definitions and train the models on these concepts. Instead, some models do not and cannot reuse non-defective concepts, but will always engineer their own defective concepts.

More significantly, however, disregarding the conceptual agency of AI tools can cause us to underestimate its harms. However, even when we regard AI as a subjective system that encodes and reflects oppression, we may merely think that AI is replicating problems as they already exist. For instance, if AI classifies someone as a criminal, we may think that this is because the model has performed a (potentially problematic) categorisation that humans would, but simply faster. We may attribute unexpected categorisations to the complexity of such models: that they are able to consider a variety of factors in their judgment that would be difficult for us to reconcile. The only critique, then, may be that AI reinforces existing injustices merely by increasing the speed at which these injustices occur. However, this is not necessarily the case. Recall that the algorithmic concept has a different extension than the human-engineered concept. As a result, the model does not merely make *faster* categorisations, but new ones—it could label someone a criminal, even when our existing (problematic) definitions would not. Understanding AI as a

conceptual agent makes this distinction clear. AI tools are not merely reenacting defective categories that exist, but changes classifications, often in unpredictable and worse ways. As Benjamin (2019) argues, these models *produce*, rather than screen for, properties like criminality.

Should we hold that AI are not conceptual agents, we may come to take on algorithmic concepts in place of our own. This is especially true given the reach of the multinational institutions that deploy AI models—their influence increases the likelihood that the concepts they use become dominant (Richard, 2020). Consider if, as one AI start-up has suggested (Metz, 2023), algorithmic concepts are used to alert security guards when people of the “wrong” gender are trespassing. The guards may manually override this alert, but nevertheless, their actions (whether to act on the alert, or dismiss it) are in response to the algorithmic concept. It may in fact be company policy that such alerts must be followed, in which case, the guards’ own concept of gender would be overridden by the algorithmic one. Or, if AI were to classify someone as a criminal, we may come to perceive that person as one. These are cases of *epistemic hijacking* by AI tools. In this way, AI tools come to not merely reflect social reality, but engineer concepts that shape our ontology—that carve society at its joints. The risk that our hermeneutical resources could be overridden by those of a model whose justificatory reasoning is inaccessible (Carabantes, 2020) is one we ought to take seriously. Our concepts, while defective, can be amend through semantic amelioration. I have already argued that in some cases, AI tools are inherently unable to engineer non-defective concepts. To allow these intrinsically defective concepts to shape our social ontology is not a reinforcement of the status quo, but a dramatic deterioration of it.

I am not entirely cynical about the use of AI models. We can imagine, for instance, that an AI model deliberately made for very localised contextual usage could engineer concepts that are not as defective. This is because, for this use case, the model does not suffer in its predictive capability with a more diverse dataset. This may reduce the divergence in its engineered concepts from our own. Similarly, AI predictions centred around natural concepts and kinds may circumvent the problem of generating inherently defective concepts.

5 Conclusion

AI tools are not merely accelerators of socially embedded conceptual defects. They can, in fact, shape the world in much worse ways. Our efforts to improve them may be fruitless. Yet, coming to these conclusions about these technologies is helpful—it lets us modify our decision-making processes to avoid harmful dependencies on these models. These conclusions are made apparent by understanding AI tools as players in the game of hermeneutics. They are conceptual agents, able to redefine the social kinds they are used to categorise, and these concepts can come apart from the representations of their programmers. Understanding conceptual engineering as a process clarifies the ontological risks associated in using AI—a necessary perspective, if it is to have an outsized role in our social lives, as many imagine.

References

- Antipov, Grigory; Berrani, Sid-Ahmed; Dugelay, Jean-Luc (2016) “Minimalistic CNN-based ensemble model for gender prediction from face images”, *Pattern Recognition Letters* 70, 59–65.
- Benjamin, Ruha (2019) *Race After Technology: Abolitionist Tools for the New Jim Code*, Cambridge: Polity Press.
- Birhane, Abeba (2023) “Algorithmic Colonization of Africa” in: Stephen Cave and Kanta Dihal (eds.) *Imagining AI: How the World Sees Intelligent Machines*, Oxford: Oxford University Press.
- Buolamwini, Joy; Gebru, Timnit (2018) “Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification”, *Proceedings of the 1st Conference on Fairness, Accountability and Transparency* 81, 77–91.
- Cappelen, Herman; Plunkett, David (2020) “A Guided Tour of Conceptual Engineering and Conceptual Ethics” in: Herman Cappelen, David Plunkett, and Alexis Burgess (eds.) *Conceptual Engineering and Conceptual Ethics*, Oxford: Oxford University Press, 1-26.
- Carabantes, Manuel (2020) “Black-box artificial intelligence: An epistemological and critical analysis”, *AI Soc.* 35(2), 309–317.
- Gebru, Timnit (2020) “Race and Gender” in: Markus D. Dubber, Frank Pasquale, and Sunit Das (eds.) *The Oxford Handbook of Ethics of AI*, Oxford: Oxford University Press, 252–269.
- Haslanger, Sally (2000) “Gender and Race: (What) Are They? (What) Do We Want Them to Be?”, *Noûs* 34(1), 31–55.
- Haslanger, Sally (2020) “Going On, Not in the Same Way” in: Herman Cappelen, David Plunkett, and Alexis Burgess (eds.) *Conceptual Engineering and Conceptual Ethics*, Oxford: Oxford University Press, 230-260.

Keyes, Os (2018) “The Misgendering Machines: Trans/HCI Implications of Automatic Gender Recognition”, *Proceedings of the ACM on Human-Computer Interaction* 2(CSCW), 1–22.

Metz, Rachel, (2019), “AI software defines people as male or female. That’s a problem.”, *CNN*.

Noble, Safiya Umoja (2018) *Algorithms of Oppression: How Search Engines Reinforce Racism*, New York: NYU Press.

Richard, Mark (2020) “The A-project and the B-project” in: Herman Cappelen, David Plunkett, and Alexis Burgess (eds.) *Conceptual Engineering and Conceptual Ethics*, Oxford: Oxford University Press, 358–378.

Russell, Stuart; Norvig, Peter (2021) *Artificial Intelligence: A Modern Approach, Global Edition*, Harlow: Pearson Education Limited.

Scheuerman, Morgan Klaus; Paul, Jacob M.; Brubaker, Jed R. (2019) “How Computers See Gender: An Evaluation of Gender Classification in Commercial Facial Analysis Services”, *Proceedings of the ACM on Human-Computer Interaction* 3(CSCW), 1–33.

Wang, Yilun; Kosinski, Michal (2018) “Deep neural networks are more accurate than humans at detecting sexual orientation from facial images”, *Journal of Personality and Social Psychology* 114(2), 246–257.

Zuboff, Shoshana (2019) *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power*, London: Profile Books.

Notes on MS 168 – “... a bag full of raisins”

Peter K. Westergaard (Copenhagen, Denmark)

Abstract

Among Wittgenstein's late notebooks from 1947–1950, MS168 attracts a certain interest. Notebook No. IV is a mystery in Wittgenstein's *Nachlass*. In the literature on Wittgenstein, it has been discussed only briefly and sporadically. There is uncertainty and disagreement about the background to the notebook and where it should be placed in the *Nachlass*. This paper presents MS168, and the few thoughts that have been expressed about it. The presentation concludes with further reflections on its content.

“On general subjects”

Among Wittgenstein's notebooks from 1947–1950, the so-called Notebook No. IV attracts a certain interest. MS168 is a mystery in Wittgenstein's *Nachlass*, albeit, admittedly, a fairly minor one. There is uncertainty and disagreement about, on the one hand, where the notebook should be placed and what role it plays within the *Nachlass*, and on the other its purpose and background. In the following, I shall take a closer look at the notebook and present some of the responses that it has elicited. I will conclude my presentation with a few thoughts resulting from my work on MS168.

In his catalogue of the *Nachlass*, Von Wright describes the manuscript thus: “**168.** Notebook. Fair manuscript copy of some remarks from the years 1947–49 on general subjects. 12pp.” (von Wright 1982a,46). This is elaborated in the commentary: “**168.** The remarks, entered in the reverse order (1949–47), are from 136–138 (MS volumes Q–S)” (von Wright 1982a,54). There are twenty-three remarks on a wide range of topics such as “recounting a dream”, religious belief and superstition, stylistic devices and aesthetic impressions, poetic mood, education, tradition, genius, and the value of Mahler's music. Indirect references are made to Kierkegaard, Goethe, and Freud, while Shakespeare, Schopenhauer, Schiller, Kraus, Busch, and Mahler are mentioned directly.

All remarks in MS168 are reproduced in von Wright's *Vermischte Bemerkungen*, with one exception. The exception, which is reproduced as §110 in the first volume of *Last Writings on Philosophy of Psychology*, reads: “Die unheilbare Krankheit ist die Regel, nicht die Ausnahme” (MS168,3r). All the remarks, including the exception, appear in the first comprehensive selection of

material that von Wright compiled and transcribed in 1965–66 under the heading of “A *COLLECTION OF REMARKS BY LUDWIG WITTGENSTEIN* on questions connected with his Life and Work; the Nature of Philosophical Inquiry; Art, Religion and the “Philosophy of Life”; the I, the Will, and the World and various other General Topics”; otherwise known as the *Urfassung der Vermischte Bemerkungen*.

MS168 is a notebook with a total of twenty lined pages. It contains a selection of twenty-one remarks from MS136 and MS137, each marked with an “Anfangsstrich-Endstrich-Sektionsmarkierung”. In addition, the notebook begins with two remarks dated, respectively, “16.1.1949” and “1.49”, and which appear only in MS168. All the other notes are consistently marked with dates that correspond to those in MS136 and MS137 indicating when the remarks were written. The total of twenty-three notes are written on the first thirteen pages starting on page two. The remaining seven pages are without text. Fifteen of the notes are from MS137,42 to MS137,122 (from 30.05.1948 to 11.12.1948) and six from MS136,16 to MS136,91 (from 21.12.1947 to 11.01.1948), i.e. they represent a time-span from 21.12.1947 to 16.01.1949, or just over one year. As copied in MS168, the notes do not follow the original chronological order. They are written in Wittgenstein’s best handwriting. It is unknown when exactly the notebook was compiled. Pinch and Swedberg are of the opinion that MS168 belongs among the “manuscripts Wittgenstein worked on during the whole year of 1949”, whereby they exclude from this period Wittgenstein’s visit to Ithaca from late July to late October (Pinch & Swedberg 2013,25n8). Rothhaupt assumes that the first date of the collection, “16.1.1949”, is identical with “der Zeitpunkt, an welchem Wittgenstein diese Auswahl und Zusammenstellung begonnen hat” (Rothhaupt 2013,69). Keicher is of the same opinion (Keicher 2014,150). However, it cannot be ruled out that MS168 was not written until after Wittgenstein’s return from America in October 1949 or perhaps even during 1950.

Thus, the remarks in MS168 are effectively a selection from MS136 and MS137, which means they could be seen as being “at home” among those of Wittgenstein’s *Denkbewegungen* that we associate with *Philosophical Investigations* Part II, insofar as MS136 and especially MS137 constitute one of the main sources for the remarks in MS144. Both manuscripts were written during Wittgenstein’s stay in Ireland. They touch on many different questions

and topics relating to what we might collectively characterise as the grammar or logic of psychological concepts. In addition to these topics, the manuscripts contain some “general remarks” on the method of philosophy and several on aesthetics. Further, there are a number of remarks of autobiographical relevance, some of them in code.

Roughly a year before he began work on the abovementioned *Urfassung* of *Vermischte Bemerkungen*, von Wright mentioned MS168 in a letter to Anscombe. In June 1964, he wrote: “This idea of a volume of general remarks is worth taking seriously. There are, for example, a number of remarks in the typescript of October 1948 to March 1949 which might go into such a volume and also in the typescript of last writings,” to which he adds, “(i.a. the whole Notebook IV [MS168])” (quoted from Erbacher 2017,89). When preparing *A COLLECTION*, however, von Wright disregarded MS168 insofar as it duplicates material that is already present in MS136 and MS137, and it was the positions of the remarks in these manuscripts that he used as identifying references. *A COLLECTION* does reference MS168, however, in connection with the first two remarks from that manuscript, since these do not occur anywhere else in the *Nachlass*.

“The whole book looks a bit like ...”

MS168 is mentioned only rarely and briefly in the literature on Wittgenstein. Schulte refers to the notebook during his discussion of Wittgenstein’s remarks on Shakespeare. There he suggests that the notebook may have been compiled with some specific occasion in mind: “The whole book looks a bit like a kind of offering, perhaps a present to a friend or a relative” (Schulte 2013,18). Rothhaupt dwells on MS168 at somewhat greater length. For him, the notebook is evidence “dass Wittgenstein wirklich den Plan hatte und dessen Realisierung tatsächlich begann, eine Anfangsstrich-Endstrich-Sammlung (ggf. zusammen mit der – vielleicht sogar bereits exzerpierend erstellten – KRINGEL-BUCH-Sammlung) herzustellen” (Rothhaupt 2013,69). Rothhaupt stresses that we do not know why Wittgenstein did not continue to work on and to complete the “collection”. Thus Rothhaupt’s view differs from that of Schulte, who seems to regard MS168 as a finished manuscript.

For Rothhaupt, MS168 is evidence of one of Wittgenstein’s decisive intentions, if not indeed a proof within the *Nachlass*, that Wittgenstein wanted to compile

a sizeable selection of “general remarks”. It is an intention that Von Wright’s selection and publication of *Vermischte Bemerkungen* obscures and fails to recognise. Rothhaupt assumes that Wittgenstein himself wanted, and had in fact begun, to compile a selection of “allgemeine Bemerkungen” – in other words, of “kulturelle Bemerkungen”. To this end, Wittgenstein used special section markings, namely: a) an “O” to identify selected remarks dated from 11.10.1929 to 01.11.1931 (from MS107 to MS112), which constitute what Rothhaupt calls “der erste KRINGEL-BUCH-Teil”, plus a selection from 11.09.1937 to 14.04.1938 (from MS118, MS120 and MS158), which constitute what Rothhaupt calls “der zweite KRINGEL-BUCH-Teil”. And b) the section marking “|...|” (the so-called “Anfangsstrich-Endstrich-Sektionsmarkierung”) to identify selected remarks written between 04.05.1931 and 18.04.1951. The remarks that make up MS168 belong to the latter group.

The two collections (a and b) cover the period from October 1929 to April 1951. The section markings of both types were added either at the time when the remarks were composed, or later. Then, early in 1949, the twenty-one remarks with the section marking “|...|” were copied into in MS168. Rothhaupt states that researchers have ignored the fact that Wittgenstein “eigenhändig über den Zeitraum 1929 bis 1951 mit insgesamt 488 Sektionen eine ansehnliche [...] Sammlung von “Allgemeinen Bemerkungen” bzw. “General Remarks” [anlegte]” (Rothhaupt 2017,121–122). On 16 January 1949, Wittgenstein himself started to compile just such a selection in MS168, “nicht nur in der Form der in den Manuskripten hinzugefügt Sektionsmarkierungen, sondern auch physisch als eigenes Manuscriptcorpus” (Rothhaupt 2017,125). In short: “MS168 ist ein von Wittgenstein begonnenes Kondensat seiner eigenen Doppelauswahl” (Rothhaupt 2017,131). Thus in Rothhaupt’s view, the smaller selection of “general remarks” in von Wright’s *Vermischte Bemerkungen* can and should be expanded to reflect the two selections (identified by means of “O” and “|...|”) that Wittgenstein himself made, whereby “die Materialbasis als auch die Beispielfülle noch einmal extensiviert und intensiviert [wäre]” (Rothhaupt 2017,130–131).

Like Schulte, Erbacher mentions MS168 only in passing. The context is Erbacher’s reconstruction of “von Wright’s reasons and motives for creating and publishing” the selection of “General remarks” from the *Nachlass* that we now know as *Vermischte Bemerkungen* (Erbacher 2017,81). For von Wright,

Vermischte Bemerkungen tells “us more than any other written source about Wittgenstein’s intellectual character and view of life, and also how he regarded his relationship with his time” (von Wright 1982b,203). Erbacher comments on Rothhaupt’s assumptions: “Whether Wittgenstein himself ever contemplated composing a book of remarks on broader cultural topics is an issue of debate. While working on and revising his texts, Wittgenstein marked sections that may be read as remarks on culture, and he transferred such remarks into fair copies (e.g. Ms 168; cf. Rothhaupt 2017). But scholars have questioned whether the philological evidence suffices to conclude that Wittgenstein worked on a selection of cultural remarks as part of his philosophical work or even that he planned a book containing such a selection [...]. In any case, with regard to Wittgenstein’s literary executors there is no indication that they thought the selection in CV would resemble anything Wittgenstein would have wanted to publish as his contribution to philosophy” (Erbacher 2017,81).

Both Rothhaupt and Erbacher comment on MS168 in the context of *Vermischte Bemerkungen*. Schulte, by contrast, suggests that MS168 has a biographical motivation. As for Denis Paul, who is also acquainted with the notebook, there may be yet another connection. Paul writes: “I knew it [MS168] in 1952 and called it “the St. John’s Street Aesthetics” (Elizabeth Anscombe lived at No. 27).” He adds: “It was Wittgenstein’s attempt to compile a collection of aesthetic remarks from his recent notebooks.” The only justification that Paul gives for characterising MS168 as a “compendium on aesthetics” is to highlight the remark on the value of Mahler’s music (Paul 2007,284-285). In brief, Paul regards MS168 as a draft for a collection of remarks that is intended to illuminate some of Wittgenstein’s reflections on aesthetics, a project that he did not, however, pursue any further. The concluding remark in MS168 even indicates Wittgenstein’s reason for “giving up that selection attempt” (Paul 2007,283). Wittgenstein writes: “Rosinen mögen das Beste an einem Kuchen sein; aber ein Sack Rosinen ist nicht besser als ein Kuchen; & wer im Stande ist, uns einen Sack voll Rosinen zu geben, kann damit noch keinen Kuchen backen, geschweige, daß er etwas besseres kann. Ein Kuchen, das ist nicht gleichsam: verdünnte Rosinen” (MS168,6v).

“It strikes me that ...”

In the literature on Wittgenstein, the remarks in MS168 have been mentioned and discussed in various contexts and in relation to their inclusion in *Vermischte Bemerkungen*. Neither Schulte, Rothhaupt, Erbacher nor Paul considers and analyses the selection of remarks as a discrete and integral collection. For my own part, I would echo the discussions already mentioned in finding that MS168 does not in and of itself allow us to draw any strong or definitive conclusions. Allow me at this point to mention a few reflections to which my own work with this collection has given rise.

A number of points should be emphasised here. Firstly, that MS168 can and should be regarded *a selection of a selection*. A selection of a selection, whereby the process of reproducing the remarks for each new selection is associated with omissions, additions, changes of wording and new underlinings. Secondly, that many remarks in MS136 and MS137 that are marked with “|...|” are not included in MS168, even though they appear at first glance to be relevant to the topics addressed in the collection. Further, it is worth noting that a number of other remarks that would seem relevant to the themes in MS168 are *not* marked with *eine Anfangsstrich-Endstrich-Markierung*. The idea that MS168 should be regarded as a selection of a selection strengthens the assumption that MS168 was compiled with some specific or special aim in mind. If this is the case, then the notebook should not necessarily be associated with the collection of “allgemeine Bemerkungen” that reflects the markings “O” and “|...|”, as Rothhaupt assumes. Seen from this perspective, the remarks should not necessarily be seen as an indication that Wittgenstein – to quote Schulte again – “dedicated his spare time to composing his own *Aphorismen zur Lebensweisheit*” (Schulte 1999,80).

One obvious question when seeking to understand MS168 is whether we can identify any clear overarching themes in the selected remarks. My answer is: No and Yes! Because, on the one hand, and as already mentioned, MS168 touches on many different topics, while on the other, two topics do stand out as particularly prominent: firstly, the grammatical investigations with respect to aesthetics, and secondly, the autobiographical reflections. Although Paul’s title for the collection – “The St. John’s Street Aesthetics” – might seem well chosen, it is nevertheless too narrow, because it isn’t only aesthetics that Wittgenstein is concerned with in MS168. While aesthetics are central to the

first two remarks, these are followed by the enquiries that address the “metaphysical” aspect of Busch’s drawings, Schiller and “poetische Stimmung” (MS168,5r), and “vergleichen” as a methodological tool when assessing an object’s aesthetic value.

The dominant question in these aesthetic reflections concerns the preconditions for an aesthetic “Eindruck”. This presupposes Wittgenstein’s critique and rejection of the notion that aesthetics is a branch of psychology that can provide causal explanations for our aesthetic experiences. In the first remark of MS168, Wittgenstein extends this critique so that it also applies to the recounting of dreams. Aesthetic impressions and “die Traumerzählung” are both embedded in “unserm gegenwärtigen Leben” and should, as such, be considered against the background of certain activities, a certain culture or form of life: “Ich will sagen, daß es nicht notwendigerweise Sinn haben muß, von einem Auffinden ihrer Ursache zu reden” (MS168,1r). In MS137 Wittgenstein notes that: “Unser Problem ist kein kausales sondern ein Begriffliches” (MS137,139b).

The strong impression made on us by dreams, the plays of Shakespeare, the drawings of Busch is due not least to the fact that each of them creates “ihre eigene Sprache & Welt”. These languages and worlds often strike us as “ganz unrichtig, absurd, zusammengesetzt, & doch ganz richtig” (MS168,1r). Even so, they make an impression on us that seems to demand interpretation. Wittgenstein adds: “Warum? Ich weiß es nicht. Und wenn Shakespeare groß ist, wie von ihm ausgesagt wird, dann muß man von ihm sagen können: Es ist alles falsch, stimmt nicht – & ist doch ganz richtig nach einem eigenen Gesetz” (MS168,1v).

In addition to these grammatical enquiries, the collection touches on other themes, including the general condition of human life and human powerlessness: one finds oneself in this world, “in der die Dinge rutschen, zerbrechen, Unheil anstiften. Und sein Körper ist natürlich eins von den Dingen” (MS168,2v). In this context, Wittgenstein mentions the constraints and injunctions of religion and offers a Kierkegaardian account of the grammar of religious faith. He introduces this account by saying: “Es kommt mir vor, als könne ein religiöser Glaube nur ein leidenschaftliches Sich-entscheiden für ein Bezugssystem sein” (MS168,4r). This is further contextualised with remarks

about philosophical debate and the temptations one faces in doing philosophical work. In one remark, he touches on the question of the function that dreams might serve by juxtaposing “Nachtträume” and “Tagträume” (MS168,2v). Another remark is concerned with the question of whether upbringing can be viewed as a way of giving life a Gestalt.

The other theme that is given prominence in MS168 is the autobiographical. This is most clearly evident in the longest of the remarks about Mahler. The autobiographical aspect is also noticeable in the foregoing remarks, which include, among other things, a discussion of the stylistic devices or figures that Wittgenstein forbids himself to use in his writing, together with a reminder to maintain a sense of *Anständigkeit* in his thought, and to focus on the qualities and content of ideas rather than their originality and value. He writes: “Es ist wohl möglich, daß meine Gedanken ihren Glanz dann nur von einem Licht, das hinter ihnen steht, empfangen. Daß sie nicht selbst leuchten” (MS168,5r-5v). Given the prominence of this autobiographical tone at many points in MS168, it is striking that Wittgenstein leaves out one particular autobiographical reflection – namely the words “Ich denke an Kraus & seine Aphorismen, aber auch an mich selbst & meine philosophischen Bemerkungen” (MS136,92a) – when transcribing the concluding remark from MS136. On the other hand, it is not surprising that Wittgenstein adds the clarifying formulation: “[Über meine philosophischen Bemerkungen.]” (MS168,4r) as a prelude to the remark from MS137: “Im Geistigen lässt sich ein Unternehmen meistens nicht fortsetzen; soll auch gar nicht fortgesetzt werden. Diese Gedanken düngen den Boden für neue Gedanken” (MS137,122a).

The note on Mahler’s music can also be viewed as part of the collection’s autobiographical material. The discussion of Mahler’s symphonic works begins with the following thematic statement: “Der Mensch kann alles Schlechte in sich als Verblendung ansehen” (MS168,5v). Following a route that takes him via comments on the value of Mahler’s compositions – “ich glaube, daß Mahlers Musik nichts wert ist” (MS168,5v) – and the circumstances that allow one to judge the value of one’s own work, Wittgenstein eventually arrives at the central theme of this section: In what way and under what conditions can philosophy continue to be practised after the collapse of a tradition – without borrowing, embellishing or repurposing inherited and worn-out material.

Here he identifies with the situation of Mahler, but considers himself more honest than the composer. The final line of the remark in MS136, “Unbestechlichkeit ist alles!” (MS136,111a), is omitted from the transcription in MS168. The remark to the effect that “[wer] eine Tradition nicht hat [ist] wie ein unglücklich Verliebter” (MS168,3v-4r) and the earlier one about the distinction between “Genie” and “Talent” (MS168,4v-5r) can be regarded as autobiographical statements that Wittgenstein arrives at via Mahler. Valuable music and honest thinking presuppose courage in practising self-insight, courage to assess one’s own talent and to confront self-deception, one’s actual situation and contemporary world. Szabados has pointed out that in this remark Wittgenstein returns to a train of thought he had already explored in MS120. On 19 February 1938 Wittgenstein noted in code:

“Sich über sich selbst belügen, sich über die eigene Unechtheit belügen, muß einen schlimmen Einfluß auf den Stil haben; denn die Folge wird sein, daß man in ihm nicht Echtes von Falschem unterscheiden kann. So mag die Unechtheit des Stils Mahler zu erklären sein & in der gleichen Gefahr bin ich.

Wenn man vor sich selber schauspielert, so muß der Stil davon Ausdruck sein. Er kann dann nicht der Eigene sein. Wer sich selbst nicht kennen will, der schreibt eine Art Betrug.

Wer in sich selbst nicht hinunter steigen will, weil es zu schmerzhaft ist, bleibt natürlich auch mit dem Schreiben an der Oberfläche. (Wer nur das Nächstbeste will, kann doch nur das Surrogat des Guten erreichen.)” (MS120,72v-73r).

References

Erbacher, Christian (2017) ““Among the omitted stuff, there are many good remarks of a general nature” – On the Making of von Wright and Wittgenstein’s *Culture and Value*”, *Sats. Northern European Journal of Philosophy* 18, 79-113.

Keicher, Peter (2014) “Die *Vermischten Bemerkungen* und das Manuskript MS 168”, in: Sonja Rinofner-Kreidl & Harald A. Wilsche (Hrsg.): *Analytische und Kontinentale Philosophie: Perspektiven und Methoden. Beiträge des 37. Internationalen Wittgenstein Symposiums* Vol. 22, Kirchberg am Wechsel: ALWS, 150-153.

Paul, Denis (2007) *Wittgenstein’s Progress 1929-1951*, Bergen: Publications from the Wittgenstein Archives at the University of Bergen No 19.

Pinch, Trevor & Swedberg, Richard (2013) “Wittgenstein’s visit to Ithaca in 1949: on the importance of details”, *Scandinavian Journal of Social Theory* 13, 2-29.

Rothhaupt, Josef G.F. (2013) “Zur Philologie des KRINGEL-BUCHES und seiner Verortung in Wittgensteins Œuvre”, in: Josef G.F. Rothhaupt & Wilhelm Vossenkuhl (Hrsg.): *Kulturen und Werte. Wittgensteins ”Kringel-Buch” als Initialtext*, Berlin: De Gruyter, 3-76.

Rothhaupt, Josef G.F. (2017) “Wittgensteins “General Remarks”. Auswahlen von Bemerkungen in oder aus seinem Nachlass”, *Wittgenstein-Studien* Vol 8, 103-136.

Schulte, Joachim (1999) “What Wittgenstein Wrote”, in: Rosaria Egidi (Hrsg.): *In Search of a New Humanism. The Philosophy of Georg Henrik von Wright*, Dordrecht: Kluwer Academic Publishers, 79-91.

Schulte, Joachim (2013) “Did Wittgenstein Write on Shakespeare?”, *Nordic Wittgenstein Review* 2, 7-32.

Szabados, Béla (2007) “Wittgenstein Listens to Mahler: How to Do Philosophy and Compose Music in the Breakdown of Tradition”, *Dialogue* XLVI, 91-113.

von Wright, Georg H. (1966) *A COLLECTION OF REMARKS BY LUDWIG WITTGENSTEIN*, The von Wright and Wittgenstein Archives at the University of Helsinki (WWA), Finland.

von Wright, Georg H. (1982a) “The Wittgenstein Papers”, in: Georg H. von Wright: *Wittgenstein*, Oxford: Basil Blackwell, 35-62.

von Wright, Georg H. (1982b) “Wittgenstein in Relation to his Times”, in: Georg H. von Wright: *Wittgenstein*, Oxford: Basil Blackwell, 201-216.

Waismann Says only Fools Rush In: The Waismannian Mechanism in Conceptual Engineering

Jennifer Whyte (Vienna, Austria)

Abstract

Much philosophical attention has recently been paid to Conceptual Engineering - or, the intentional modification of concepts by philosophers in order to achieve epistemic, ethical, or political ends. The most controversial aspect of conceptual engineering are its aims, but there is also much debate over the mechanism by which the process could even be carried out. In this paper, I propose a mechanism derived from the work of Friedrich Waismann by which intentional conceptual amelioration can - and does - occur. I will then argue that this account of conceptual engineering solves a number of standing problems in the literature, but in doing so, places some key checks on the ambition of conceptual engineering as a project.

In this paper I identify what I call the ‘Waismannian Mechanism’ for conceptual engineering, drawn from the work of Friedrich Waismann. Waismann’s role as a Wittgenstein collaborator, both during and after his time in the Vienna Circle, is widely known, but Waismann’s own philosophical work has (until recently) received little attention[1]. I believe that the context of conceptual engineering gives ample reason to change that. The Waismannian Mechanism is the method of altering the extension and intension of a concept by making deliberate decisions at the point at which the concept’s open texture is discovered. I argue that this mechanism solves the Implementation Problem and suggests solutions to other problems in the conceptual engineering literature. Though the Waismannian Mechanism only accounts for a small proportion of conceptual engineering, it is reliable and historically well-attested. It is my hope that focusing on a single mechanism of conceptual engineering, rather than putative outcomes, will provide the prospective conceptual engineer with a new tool in their conceptual toolkit. The social ends of conceptual engineering are of sufficient importance that they deserve to be approached with clear, reliable, and well-defined means.

Though the term ‘conceptual engineering’ is new, the idea is not. We can think of ‘conceptual engineering’ as a term that has itself been engineered to draw our attention towards an aspect of philosophical practice that deserves greater focus: the role of a philosopher in shaping - rather than merely describing - the concepts of the language they use. Conceptual engineers are those who think that one of the goals of philosophy is to make those concepts better - either

better at fulfilling their existing linguistic functions, or better when judged according to broader social, pragmatic, or epistemic goals. The conceptual engineer is not content to analyze language when they could be intervening in it. Paradigmatic projects of conceptual engineering include Sally Haslanger's ameliorative proposals for social categories (eg. woman) (Haslanger, 2000), Amie Thomasson's therapeutic revision of metaphysical concepts (Thomasson, 2025), and Carnap's method of explication.

Conceptual Engineering raises a number of distinctive philosophical questions. The most familiar is the question of how synonymy can be preserved over the change of a concept's meaning (sometimes termed 'Strawson's Challenge') but the most pressing is the question of how – or whether - such changes can be implemented (Cappelen, 2020). If conceptual content cannot be intentionally altered, or cannot be altered in a manner we can control, the whole project is pointless. So, the task of finding mechanisms by which concepts can be intentionally altered should be at the forefront of the putative engineer's mind, with questions of synonymy answered once such a mechanism has been found. Let us put one on the table.

The Waismannian Mechanism

Friedrich Waismann first defines the idea of Open Texture in print in his 1945 essay 'Verifiability'. Open texture (porosität) is a property that Waismann thinks is found in all empirical concepts. If a concept exhibits open texture, there is some possible future phenomenon that the concept would fail to classify as either falling within or without it. No matter how precisely the concept is specified, it simply contains no conceptual resources that can settle the question of whether the new phenomenon counts as part of the concept or not. Waismann illustrates this phenomenon with the concept GOLD.

"The notion of gold seems to be defined with absolute precision, say by the spectrum of gold with its characteristic lines. Now what would you say if a substance was discovered that looked like gold, satisfied all the chemical tests for gold, whilst it emitted a new sort of radiation? 'But such things do not happen.' Quite so; but they might happen, and that is enough to show that we can never exclude altogether the possibility of some unforeseen situation arising in which we shall have to modify our definition. Try as we may, no concept is limited in such a way that there is no room for any

doubt. We introduce a concept and limit it in some directions; for instance, we define gold in contrast to some other metals such as alloys. This suffices for our present needs, and we do not probe any farther. We tend to overlook the fact that there are always other directions in which the concept has not been defined. And if we did, we could easily imagine conditions which would necessitate new limitations. In short, it is not possible to define a concept like gold with absolute precision, i.e. in such a way that every nook and cranny is blocked against entry of doubt. That is what is meant by the open texture of a concept.” (Waismann, 1945)

The key upshot of open texture is that, when the open texture of a concept is discovered, it leaves the language-user with a decision to make. The decision is between three options, which are exhaustive. First, the aberrant phenomenon can be accepted into the concept (the radioactive gold is simply gold with an unusual property). Second, the aberrant phenomenon can be excluded from the concept (the radioactive metal looks like gold, but it isn’t gold). Third, the language user can give up on using the concept gold in a precise way, and allow it to become vague. The choice between these three options must be made, but it also must be made freely. The concept itself has, by definition, no resources within it that could settle the question logically. We may bring any reasons we choose to bear on our decision – practical, aesthetic, ethical – but we do not get to choose whether the decision gets made. It must be made.

I claim that this decision process is, by any reasonable standard, conceptual engineering. No matter which choice we make in the face of open texture, the concept that we will have after that decision is made will be different from the one we had before. If we take the first option, and admit the aberrant phenomenon, the extension of the concept will have changed. If we take the second option, and bar the aberrant phenomenon, we must introduce new content to the concept to justify that exclusion, and so the intension will have changed. If we take the third option, and allow the concept to become vague, the character of the concept will have changed (this is, I claim, what Waismann means when he says that Open Texture is not vagueness, but it is “something like the possibility of vagueness” (Waismann, 1945)). This decision-making process is what I term the ‘Waismannian Mechanism’ of conceptual engineering.

The Waismannian Mechanism is one possible answer to the challenge of implementation. Conceptual change according to this model is historically very well-attested (see (Makovec, 2025) for a plethora of examples). This method of conceptual engineering is, in essence, an intentional intervention into a natural process of language evolution – less akin to engineering a building, and more akin to building embankments to divert the path of a river. We do not get to choose that the river flows down from the mountain. We do get to choose where it flows from there. In fact, we must.

Using the Mechanism

Waismann initially deployed the notion of open texture towards empirical concepts, because the context in which he proposed it was verificationism. Scientific concepts are clearly prototypical in Waismann's thinking about open texture – his go-to examples are the aforementioned gold and its fellow candidate for natural kind status, cat. However, Waismann does not think that open texture is only a feature of scientific concepts. In his later discussion of what he terms Language Strata, he argues that many different parts of language are characteristically open[2]. Where does this leave the social and political concepts with which conceptual engineers are typically concerned? Though the literature on conceptual engineering has expanded enormously over the past decade, the main focus remains on the substantive ethical and political project of designing better social concepts for the sake of a better world. Concepts like marriage and woman are the typical suggested loci of intervention (or in the case of marriage, of successful past intervention). If these concepts are to be modifiable with the Waismannian Mechanism, two things must be the case: first, they must exhibit open texture; and second, a novel phenomenon that exposes that open texture must be present.

I would argue that, for Waismann, neither woman nor marriage exhibits open texture – or at least not woman or marriage *simpliciter*. Recall that, for a concept to exhibit open texture, it must first be reasonably precise. Waismann distinguishes open texture from vagueness by this criterion. If a concept is used in a variable way, like bald or pink or beautiful, where one speaker saying 'that man is bald' and another saying 'that man is not bald' is not a contradiction and neither must be wrong, that concept is vague, not open. Open texture arises in the contexts in which we are not satisfied to be vague. For Waismann, this is a difference of Language Strata – a domain in which the

majority of concepts exhibit open texture is a different stratum than one in which those concepts do not. Each stratum has its own characteristic logical, alethic, textural, and evidential features. This means that the Waismannian Mechanism is more suited to some domains than others.

Let us take marriage as our case[3]. There are many different senses in which people can be said to be married. Most states define a legal notion of marriage particular to their state. Many religions do the same. The colloquial sense of the term varies from place to place, and its relation to these more formally-established definitions also vary. It is neither surprising nor contradictory to say of a couple of your friends who became legally married solely for tax purposes that ‘They’re married, but they’re not *married*’. Similarly, a member of a church that does not recognize same-sex unions can, without contradiction or irrationality, say of a same-sex couple that ‘they are legally married but I do not consider them to be married’. However, that kind of statement applied internally to the legal or religious context would no longer be rational. When presented without further qualification, marriage is not a concept that exhibits open texture, because it is vague. It is not a reasonable target for the Waismannian Mechanism. However, many religious concepts of marriage are precisely specified and exhibit open texture, and every legal concept of marriage does as well. In fact, one can argue that the Waismannian Mechanism was part of the way the legal concept of marriage changed in Canada.

Between the years 2003 and 2005, most of Canada’s provinces and one of its territories independently reached provincial/territorial Supreme Court decisions that same-sex marriage must be allowed under the protections of the Charter of Rights and Freedoms. This put the Federal government, which did not recognize same-sex marriages, in a somewhat awkward position: either these provincial marriages had to be included in the federal concept of marriage (which would require changing the definition of that concept), or a new kind of marriage – provincial -but-not-federal marriage – would have to be introduced. The option to allow the term ‘marriage’ to become vague was not open to the Federal Government, which has a mandate to avoid vagueness. The Parliament of Canada was in precisely the situation of forced-but-free choice characteristic of the Waismannian Mechanism. The eventuating Civil Marriage Act (S.C. 2005, c. 33) decided in favour of the first option: there was to

be no special category of provincial-but-not-federal marriage, and same-sex marriage would be incorporated into the legal definition of marriage. The Act cites a number of justifications for the choice they made, including ethical considerations and other legal precedent. The most relevant justification for our purposes is this one:

“WHEREAS the Supreme Court of Canada has determined that the Parliament of Canada has legislative jurisdiction over marriage but does not have the jurisdiction to establish an institution other than marriage for couples of the same sex;” (Government of Canada, 2005)

In which the Supreme Court explicitly rules out Option 2, of creating a new kind of marriage-like-concept instead of incorporating same-sex marriage into the existing institution of marriage.

The characterization of the legalization of same-sex marriage in Canada in terms of the Waismannian Mechanism is incomplete, of course – it does not include the long legal history of the case or the many decades of activist effort that preceded it. However, I do think this case shows a way in which the use of the Waismannian Mechanism can produce the change that the literature on conceptual engineering seeks.

Conclusions

Engineers carry many tools, and there is no reason conceptual engineers should carry fewer. My goal in this essay was to characterize one such tool, so as to draw the attention of the prospective conceptual engineer to its function and operation. Conceptual change according to the Waismannian Mechanism happens naturally as language evolves, but even in its natural habitat, it is inescapably bound up in the value judgements of agents. Philosophers who wish to alter concepts that feature open texture should, therefore, seek that open texture out. We should make novel phenomena manifest so that our concepts cannot remain the same, and provide reasons for why they should change in the manner we think is best. These changes will be small, for the most part, but a small change can still make a big difference. Conceptual Engineering characterized by the Waismannian Mechanism is opportunistic and active. We cannot force changes onto concepts that are functioning without trouble. We also may not conceptually engineer from the armchair – the domains in which the Waismannian Mechanism can be used are typically

technical domains with their own distinctive reasons and methods of reasoning. To actively participate requires patience, keen observation, and sensitivity to all relevant concerns. After all, Waismann says only fools rush in.

[1] For further reading, see (Makovec & Shapiro, 2019).

[2] A rare counterexample is his insistence that mathematical concepts are not open-textured. This claim has come under scrutiny in recent years. In several publications, Stuart Shapiro and Craig Roberts have argued that mathematical concepts do exhibit open texture (Shapiro, 2006; Shapiro & Roberts, 2019, 2021). Fenner Tanswell has also argued that open texture can be used as a distinctive method for conceptual engineering within mathematics (Tanswell, 2018).

[3] With no loss of generality. All relevant features of the concept marriage also apply to woman. It is, however, more expedient for the purposes of demonstration to speak of a past case, rather than an ongoing one.

References

Cappelen, H. (2020). Conceptual Engineering: The Master Argument. In A. Burgess, H. Cappelen, & D. Plunkett (Eds.), *Conceptual Engineering and Conceptual Ethics* (1st ed., pp. 132–151). Oxford University Press/Oxford. <https://doi.org/10.1093/oso/9780198801856.003.0007>

Government of Canada. (2005, July 20). *Consolidated federal laws of Canada, Civil Marriage Act*. <https://laws-lois.justice.gc.ca/eng/acts/c-31.5/page-1.html>

Haslanger, S. (2000). Gender and Race: (What) Are They? (What) Do We Want Them To Be? *Noûs*, 34(1), 31–55. <https://doi.org/10.1111/0029-4624.00201>

Makovec, D. (2025). Open Texture in Science and Philosophy. In E. Hienrich-Ramharter, A. Pilcher, & F. Stadler (Eds.), *100 Years of 'Tractatus Logico-Philosophicus'—70 Years after Wittgenstein's Death*. De Gruyter.

Makovec, D., & Shapiro, S. (Eds.). (2019). *Friedrich Waismann: The Open Texture of Analytic Philosophy*. Springer International Publishing. <https://doi.org/10.1007/978-3-030-25008-9>

Shapiro, S. (2006). Computability, Proof, and Open-Texture. In *Church's Thesis After 70 Years* (pp. 420–455). De Gruyter. <https://www.degruyter.com/document/doi/10.1515/9783110325461.420/html>

Shapiro, S., & Roberts, C. (2019). Open Texture and Analyticity. In D. Makovec & S. Shapiro (Eds.), *Friedrich Waismann* (pp. 189–210). Springer International Publishing. https://doi.org/10.1007/978-3-030-25008-9_9

Shapiro, S., & Roberts, C. (2021). Open Texture and Mathematics. *Notre Dame Journal of Formal Logic*, 62(1). <https://doi.org/10.1215/00294527-2021-0007>

Tanswell, F. S. (2018). Conceptual engineering for mathematical concepts. *Inquiry*, 61(8), 881–913. <https://doi.org/10.1080/0020174X.2017.1385526>

Thomasson, A. (2025). *Rethinking Metaphysics*. Oxford University Press.

Waismann, F. (1945). Verifiability. *Proceedings of the Aristotelian Society*, 19, 100–164. <https://doi.org/10.2307/2267243>

Why feminist epistemology needs feminist aesthetics. The feeling of unease as an example

Katrin Wille (Hildesheim, Germany)

Abstract

Feminist epistemology's claim to be critical of inequitable knowledge practices, whilst concomitantly establishing new knowledge practices, necessitates an examination of the feeling of unease. This feeling frequently serves as a catalyst for critical examination of knowledge practices, including the attribution or denial of credibility to individuals as sources of knowledge. Unease is an immediately perceptible situational feature that requires aesthetic forms of analysis. In this context, aesthetics is not to be understood in the sense of a theory of art, but rather as a theory of qualitative perception, as also emphasised in the pragmatist tradition. The aesthetic is thus considered a fundamental dimension of all experience, rather than a domain that can be confined to specific aesthetic experiences or artistic activities and their reception. This article therefore calls for the development of a critical feminist aesthetics following the pragmatist theoretical framework in order to provide analytical means for aesthetic-epistemic research and criticism. In this context, the feeling of unease functions as a kind of compass for navigating the fields of feminist epistemology and pragmatist aesthetics.

Introduction

Feminist epistemology, purporting both to critique inequitable knowledge practices and establish alternative knowledge practices, calls for an examination of the feeling of unease. In this context, the feeling of unease frequently serves as a catalyst for critical examination of knowledge practices, including the attribution or denial of credibility to individuals as sources of knowledge. Unease frequently arises in individuals who are ascribed less credibility due to their possessing characteristics deemed irrelevant in the given context. In this sense, the feeling of unease can function as an indicator of epistemic injustice and is therefore a relevant phenomenon for critiquing the forms in which it occurs.

Unease is a directly perceptible situational feature and exemplifies a type of relation to the world in contradistinction to knowledge relations. Consequently, alternative analytical methods are necessary to elucidate its relevance to knowledge relations. The pragmatist aesthetic, which prioritises the felt qualities of situations, offers a suitable theoretical framework for examining the intrinsic character and relevance of feelings such as unease. In the context of pragmatist theory, aesthetics is not conceived as a theoretical framework for art. Rather, the aesthetic is regarded as a fundamental

dimension of all experience, not as confined to specific aesthetic experiences or artistic activities and their reception. The aesthetic dimension refers to the immediate involvement in meaningfulness.

This conception of aesthetics provides the necessary framework for elucidating the significance of the feeling of unease for a critical feminist epistemology. This paper therefore calls for the development of a critical feminist aesthetics modelled on the pragmatist theoretical framework in order to provide the analytical means for aesthetic-epistemic research and critique. Unease thus serves as a kind of compass, guiding us through the various domains of feminist epistemology and pragmatist aesthetics. In what follows, the significance of the feeling of unease for feminist epistemology and its limits are first worked out. Then, the profile of pragmatist aesthetics is delineated, with the objective of relating these analytical means to the feeling of unease. An appropriate treatment of unease is an essential component of feminist-pragmatist aesthetics.

1. The Significance of the Feeling of Unease within the Context of Feminist Epistemology

In the domain of feminist epistemology, a significant concern is the range of constraining and facilitating factors that influence the articulation of the experiences of individuals who, in view of their social standing, are assigned diminished credibility in the generation of knowledge (Anderson 1995). This analysis gives rise to profound critical transformations of traditional epistemology into social epistemology. A pivotal element in this is the conversion of the abstract and individualistically conceived notion of the knowing subject into a situated and relational entity. This shift has far-reaching implications for the nature of knowledge, as evidenced by the seminal contributions of Harding (1986), Hartsock (1983), and Longino (1990). The incorporation of situatedness and relationality facilitates the analysis of practices of valorisation and devaluation of subjects, objects and modes of knowledge, allowing forms of epistemic injustice to be studied (Fricker 2007; Dotson 2014; Hänel 2024). This encompasses the conditions of attribution of epistemic authority, as well as the examination of epistemic privileges based on specific standpoints. Consequently, strategies employed for the maintenance of power, epistemic disparities, and survival in subordinate roles can be identified (Hill Collins 1990/2008). A strategy for maintaining power

that has been the focus of much discussion consists of practices of deliberate ignorance, whereby proposals from marginalised groups are forestalled by displacing forms of language and conceptual categories (Pohlhaus 2012).

This is founded on the objective of diminishing impediments to voicing concerns and of pluralising and democratising participation in the establishment of epistemic authority. However, an exclusive focus on epistemology disregards specific dimensions of experience that remain unarticulated but nonetheless play a significant role in questions of knowledge, even if they elude theoretical access. One such dimension of experience is the feeling of unease, characterised by a diffuse sensation of something being amiss or absent, distorted or concealed, of a situation that is not equally habitable for all participants. This phenomenon manifests in a wide variety of contexts, yet it is of particular significance for feminist and anti-racist concerns. Feelings of unease frequently serve as the impetus for the pursuit of appropriate articulations and the acquisition of knowledge. A substantial corpus of literature exists that employs vivid and compelling descriptions of unease (Perkins Gilman 1892/1999). Experiential reports constitute a substantial source of material, utilised by Miranda Fricker in her work *Epistemic Injustice. Power & the Ethics of Knowing*, as they provide dense descriptions of unease, but also emotions such as fear and anger, which can give rise to a collective search for expression. A significant example can be found in the seventh chapter of Fricker's book on hermeneutic injustice, in which she quotes an extensive passage from Susan Brownmiller's memoir of the second American women's movement (Fricker 2007: 150). This passage describes Carmita Wood's experiences and the collective search for a suitable expression for the experience of sexualisation. Carmita Wood's feeling of unease in the workplace, which gradually escalates into a threat, illustrates the significant importance of unease for feminist epistemology.

Since unease is not an epistemic relationship to the world, it is not possible for a purely epistemological perspective to render it transparent as an experience. Consequently, such a perspective entails a reductionist interpretation of unease. In this perspective, the feeling of unease is reduced to an epistemically deficient form of not yet being able to articulate or comprehend. An appropriate examination of situational qualities such as unease also reveals the productive function that sensations have not only before but also during

the search for expression. Furthermore, perceptions of qualitative features function to orient processes of articulation until the end. This productive function becomes very clear in the memoir cited by Fricker. The process of collective articulation culminates in a shared sense of coherence, as exemplified by the following assertion: “*Sexual harassment!* Instantly we agreed. That's what it was” (Fricker 2007: 150).

In order to undertake a comprehensive analysis of the role of qualitative features in processes of articulation, it is necessary to adopt a more expansive disciplinary perspective. Qualities must be understood not only as precursors to knowledge, but as playing a decisive role always and everywhere. In order to demonstrate this, it is necessary to adopt an aesthetic conception that facilitates this understanding. In the following section, I draw upon pragmatist aesthetics to offer a theorisation of the significant feeling of unease as a final step.

2. The aesthetic dimension in meaning-making processes

From a pragmatist standpoint, the aesthetic is considered omnipresent, not a specific type of experience or a phenomenon of art that can subsequently be theorised within the philosophical discipline of aesthetics (Johnson 2018: 6). Instead, aesthetics focuses on the structures, processes, qualities and sensations that make meaningful experiences possible (Johnson 2018: 210).

Pragmatist thinking is distinguished by its examination of the phenomenon of meaning at its most fundamental levels. In this sense, the perception of qualities is a characteristic of all living things and life is in the most general sense meaning-making. Living organisms interact with their environment in a variety of ways, sensing what is and is not conducive to the preservation and development of life. This direct sensory-motor and sensory-motor-affective exchange is a constitutive feature of life. These processes enable organisms to affect and be affected. When individuals seek to slake their thirst, they engage in the muscular exertion necessary to lift and transport the glass of water to their mouths. Biological and material processes of this kind combine with social and historical conditions and develop a formative force for linguistic meaning. These complex feedback relationships and the formative power of bodily and material processes for all kinds of meaning comprise what is meant by the aesthetic dimension of experience. The concept of the qualitative, as

proposed by the pragmatist philosopher John Dewey, pertains to this aesthetic dimension of all experience (Dewey 1934/1987).

In the realm of pragmatist philosophy, qualities are conceived as modes of interaction through which organisms discern the significance of situations for themselves and their environment. These modes of interaction are direct, and the qualities of situations are felt. Dewey's idea is that we can grasp the quality of entire situations and that situations receive their specific character through such pervasive qualities. For instance, situations can be felt as frightening or exuberant. These qualities, in turn, influence human behavior and thought processes, and can occasionally provide the impetus to adopt a relation of knowledge toward them. In such knowledge relations, however, specific aspects of the situation are always foregrounded and others pushed into the background. The selection processes are controlled and accompanied by qualities, which orient selections in terms of appropriateness.

Building upon the work of Dewey, Matthias Jung (Jung 2023: 71) has developed a phase model of qualitative experience that provides a crucial corrective to interpretations that either neglect the role of the qualitative or confine it to the initial stages of processes of articulation. Jung elucidates the diverse functions attributed to the experience of qualities in all phases of such processes. Prior to their initiation, qualities emerge in every practical interaction, perceived as meanings relevant to life and facilitating the comprehension of situations experienced as contexts and their diverse situational potentials (1). In the event of a stagnation or deviation from established processes, or when confronted with a challenge, these qualities are perceived as situational, thereby facilitating the initiation of expression (2). Felt qualities serve a pivotal role in guiding and orienting the articulation of goals of action and the selection of means, thereby facilitating the progression of the process (3). Furthermore, felt qualities serve as indicators of the culmination of a process of articulation, manifesting as a sense of fulfillment, consummation, completeness, or coherence (4).

Two levels of interaction are particularly salient for situational qualities. On the one hand, these are direct physical exchange processes between organisms and their environment, for which the specific body logics of the organisms are decisive. In his pragmatic aesthetics, Mark Johnson accords particular

importance to verticality and scalar intensity (Johnson 2018: 18). The second level of interaction pertains to the coordination between organism and material-spatial environment. This coordination can be analyzed using the concept of affordance, which is a term coined by the perceptual psychologist James Gibson (Gibson 1979/2015).

In order to account philosophically for the significance of the feeling of unease, the practice of theorizing must change in such a way that the function of the aesthetic in thought and knowledge receives increased theoretical attention.

3. Unease as the subject of aesthetic-epistemic research

It is important to note that unease is conceptualised as a feeling, not an emotion or affect. The reason for this is that feelings are characterised by the interplay of bodily and affective aspects. In the domain of theories of embodiment and affect, various divergent proposals for conceptual differentiations have given rise to contentious debate. The subsequent discussion of this interplay is founded on the contributions of Virginia Woolf (Woolf 1938/2015) and Sara Ahmed (Ahmed 2017), from which significant insights into the feeling of unease can be gleaned.

If a situation is experienced as uneasy, this uneasiness is felt as a pervasive quality. Through the sensory-motor-affective direct exchange processes and interactions, those involved feel the significance of the situation as a whole. The situation as a whole becomes uneasy and uncomfortable, and the various aspects of the situation, from things it contains to spatiality itself and available modes of communication within it, are colored. Atmospheres also have this all-pervading, coloring character in a situation, and it is useful to look at the transitions between the two concepts. In this regard, Virginia Woolf's work offers a seminal reference, as she defines her critical literary task as the capture of atmospheres that are formative and, concomitantly, seem intangible due to their composition of numerous intricacies across multiple levels, a feat that would require ten or fifteen volumes to fully articulate (Woolf 1938/2015 :134-135).

A situation that is felt as uneasy is one in which those involved are unable to develop. Tension is observed within the body, manifesting as a shallower respiratory cycle and muscle contraction. The range of movement may be restricted, arm movements may be less expansive, and stance may not be as

stable. The gaze may become more evasive. In her critical reflection on diversity work, Sara Ahmed frequently alludes to feelings of discomfort and unease, emphasising the significance of the dimension of embodiment. Ahmed's assertion that feminism is sensorial signifies her commitment to underscoring the significance of bodies in the abstract and to the serious consideration of physical sensations. The metaphor of dwelling and inhabiting recurs strikingly often in her critical analyses, tapping into the situational quality of unease and discomfort (Ahmed 2012: 24-25). She identifies a pivotal distinction between individuals' ability to inhabit a situation, encompassing its spatial characteristics, objects, and institutional contexts (Ahmed 2017: 24-25). Individuals who find themselves in less favourable situations may experience a diminished sense of freedom to occupy, move and behave in a space. This alteration in their capacity to navigate and engage with their environment, both tangible and intangible, has a discernible impact on the affordances of their immediate surroundings. The consequences of this phenomenon are both visible and invisible, thus allowing for the concrete physical processes through which Ahmed's more metaphorical reference to living space can be delineated. She writes: "I often think through the politics of comfort through chairs: furniture is always good to think with. Think of how it feels to be comfortable: say you are sinking into a comfortable chair. Comfort is about the fit between body and object [...]" (Ahmed 2017: 122-123). If situations are experienced as uneasy or uncomfortable, one's own body does not fit well with the surroundings and the dispositional cues they offer become a problem for the body. Chairs, for instance, may be either oversized or undersized, hindering a body's ability to fully sink into them and providing little support or invitation to relax. Conversely, the bodies of some participants get lost in the armchairs, or are unable to find a comfortable position within them, instead moving restlessly from one corner to another, forward and back, as they are unable to locate a suitable, inviting or appropriate position. Ahmed's observations, made within the same theoretical framework, underscore this phenomenon. "Discomfort involves this failure to fit. A restlessness and uneasiness, a fidgeting and twitching, is a bodily registering of an unexpected arrival" (Ahmed 2017: 129).

Such feelings of unease can initiate processes of articulation with the aim of finding a language adequate to experience, and thus also open up the

possibility of analysing the conditions that lead to the situation, which only offers a comfortable place to some participants and not to others. However, this is not the only function of perceptions of qualitative features that make it necessary to process them using aesthetic means. The articulation of theoretical or political analyses does not supersede or replace the concept of quality-feelings. Instead, perceptions of quality and conceptual articulations enter into a reciprocal relation by which both are altered. The feeling of unease, therefore, serves as an important compass for many concerns of feminist epistemology, as it determines which conceptual proposals and analytical means are appropriate and which are not. Unease plays a significant role in the critique of unjust social and political conditions, as well as in theoretical work. This provides an opportunity to contribute to the further development of aesthetic-epistemic means to improve the theoretical comprehension of expressions such as these: “*Sexual harassment!* Instantly we agreed. That’s what it was.” (Fricker 2017: 150).

References

- Ahmed, Sara (2017) *Living a feminist life*, Durham/London: Duke University Press.
- Anderson, Elizabeth (1995) “Feminist Epistemology: An Interpretation and a Defense”, *Hypatia* 10:3, 50-84.
- Dewey, John (1934/1987) *Art as Experience*, in: Jo Ann Boydston (ed.) *John Dewey. The Later Works 1925-1953* Vol. 10, Carbondale, IL: Southern Illinois University Press.
- Dotson, Kristie (2014) “Conceptualizing Epistemic Oppression”, *Social Epistemology* 28:2, 115-38.
- Fricker, Miranda (2007) *Epistemic Injustice: The Ethics and Power of Knowing*, Oxford: Oxford University Press.
- Gibson, James J. (1979/2015) *The Ecological Approach to Visual Perception*, New York/London: Taylor & Francis.
- Hänel, Hilkje Charlotte (2024) *Epistemische Ungerechtigkeiten*, Berlin/Boston: de Gruyter.
- Harding, Sandra (1986) *The Science Question in Feminism*, Ithaca: Cornell University Press.
- Hartsock, Nancy (1983) “The Feminist Standpoint: Developing the Ground for a Specifically Feminist Historical Materialism”, in: Sandra Harding and Merrill
- Hintikka (eds.) *Discovering Reality: Feminist Perspectives on Epistemology, Metaphysics, Methodology, and the Philosophy of Science*, Dordrecht: D. Reidel, 283-310.

Hill Collins, Patricia (1990/2008) *Black Feminist Thought: Knowledge, Consciousness, and the Politics of Empowerment*, New York: Routledge.

Johnson, Mark (2018) *Aesthetics of Meaning and Thought. The Bodily Roots of Philosophy, Science, Morality, and Art*, Chicago/London: Chicago University Press.

Jung Matthias (2023) *Leben und Bedeutung. Die verkörperte Praxis des Geistes*, Berlin/Boston: de Gruyter.

Longino, Helen (1990) *Science as Social Knowledge: Values and Objectivity in Scientific Inquiry*, Princeton: Princeton University Press.

Perkins Gilman, Charlotte (1999) "The Yellow Wall-Paper", in: Denise K. Knight (ed.) *The Yellow Wall-Paper, Herland, and Selected Writings*, New York: Penguin Books, 166-182.

Pohlhaus, Gaile (2012) "Relational Knowing and Epistemic Injustice: Toward a Theory of Willful Hermeneutical Ignorance", *Hypatia* 27:4, 715-735.

Woolf, Virginia (1938/2015) "Three Guineas", in: Anna Snaith (ed.) *A Room of One's Own and Three Guineas*, Oxford: Oxford University Press, 87-215.

Criticism as Third-Order Change—A Feminist Examination of Wittgensteinian Practices

Yara Windmüller (Kiel, Germany)

Abstract

Wittgenstein's (later) philosophy may well be read as encouraging a conservative standpoint towards forms of life. Especially in his remarks on rule-following and in *On Certainty*, Wittgenstein implies that our practices create an unchangeable functional whole, which we must accept as providing a foundation for everything else. However, this is certainly not the only way to read Wittgenstein: Especially from a feminist standpoint, Wittgenstein seems to offer plenty of room for allowing constant development of practices, and even for substantial change amidst criticism that profoundly questions a given practice. Such change doesn't just happen accidentally, but is rather brought about by active intervention on the part of those practitioners who see the need for a rethinking of certain practices.

To make this thought plausible, I draw on Kristie Dotson's contributions on epistemic oppression, particularly on her concept of third-order change, which provides exactly the kind of reflection on the mechanisms of unjust practices that ideally leads to the change I have in mind. I argue that the implementation of third-order conceptual change by means of such critical awareness is itself a practice, and that, therefore, a Left Wittgensteinian and feminist-inspired reading of *On Certainty* is not only consistent with Wittgenstein's thinking, but also especially desirable where marginalized perspectives are concerned.

1 The Role of Tradition in Practice

Wittgenstein's (later) philosophy may well be read as encouraging a conservative standpoint towards forms of life: All practicing heavily relies on rule-following, and all rule-following is based on an existing set of rules that have been established over time in the past. This set of rules now seems more or less stable, as if notable changes to these practices were highly unlikely to be witnessed.

Especially in his remarks on rule-following and, later, in substantial parts of *On Certainty* (OC), Wittgenstein implies that some things seem to be just there—like a foundation that everything else is built on: “*This* is how one calculates. Calculating is *this*. What we learn at school, for example.” (OC 1969: 47; italics in the original); “(...) [M]y picture of the world (...) is the inherited background against which I distinguish between true and false.” (OC 1969: 94; see also OC 1969: 87–97, 156, 559). The conservative thought lies in taking these remarks quite literally, suggesting that change, if it happens at all, will only result from slow-paced, organically caused alterations to or within these solid foundations

(Nyíri 1976: 509). The traditionalist takes the world to be already finely calibrated, not leaving much room to encourage substantial development (Queloz and Cueni 2021: 764).

And indeed, many passages of OC may be interpreted that way. Take, for example, Wittgenstein's analogy of the riverbed (OC 1969: 96–99), where he describes the riverbed as providing the boundaries, the system of (more or less) invariable certainties (*Gewissheiten*), which alone generate the framework for (mere) everyday knowledge to come into being, knowledge that is more dynamic and flexible—thus flowing within the riverbed—, but ultimately founded on the certainties that the riverbed itself provides.

It is certainly accurate to read these certainties as indeed necessary to draw some boundaries around (mere) knowledge. However, even these boundaries are subject to gradual change: Wittgenstein himself formulates that “what men consider reasonable or unreasonable alters. At certain periods men find reasonable what at other periods they found unreasonable. And vice versa.” (OC 1969: 336) *That* such change occurs is acknowledged by both the conservative Right Wittgensteinian as well as their more liberal Left counterpart. However, whilst the conservative interpreter isn't especially concerned with the role of change, the Left Wittgensteinian emphasizes that constant change is part of the human form of life. Practices are anything but frictionless and fully developed, just like rivers naturally alter their shapes over time.

However, what is often neglected is that change itself may come in different forms. In this paper, I want to argue that this sort of gradual change that both Right and Left readers may accept as being present in Wittgenstein's thinking is not the only kind of change that we may draw inspiration on from Wittgenstein. By contrast, I suggest that there is indeed room for more radical change—change that doesn't just happen accidentally, but is rather brought about by active intervention on the part of those practitioners who have the bigger picture in view. To make this thought plausible, I will draw on Kristie Dotson's contributions on epistemic oppression, particularly on her concept of third-order change, which provides exactly the kind of reflection on the mechanisms of unjust practices that ideally leads to the change I have in mind. I argue that the implementation of third-order conceptual change by means of

such critical awareness is itself a practice that constitutes the riverbed, and that, therefore, a Left Wittgensteinian reading of OC is consistent with Wittgenstein's thinking, and indeed especially desirable where marginalized perspectives are concerned.

First, I will roughly sketch a Left Wittgensteinian critique of its conservative counterpart. Second, I will develop a feminist argument emphasizing the need of a critical engagement with existing practices in order to promote epistemic justice. Finally, by building on the previous parts, I will sketch a Left Wittgensteinian and feminist-inspired way of doing proper justice to the potentials of third-order change in OC.

2 Some Perspectives of Left Wittgensteinianism

Historically, Left Wittgensteinianism has emerged as a family of conceptions challenging Right interpretations of Wittgenstein's remarks on practice and tradition. Those conceptions, broadly speaking, understand life as generating a holistic set of practices and traditions—a form of life—that effectively suppresses any external criticism directed towards it.

Right Wittgensteinianism can be summed up as essentially following four principles:

- (a) to view our form of life as a finely calibrated functional whole;
 - (b) to assume that it is tensionless;
 - (c) to think of it in static and ahistorical terms; and
 - (d) to view departures from current practice not just as unmotivated, but as rationally unintelligible unless and until they have already happened.
- (Queloz and Cueni 2021: 764)

In this framework, the existing form of life must be adhered to; change is, if at all, only possible from *within* those already existing practices—there is no method of conscious transcendence (Crary 2000: 122). Any such *substantial* criticism would shake up the functional whole to an extent that would threaten the entire network. Therefore, if criticism can be addressed at all, its extent can only be small-scaled and gradual, for anything else would be plainly unintelligible within the holistic system (Williams 2005, 34). This way, Right Wittgensteinianism is essentially foundationalist in the sense that practice is tied to this already existing whole from which there is no escape.

Now, Left Wittgensteinianism rejects all four tendencies. Neither do practices and traditions produce (a) a functional whole that is (b) tensionless, (c) static and ahistorical, nor need criticism and changes in practice be (d) unmotivated. Amidst a certain blindness to political and ethical matters which precludes actors from seeing the significance of criticism and change in and of certain forms of life, a wider understanding of the community as including everyone who seeks to take part in the practice from their very own local and historical situation (rather than as a term of exclusion of those who don't fit in) puts emphasis on the *possibility* of constant development (Williams 2005: 37; Mollema 2024: 604), and indeed a *need* for adjustment and change because of the tensions and developments inherent in our practices (Queloz and Cueni 2021: 770f.). The relevant question, then, is to ask for the *point* or *goal* of each criticism so that conceptual change can address *specific* tensions and needs within the practicing community (Queloz and Cueni 2021: 767). This way, the idea of a form of life as an untouchable whole is replaced by an understanding of practices as vivid and constantly changing. In other words: “‘meaning’ and ‘criticism’ [are brought] back from their metaphysical to their everyday use” (Deegan 2023: 5). Left Wittgensteinianism, by emphasizing the role of situated individuals taking part in each practice, reveals the importance of an ongoing critical practice with regard to the need for conceptual change.

3 A Feminist Perspective

Of course, Wittgenstein isn't the only one to discuss social practice and change. In what follows, I will sketch a feminist-inspired conception of conceptual change that ultimately helps understand how critical engagement of individuals figures in the change of existing practices. Since I claim that the focus on small-scaled change does not exhaust Left Wittgensteinian potentials, I will explain how the space for higher-order change that is already implicit in Wittgenstein results in a conception of practice that is even richer than what is traditionally covered by Left Wittgensteinianism.

In *Conceptualizing Epistemic Oppression*, Kristie Dotson distinguishes three levels of change concerning collective epistemic resources. The overall aim of these changes is to eradicate existing injustices inherent in practices of knowledge production. (Note that the third level will be the most important one for our purposes.)

First-order change is concerned with “the tacit reinforcement of present understandings”, which “leaves intact already present operative schemata” (Dotson 2014: 118; cf. Bartunek and Moch 1987: 486). Such change occurs when small adjustments are made in order to better adhere to present beliefs and values of a given society, whereas the beliefs and values themselves remain unquestioned—it is aimed at the *efficiency* of shared epistemic resources (Dotson 2014: 123). First-order change may be required, for example, where instances of what Miranda Fricker calls *Testimonial Injustice* are present, which Fricker characterizes as being in place when a speaker “is *wronged specifically in her capacity as a knower*” (Fricker 2007: 20; italics in the original). The harmful result is that

[t]estimonial injustice [prejudicially] excludes the subject from trustful conversation. Thus it marginalizes the subject in her participation in the very activity that steadies the mind and forges an essential aspect of identity—two processes of fundamental psychological importance for the individual. (Fricker 2007: 53f.)

Such testimonial injustice may be operative, for example, when a hearer did not accurately evaluate the credibility of a (marginalized) speaker under a given sensible schema in which such speakers should usually be granted special credibility. A first-order change of the kind that Dotson has in mind, then, leads the hearer to adjusting their credibility judgment towards the speaker, leading to better achieving the common goal of accurately valuing such reports of marginalized speakers. This, then, is the exercise of testimonial *justice* in the sense Fricker intends.

Second-order change is characterized as “the conscious modification of present schemata in a particular direction” (Dotson 2014: 118; cf. Bartunek and Moch 1987: 486). It is in place when a certain schema is deemed *insufficient* to contribute to the collective goal (Dotson 2014: 126), and is therefore replaced by a better schema. Again, the change occurring here can be thought of as the kind of change that Fricker has in mind in her conception of *Hermeneutical Justice*. It again addresses a kind of injustice, this time hermeneutical injustice, which Fricker characterizes as

having some significant area of one's social experience obscured from collective understanding owing to a structural identity prejudice in the collective hermeneutical resource. (Fricker 2007: 155; italics in the original)

Hermeneutical injustice causes harm specifically when an existing lack of concepts for the specific type of marginalization, discrimination or mistreatment that someone experiences prevents them from articulating their experiences in such a way that they have no means to address, or let alone stop it (Fricker 2007: 151). However, when certain hermeneutical resources are developed that conceptualize certain experiences of marginalization, that experience is removed from its unintelligibility—the experience can be put in words, and both the speaker and the hearer can better understand what is happening. In this process, the present schema is altered in such a way that the shared goals and values can be achieved more sufficiently—a second-order change is in place.

A third-order change, finally, operates on yet a different level, which Dotson claims is not reducible to the other two (Dotson 2014: 132–133) and is thereby much more substantial for our present project. It involves “the training of organizational members to be aware of their present schemata and thereby more able to change these schemata as they see fit” (Dotson 2014: 118; see also Bartunek and Moch 1987: 486). Such awareness involves a sensitivity to the “overall epistemological systems that are preserving and legitimating inadequate epistemic resources.” (Dotson 2014: 131) Third-order change thus educates members of the community about the wider contexts surrounding a given schema, and to see possible pathways for changing these schemata accordingly, i.e. when an epistemic resource is found to be *inadequate* to properly address some features of a present collective value (Dotson 2014: 129). By means of reference to the overall system, it is possible to find more adequate alternatives that may replace a present schema. However, that process demands the willingness of the members of a given community to contribute to that change. Since existing schemata have an inherent resilience towards change, it takes quite some effort to overcome that threshold enabling third-order change (Dotson 2014: 132). It takes continuous efforts on the part of both speakers and hearers to gradually convince proponents of a present schema that a different one better suits the requirements of a given epistemic discourse. That effort must involve all three levels of change (Walsh 2004: 307),

since the new schema won't be fully intelligible to hearers loyal to the present schema, whereas at the same time the change that is ultimately sought won't be fully articulable in terms of the present schema (which is exactly why the change is needed). However, once the effort has been made and the new schema is established, a major shortcoming of the previous set of epistemic resources of the community will have been removed—careful reflection and search for alternatives outside the current way of thinking made it possible to develop a new schema that is capable of providing the resources necessary to articulate the special experiences of its members.

4 Third-Order Change in Wittgensteinian Practices

With these conceptual tools in place, we can once again see that the criticism of the four principles of Right Wittgensteinianism that I discussed above is indeed the more desirable stance amidst real-world issues: In contrast to what a foundationalist reading of Wittgenstein suggests, one single glance at persistent social injustices reveals that our form of life is anything but tensionless. Once we acknowledge that members of the community introduce many different perspectives, situations, and circumstances, it becomes obvious that these perspectives will inevitably lead to conflicting interests and needs. In this sense, the functional whole can only *seem* finely calibrated from a privileged standpoint, and the inherent historicity of our form of life can be glossed over by simply ignoring the diversity of our own traditions. It takes only little awareness to see that the correction of existing injustices is anything but unmotivated even within our own tradition. Indeed, it is a symptom of ignorance to pretend that there is no room for such change.

The traditional rationale may be right about that certain difficulty in articulating the kind of change that is needed—for the simple reason that our historical ignorance has generated a lack of proper concepts allowing us to understand such proposals for change. This, however, is no excuse for ignoring the adequacy of such complaints: We are as much responsible as we are able to reflect on the best performance of our practices.

In this sense, *the changing of practices is itself a practice* (Mollema 2024: 604). It is one that goes beyond daily business, and instead evaluates how well we are currently performing our practices, thereby overcoming the *blindness of practices to political and ethical matters*. When such a self-reflexive practice

detects some persistent injustice, we must ask: What is the *point* of that criticism raised and how do we need to change a given practice in order to address it?

Let me note here that Right Wittgensteinianism, as characterized above, only admits the first two Dotsonian forms of change. In the conservative framework, any criticism beyond the second order would count as coming from outside the existing holistic form of life, thus being dismissed as unintelligible before being heard. However, this very limitation of the Right reading shows the relevance of Dotson's third-order change for Left Wittgensteinianism: By establishing a sensitivity to the surrounding context of a given schema, it is possible to look beyond existing practices and to see room for change in the relevant sense. Of course, such sensitivity is nothing that people merely *possess*—rather, it is a capacity to perform constant analysis, reflection, and criticism of existing practices, which requires that the participants in this practice constantly force themselves to question those goings-on that seem so attractive. In one sense—the one relevant for our criticism of Right Wittgensteinianism—, this is, by means of introducing change even though it seems yet unintelligible, a way of reaching *beyond* an existing practice in order to change it for the better. In another sense—in accordance with a Left account—, this is by no means an overstepping of the form of life. Instead, speaking in a Thompsonian manner, it is very much part of the human form of life that it already *entails* its own criticism (even that on the third level).

Taking another look at the riverbed analogy can, I think, provide some helpful illustration. Roughly speaking, first and second order changes are located *within* the riverbed—they are concerned with the everyday knowledge, the water that flows down the hill. They debate how existing practices can be made more efficient, whom to listen to under what circumstances, and where to provide conceptual tools tailored to marginalized perspectives. All of this requires relatively small-scaled change—erosion and sedimentation of particles—that leads to only slight impingement on the shape of the river itself. Third order change, by contrast, provokes a reflexive awareness towards that shape. If that critical practice diagnoses that a certain practice poses a major threat to the functioning of the whole system, then it has to work towards more invasive change. In that case, it is not enough to let natural processes (i.e.

ongoing unreflected everyday practicing) determine the direction of flow; instead, the river must be redirected by means of active and engaged digging. Such reflection creates awareness of what is required to replace unjust practices with new schemata that provide room for all its participants.

I claim that such an interpretation of the riverbed analogy is consistent with Wittgenstein's understanding of practices. In contrast to what conservative readers may claim, critical reflection, which is itself a practice on its own, makes it possible to look beyond the illusion of the seemingly tensionless functional whole (indeed, only the privileged perspective is under that illusion). This way, there is no need for construing the human form of life as a fully enclosed system of practices, but rather as a space of inherent flexibility. Seriously using the potentials of third-order change in this sense is the first step towards creating a more inclusive environment for even the most marginalized perspectives.

References

- Bartunek, Jean M.; Moch, Michael K. (1987) "First-order, second-order, and third-order change and organization development interventions: A cognitive approach", *The Journal of Applied Behavioral Science* 23(4), 483–500.
- Crary, Alice (2000) "Wittgenstein's Philosophy in Relation to Political Thought", in: Alice Crary and Rupert Read (eds.): *The New Wittgenstein*. London/New York: Routledge, 118–145.
- Deegan, Marc James (2023) "A Marxist reading of Ludwig Wittgenstein: Making the case for social and political change", *Philosophy and Social Criticism* 51(2), 187–207. <https://doi.org/10.1177/01914537231170907>
- Dotson, Kristie (2014) "Conceptualizing Epistemic Oppression", *Social Epistemology* 28(2), 115–138. <https://doi.org/10.1080/02691728.2013.782585>
- Fricker, Miranda (2007) *Epistemic Injustice. Power & the Ethics of Knowing*. Oxford/New York: Oxford University Press.
- Mollema, Thomas (2024): "On certainty, Left Wittgensteinianism and conceptual change", *Theoria*, 90(6), 603–623. DOI: <https://doi.org/10.1111/theo.12558>
- Nyíri, Kristóf [J. C.] (1976) "Wittgenstein's new traditionalism", *Acta Philosophica Fennica* 28(1–3), 501–512.
- Queloz, Matthieu; Cueni, Damian (2021) "Left Wittgensteinianism", *European Journal of Philosophy* 29(4): 758–777. <https://doi.org/10.1111/ejop.12603>

Walsh, Kate (2004) "Interpreting the Impact of Culture on Structure: The Role of Change Processes." *The Journal of Applied Behavioral Science*, 40(3), 302–322. <https://doi.org/10.1177/0021886304266845>

Williams, Bernard (2005) *In the Beginning Was the Deed: Realism and Moralism in Political Argument*. Geoffrey Hawthorn (ed.). Princeton: Princeton University Press. <https://doi.org/10.1515/9781400826735.29>

Wittgenstein, Ludwig (1969) *On Certainty / Über Gewißheit*. G. E. M. Anscombe and G. H. von Wright (eds.), translated by Denis Paul and G. E. M. Anscombe. Oxford: Basil Blackwell.

Beyond a question that has no name: Perspectival sense-making and hermeneutical Injustice

Yinmei Wu (Edinburgh, United Kingdom)

Abstract

How the distortion of social power influences women and other subordinate groups in making sense of their experiences is a central concern of feminist epistemology. The notion of hermeneutical injustice is closely tied to this concern. As coined by Miranda Fricker (2007: 1), hermeneutical injustice is a form of epistemic injustice that occurs when “a gap in collective interpretive resources puts someone at an unfair disadvantage in making sense of their social experiences.” Much literature on hermeneutical injustice focuses on how the unavailability of apt terms and concepts—or the overabundance of distorted ones—hinders subordinate groups from making sense of their experiences. In this paper, I argue that this concept-centred approach is too restrictive, as sense-making is more than a matter of concept application. Following Paulina Sliwa (2024), I argue that making sense of a situation requires a perspective: a set of interlocking attentive, cognitive, affective, and inquisitive dispositions. While conceptualisation could be a crucial aspect of sense-making, it is not the whole picture. To illustrate this, I show two central cases: individuals may be able to make sense of their experience to some extent without possessing the apt concept, and they may still struggle even after applying an apt one. I further argue that a perspectival model of sense-making better serves our epistemic justice agenda by a) more accurately capturing marginalised agents’ hermeneutical capacities and the unfair disadvantages they encountered and b) pointing to more possible avenues of epistemic resistance.

1. Introduction: What is hermeneutical injustice?

How the distortion in social power influences women and other subordinate groups to understand their experiences is one of the central concerns of feminist epistemology. The notion of hermeneutical injustice, coined by Miranda Fricker (2007), is closely relevant to this concern. In Fricker's account, hermeneutical injustice occurs when individuals face unfair difficulties making sense of their social experience due to hermeneutical marginalisation—unequal participation in the practices through which social meanings are generated.

Fricker (2007) offers two central cases to illustrate hermeneutical injustice, both of which are quoted from Susan Brownmiller's (1990) memoir of the second-wave feminist movement.

A. Sexual harassment:

Carmita Wood suffered from sexual harassment before the existence of this concept. Lacking this concept renders Wood unable to make sense of her experience and communicate it to others.

B. Postpartum depression

Lacking the concept of "postpartum depression" hindered Wendy Sanford from seeing the distress she felt after giving birth as a combination of biological factors and social isolation, instead of her personal failure.

Both of Fricker's central cases involve "a lacuna where the name of a distinctive social experience should be" (Fricker 2007: 151) in collective hermeneutical resources. As Medina (2017) puts it, in both cases, understanding fails because words and concepts are lacking. In her latter writings, Fricker also refers to hermeneutical injustice as a "poverty of concepts", "a paucity of shared concepts", and "lack of conceptual-interpretative resources" (2013: 51; 2016: 12).

If the problem of hermeneutical injustice is that the needed concepts and names are lacking, then naturally, the corresponding solution is to create and circulate the lacking names and concepts. Following Fricker's approach, much of the current discussion on hermeneutical injustice shares this linguistic and conceptual-centred pattern (see the critiques from Medina 2017; Crerar 2016; Catala 2020). The underlying assumption of this approach is that sense-making is essentially a conceptualising process. While I agree that the application of fitting concepts can largely facilitate epistemic agents making apt sense of their social experience, I want to highlight that this is not what sense-making is all about, and the concept-focused approach of hermeneutical injustice remains too narrow for our epistemic justice agenda. For the case at hand, I understand the requirements of epistemic justice as at least two-fold: first, it requires us to recognise and respect socially subordinated knowers' capabilities; second, it requires us to capture the difficulties faced by knowers due to unfair social arrangements.

In the following sections, I will argue that making sense of an experience is more than applying a set of apt concepts to it; it requires building an apt perspective on it instead. After introducing what a perspective is in section 2, I will demonstrate two cases where sense-making and conceptualising do not go

hand in hand in sections 3 and 4. Following that, in section 5, I will analyse how the perspectival model of sense-making aligns well with the epistemic injustice agenda.

2. What is a perspective?

Let us start by exploring what a perspective is. My notion of perspective is built on Elisabeth Camp, Rachel Fraser, and Paulina Sliwa's works. According to Camp (2017), A perspective is an open-ended disposition to characterise: to encounter, interpret, and respond to some parts of the world in certain ways. Generally, a perspective is a certain way of looking at the world. Specifically, a way of looking may include the following dispositions (Fraser 2021; Sliwa 2024):

Salience. Dispositions to find certain features within that subject matter salient.

Concepts. Disposition to employ certain sets of concepts to think about the subject matter, to taxonomise features of the situation in particular ways.

Associative thought. Dispositions to draw particular inferences, to engage one's imagination, to see similarities and differences

Affect. Dispositions to respond affectively to certain features.

Inquisitive. Dispositions to find certain kinds of inquiry worthwhile.

Action. Dispositions to take actions in certain ways.

The perspectival model highlights the plural structure of one's sense-making process. In this picture, making sense of an experience is not only about how you name or taxonomise it, but is also about how you feel about it, what features strike you as salient in a given situation, how much significance you intend to attribute to these features, what are the following questions you inclined to ask, what are the actions you tend to take etc. These dispositions are interlocked with each other in the sense that they motivate, reinforce and justify each other. For example, the anger you felt toward a rape case may interact with your inquisitive dispositions: motivate your inquiry on this social phenomenon, lead you to see the gender-based pattern of it, make the systematic factors of the given case more salient to you than the individual

factors in a given case. Your inquiry may further interact with your affective dispositions, reinforce and justify your anger: now you know that there is a solid, systematic injustice that deserves your anger. Further, the above cognitive and affective dispositions may motivate you to take certain actions to alleviate gender injustice.

Above is one short story of how your affective, inquisitive, attentive, and action-related dispositions may interlock and mutually reinforce each other. When one built a bundle of dispositions interlocked like that, she got a unified perspective on it in the Campian sense. A perspective on an experience is relevant to but not reducible to the way we conceptualise that experience. To make this point clear, let us consider two cases where sense-making and concept application do not go hand in hand.

3. Carmita Wood's case revisited

Taking concept application as one of the multiple dimensions of sense-making opens the theoretical possibility that an individual may be able to make sense of their experience to some level before the availability of apt concepts in collective hermeneutical resources. This possibility is closely relevant to the controversies regarding Fricker's influential analysis of Carmita Wood's case.

Let us start by taking a closer look at Fricker's (2007) analysis of Carmita Wood's case, which is based on Susan Brownmiller's (1990) account. Brownmiller narrates Wood's case as follows:

"Carmita Wood, age forty-four, born and raised in the apple orchard region of Lake Cayuga, and the sole support of two of her children, had worked for eight years in Cornell's department of nuclear physics, advancing from lab assistant to a desk job handling administrative chores. Wood did not know why she had been singled out, or indeed if she had been singled out, but a distinguished professor seemed unable to keep his hands off her. As Wood told the story, the eminent man would jiggle his crotch when he stood near her desk and looked at his mail, or he'd deliberately brush against her breasts while reaching for some papers...She requested a transfer to another department, and when it didn't come through, she quit. She walked out the door and went to Florida for some rest and recuperation. Upon her return, she applied for unemployment insurance. When the claims investigator asked why she

had left her job after eight years, Wood was at a loss to describe the hateful episodes. She was ashamed and embarrassed. Under prodding—the blank on the form needed to be filled in—she answered that her reasons had been personal. Her claim for unemployment benefits was denied."

After Wood's claim for unemployment benefits was denied, Wood sought out Lin Farley, who ran the "women and work" seminar at Cornell at that time, for help. After Wood shared her experience, the women in the seminar realised that they all had gone through an experience like that, and they decided to hold a speak-out event about this phenomenon and came up with a name for it: sexual harassment. This is also one version of the origin stories of this term.

Fricker interprets Wood's case as "a story about how extant collective hermeneutical resources can have a lacuna where the name of a distinctive social experience should be" (2007: 150). She argues that both the harasser and Carmita Wood are "cognitively handicapped by the hermeneutical lacuna" (ibid), while only Wood's "cognitive disablement" is seriously disadvantage to her: "her hermeneutical disadvantage renders her unable to make sense of her ongoing mistreatment, and this in turn prevents her from protesting it, let alone securing effective measures to stop it." (ibid).

Fricker's analysis of Carmita Wood's case is controversial. Some philosophers argue that Fricker underestimates Wood's hermeneutical capabilities of her experience prior to holding a novel concept of it. For example, Mason (2011) argues that despite being advantaged by the conceptual lacuna, Wood's actions of seeking help from Lin Farley and sharing her experience in the consciousness-raising group demonstrate that she understands her experience to a certain level prior to having a name for it. She argues that while Wood may not fully grasp the social significance of her experience, it is still inaccurate to judge Wood as unable to make sense of her experience.

The evidence from some other narrators of Carmita Wood's case echoes Mason's judgement that Wood's hermeneutical capabilities may be underestimated in Fricker's account. For example, Baker (2008) offers a detailed account of how Wood asked for a hearing prior to attending the consciousness-raising seminar. After Wood's initial application for

unemployment compensation was denied. At the hearing, she offered a detailed report of how the harasser in her case, McDaniel, treated women as "second-class citizens, and inferior beings" (Baker 2008: 28) and constantly inappropriately touched female employees in the office. Besides, Wood brought two witnesses who had gone through similar experiences to the hearing to confirm her account. However, Wood's claim was again denied by the hearing officer.

Baker's account offers us more ground to rethink Carmita Wood's hermeneutical capabilities before the existence of the novel concept of sexual harassment, and the perspectival model of sense-making is well-suited to capture that evidence. Before holding a novel concept of sexual harassment, McDaniel's attitude and behavioural pattern towards women struck Wood as salient and problematic, and she linked this to McDaniel's sexist worldview of viewing women as inferior beings. Wood noticed the similarities between her experience and the other women's experiences in the lab. After her claims were denied again by the hearing office, she was inclined to take further inquiry and actions.

According to the perspectival model of sense-making, the mentioned information is crucial evidence that Wood's attentive, evaluative, and inquisitive dispositions are interlocking together and influencing each other, which means she was able to make sense of her experience to some extent. The plural structure of the perspectival model of sense-making allows the possibility that individuals can be disadvantaged along some dimension in their hermeneutical inquiry while supported by others. Wood may be disadvantaged by the availability of the concept of sexual harassment, but the discomfort and anger she felt, her salience perspective, and her connection with the other women in the lab keep her going on her sense-making inquiry.

In contrast, if we tie an individual's hermeneutical capabilities too closely to concept application, we risk ignoring the above information as relevant to Wood's sense-making abilities and may lead to a rushed conclusion that Wood was unable to make sense of her experience due to a conceptual lacuna. This single-axis methodology of evaluating an individual's hermeneutical capabilities is dangerous for our own epistemic justice agenda of recognising and respecting knowers' capabilities.

As discussed above, I argue that individuals may be able to make sense of their experience before applying apt concepts to it. The perspectival model of sense-making offers us a more delicate approach to capturing an individual's sense-making abilities along plural dimensions.

4. Hermeneutical struggles after conceptualising

Let us then consider another way that sense-making does not go hand in hand with concept application: individuals may struggle with making sense of their experience after successful concept application.

An important case of this kind is Susan Brison's case. Susan Brison is a philosophy professor at Dartmouth University. On July 4, 1990, at 10:30 in the morning, A man assaulted Brison when she went for a walk along a peaceful-looking country road in a village outside Grenoble, France. As Brison (2022: 237) writes: "He jumped me from behind, threw me into the underbrush, beat me, raped me, strangled me into unconsciousness several times, dragged me into a ravine, hit me on the head with a rock, and left me for dead."

While the attacker of Brison's case was convicted of rape and attempted murder, and she had no difficulties applying the concept "rape" to her experience, Brison felt "things had stopped making sense" (ibid) to her.

For Brison, one major difficulty in making sense of her experience lies in situating this incident within a coherent life narrative. In Brison's words (2002: 103): "I thought I had made a certain sense of things until the moment I was assaulted. At any rate, I thought I knew how to carry on with my life—to project myself, through action, into an imagined future...Trauma shatters this assumption by introducing an event that fits no discernible pattern. Not only is it now impossible to carry on with the series, but whatever sense had been made of it in the past has been destroyed." On this point, Brison also invokes Charlotte Pierce-Baker's statement: to convey the trauma she experienced, Pierce-Baker (1998: 29) states that when she heard a crash or a thud one evening, shortly before being raped by two assailants who had broken into her home: "At that moment I knew the meaning of the word 'nonsense'". For both Brison and Pierce-Baker, the feeling of 'nonsense' does not lie in being unable to conceptualize or name their experience as rape, but lies in situating this experience into their general life narration: a traumatic experience was inserted into their life as a "surd" (Brison 2002: 103), which she cannot see any

similarities between this incident and things that happened to her in the past and the things she once projected for her future.

The perspectival account of sense-making opens the theoretical space for accommodating the unfair hermeneutical difficulties encountered by Brison. Since the individual sense-making process is a more complicated process than concept application, we may succeed at applying a concept to it yet struggle with searching for a coherent perspective on it. Brison's feeling of her life narration was interrupted reminds us that we do not simply make sense of our experiences as isolated cases, by making sense of a significant experience, we sometimes need to search for a way of situating them into our life narration, which requires the dispositions to notice the similarities and the differences between it and the other life events, to evaluate it in a way that aligns with our value system, to figure out how to feel about it and have a perspective on it that fits our general perspective on our lives.

Admittedly, not all of our sense-making inquiries are that demanding. The application of simple names may set us up well in many life scenarios. However, I argue it is important to leave theoretical space for the more complex scenarios, especially for hermeneutical justice purposes, given that many central cases of hermeneutical injustice involve a deep, unjust mistreatment received by the members of the socially subordinated groups. This type of experience is typically heavy, traumatic, and hard to make sense of along any single axis.

5. Lessons and Implications

So far, I have analysed two cases of hermeneutical injustice in detail with the lens of the perspectival account of sense-making, to demonstrate that making sense of an experience is irreducible in conceptualising it. On the one hand, one may be able to make sense of an experience to some extent before holding an apt concept of it. On the other hand, one may still struggle to make an apt sense of her experience despite having no difficulties conceptualising it. Following this line of reasoning, if we view sense-making merely as a conceptualising process, then, on the one hand, we may underestimate some individual's hermeneutical capabilities due to the unavailability of certain

concepts and articulations. On the other hand, we risk overlooking the unfair hermeneutical difficulties encountered by epistemic agents other than concept-application difficulties.

With the understanding that hermeneutical justice requires both recognising an individual's hermeneutical capability and the unfair difficulties individuals encounter in terms of making sense of their social experiences, I argue that the perspectival account of sense-making is particularly well-suited for our hermeneutical justice agenda. The holistic, multidimensional characteristics of perspective make it well-suited to capture individuals' hermeneutical capabilities and difficulties in terms of sense-making along different dispositions. In this light, it aligns well with our two-fold epistemic justice tasks to both respect marginalised knowers' capabilities and acknowledge the unfair disadvantages they encountered as knowers.

Furthermore, the perspectival account also points out more possible avenues of epistemic resistance. When sense-making is viewed essentially as a concept-application process, most discussion of hermeneutical resources is focused on concepts (Crerar 2016). However, if we view sense-making as a perspective-building process, which involves a whole set of interlocked dispositions, then all the resources that may effectively interact with sense-making dispositions should be considered as hermeneutical resources. Camp's works on frames (2013; 2017; 2023) can be useful for us to expand the scope of hermeneutical resources. A frame is a crystallisation of a perspective which can take a variety of forms, including but not limited to concepts, critical theories, narratives, slurs, slogans, memes, and various forms of visual tropes. I argue that an expansive perspectival understanding of sense-making and hermeneutical resources can add more flexibility and possibilities to our epistemic resistance.

Conclusion:

In this paper, I have argued that hermeneutical injustice is not merely a matter of missing apt concepts or difficulties in concept-application, but of unfairly disrupted sense-making across multiple dimensions. The perspectival model of sense-making better captures both the capacities and struggles of marginalised knowers and broadens our scope of hermeneutical resources and epistemic resistance.

References

- Baker, C. N. (2008) *The women's movement against sexual harassment*. Cambridge University Press.
- Brison, S. J. (2002) *Aftermath: Violence and the remaking of a self*. Princeton University Press.
- Brison, S. J. (2022). Beyond consent. In *The Routledge Handbook of Philosophy of Sex and Sexuality* (pp. 235-246). Routledge.
- Catala, A. (2020) "Metaepistemic injustice and intellectual disability: A pluralist account of epistemic agency", *Ethical Theory and Moral Practice* 23(5), 755–776.
- Camp, E. (2013) "Slurring perspectives", *Analytic Philosophy* 54(3).
- Camp, E. (2015) "Logical concepts and associative characterizations", in Margolis, E. & Laurence, S. (eds.) *The Conceptual Mind: New Directions in the Study of Concepts*, 591–621. MIT Press.
- Camp, E. (2017) "Perspectives in imaginative engagement with fiction", *Philosophical Perspectives* 31(1), 73–102.
- Crerar, C. (2016) "Taboo, hermeneutical injustice, and expressively free environments", *Episteme* 13(2), 195–207.
- Flores, C.; Camp, E. (2023) "That's all you really are": Centering identities without essentialist beliefs, in Haslanger, S., Schroeter, L., & Schroeter, F. (eds.) *Mind, Language, and Social Hierarchy*. Oxford University Press.
- Fraser, R. (2021) "Narrative testimony", *Philosophical Studies* 178, 4025–4052.
- Fricker, M. (2007) *Epistemic injustice: Power and the ethics of knowing*. Oxford University Press.
- Fricker, M. (2013) "Epistemic justice as a condition of political freedom?", *Synthese* 190, 1317–1332.
- Fricker, M. (2016) "Epistemic injustice and the preservation of ignorance", *The Epistemic Dimensions of Ignorance*, Vol. 1, 144–159. Cambridge: Cambridge University Press.
- Jenkins, K. (2017) "Rape myths and domestic abuse myths as hermeneutical injustices", *Journal of Applied Philosophy* 34(2), 191–205.
- Mason, R. (2011) "Two kinds of unknowing", *Hypatia* 26(2), 294–307.
- Sliwa, P. (2024) "Making sense of things: Moral inquiry as hermeneutical inquiry", *Philosophy and Phenomenological Research* 109(1), 117–137.

Acknowledgement

I am deeply grateful to my supervisors, Aidan McGlynn and Filipa Lopes, for their crucial guidance and support. This paper has also benefited from valuable conversations with Elisabeth Camp, Ethan Nowak, and Irene Sanchez Lopez. I am also thankful for the generous encouragement from my former supervisor, Ella Whiteley.

Wittgenstein und J. H. Newman über die Rolle von „Bildern“ im religiösen Glauben.

Isacco Zampini (Jena, Deutschland/Venedig, Italien)

Abstract

In letzter Zeit erscheinen in der Literatur über Wittgenstein immer mehr Studien, die den Einfluss beleuchten, den John Henry Newman auf die letzte Phase des Denkens des Wiener Philosophen gehabt hat. Im vorliegenden Aufsatz möchte ich die bisher wenig erforschte Frage beantworten: Was hat Wittgenstein an Newmans Sicht des religiösen Denkens überzeugend gefunden? Bei der Beantwortung dieser Frage konzentriere ich mich auf eine der wichtigsten Gemeinsamkeiten der beiden Denker in diesem Bereich: das Gewicht von „Bildern“ im religiösen Glauben.

Zunächst zeige ich, welche Rolle „Bilder“ in Wittgensteins spätem Denken in Bezug auf das religiöse Thema spielen und welche Bedeutung sie in Newmans *Grammar of Assent* haben. Anschließend zeige ich, dass, trotz der unbestreitbaren Bedeutungsunterschiede in der Verwendung des Bildbegriffs, Beide eine allgemeine philosophische Sensibilität teilen. Sie halten eine Untersuchung des Gewichts, das Bilder im religiösen Glauben und im Leben der Gläubigen haben, für relevanter als einen metaphysischen Diskurs über Religion. Schließlich weise ich nach, dass diese Gemeinsamkeit nicht nur in der allgemeinen Herangehensweise an religiöse Fragen liegt, sondern auch in einigen grundlegenden Ergebnissen der Analysen der Beiden: zum Beispiel, dass religiöse Bilder als Prinzipien der Belebung, Regulierung und Motivation des moralischen Lebens der Gläubigen dienen, oder dass es eine enge Beziehung zwischen dem Gebrauch mit Recht bestimmter religiöser Bilder und einer bestimmten Lebensweise der Gläubigen besthet.

0. Einleitung

In letzter Zeit erscheinen in der Literatur über Wittgenstein immer mehr Studien, die den Einfluss beleuchten, den John Henry Newman auf die letzte Phase des Denkens des Wiener Philosophen gehabt hat. Einige dieser Studien zielen darauf ab, besser zu verstehen, worauf Wittgenstein in § 1 von *Über Gewissheit* anspielt dem Absatz, in dem der Name „H. Newman“ auftaucht (z. B. Kienzler 2006). Andere haben versucht, dieses Spätwerk so zu lesen, als ob es ein Werk wäre, das von Wittgenstein in erster Linie im Dialog mit Newmans *An Essay in Aid of a Grammar of Assent* (GA) konzipiert wurde (Pritchard 2015), und noch andere haben einige allgemeine Gemeinsamkeiten in der Art und Weise aufgezeigt, wie die beiden Denker Philosophie betrieben (Kfoury 2023, aber auch z. B. Pichler 2024, Pichler und Sunday Grève 2025, Pritchard forthcoming).

Im vorliegenden Aufsatz möchte ich jedoch die bisher wenig erforschte Frage beantworten: Was hat Wittgenstein an Newmans Sicht des religiösen Denkens überzeugend gefunden?

Bei der Beantwortung dieser Frage konzentriere ich mich auf eine der wichtigsten Gemeinsamkeiten der beiden Denker in diesem Bereich: das Gewicht von „Bildern“ im religiösen Glauben. Dabei zeige ich zunächst 1. die Bedeutung, die dieses Thema in Wittgensteins Betrachtungen zur Religion ab 1933 spielt, dann 2. die zentrale Rolle, die es in GA spielt, und schließlich 3. die Anknüpfungspunkte bei Newman für Wittgenstein.

1. Die Rolle von „Bildern“ im Denken über Religion des späten Wittgenstein

Bekanntlich hat Wittgenstein relativ wenige Bemerkungen hinterlassen, die direkt mit dem Thema Religion zu tun haben, und viele dieser Beobachtungen finden sich in einigen im Wortlaut nicht sehr zuverlässigen Vorlesungen, in verschiedenen Notizen in Tagebüchern, die nicht offenkundig philosophischer Natur sind, oder in Gesprächen mit SchülerInnen, die nach seinem Tod veröffentlicht wurden. Nichtsdestotrotz haben diese wenigen Spuren verschiedene Interpretationsversuche und zahlreiche Studien inspiriert und regen weiterhin die Überlegungen von Philosophen und Theologen zu diesem Thema an.

Ein Motiv, das in diesen verschiedenen Quellen wiederkehrt und das in der Literatur nicht unerwähnt geblieben ist (zwei berühmte Beispiele: Putnam 1992: 134-157; oder auch Diamond 2005), ist das Gewicht, das der Begriff „Bild“ in diesen Untersuchungen erhalten. Und tatsächlich taucht das Wort „Bild“ im Zusammenhang mit dem religiösen Diskurs bereits in einigen der Tagebuchaufzeichnungen Wittgensteins von 1936 auf (DB: 75, 80). In einer dieser Notizen lesen wir zum Beispiel:

“Ich könnte also, scheint es, alle Ausdrücke brauchen, die die Religion hier tatsächlich gebraucht. Es drängen sich mir Bilder auf. [...]. Vor allem sind es natürlich nicht Gleichnisse. Denn was sich durch ein Gleichnis sagen läßt, das auch ohne Gleichnis. Diese Bilder und Ausdrücke haben ihr Leben vielmehr nur in einer hohen Sphäre des Lebens nur in dieser Sphäre können sie mit Recht gebraucht werden.“ (DB: 80)

Wir finden diese Verbindung zwischen Bild und Religion in mehreren späteren Bemerkungen entwickelt. Wenn wir uns hier auf einige wenige Zitate aus *Lectures and Conversations* (LA 1966) beschränken, finden wir zum Beispiel geschrieben:

“[B]ut he has what might call an unshakeable belief. It will show, not by reasoning or by appeal to ordinary grounds for belief, but rather by regulating for all in his life. This is a very much stronger fact – forgoing pleasures, always appealing to this picture.” (LA: 54)

“[S]uppose we said that a certain picture might play the role of constantly admonishing me, or I always think of it. Here, an enormous difference would be between those people for whom the picture is constantly in the foreground, and who just didn’t use it at all.” (LA: 56)

Und wieder:

“‘God’s eye sees everything’ - - I want to say of this that it uses a picture [...] Of certain pictures we say that they might just as well be replaced by another”, während im Falle religiöser Bilder: “The whole *weight* may be in the picture” (LA: 71, 72).

Religiöse Bilder haben „ihre Heimat“ in einer bestimmten Lebensweise, in einer bestimmten Form des Lebens. In ihnen fühlt man sich heimisch, man benutzt sie mit Recht, wer auf eine bestimmte Art und Weise lebt. Gleichzeitig spielt das Bild gerade bei der Ausrichtung einer bestimmten Lebensweise, bei der Regulierung des Handelns des Gläubigen eine zentrale Rolle. Diese Rolle beeinflusst das gesamte Leben bestimmter Menschen. In Putnams Worten: “Pictures are important in life. The whole weight of a form of life may lie in the pictures that that form of life uses” (Putnam 1992: 156).

Wer ständig mit dem Bild eines Jüngsten Gerichts im Hintergrund lebt, wird, um ein Beispiel für das Gesagte zu geben, auf bestimmte Vergnügungen und Handlungen verzichten: Er wird also entsprechend dem Gebrauch, den er von diesem Bild macht, sein Leben moralisch regulieren.

Ich möchte noch zwei weitere grundlegende Gedanken hervorheben, die sich aus den Wittgensteinschen Überlegungen ergeben, und die schon von einigen PhilosophInnen beobachtet wurden:

(a.) Bilder sind in den meisten Fällen durch andere Bilder oder durch andere Beschreibungen ersetzbar. Die Bilder, die die Gläubigen verwenden, sind dagegen oft „irreplaceable pictures“ (Diamond 2005: 118, Schönbaumsfeld 2007: 182 folg.). Und sie sind in mindestens zweierlei Hinsicht unersetzlich. Zum einen sind sie unersetzlich, weil sie kein Abbild von etwas sind, zu dem wir unabhängig vom Abbild Zugang haben. Im religiösen Fall haben wir nur das Bild, ohne die Möglichkeit, direkt zu sehen, was das Bild darstellt (LA: 66-72). Zum anderen sind sie unersetzlich, da nur dieses bestimmte Bild, dieser bestimmte Ausdruck, eine besondere und einzigartige Rolle im Leben eines Menschen spielt, so dass die Person, die dieses bestimmte Bild benutzt, nicht willens oder in der Lage ist, es durch ein anderes zu ersetzen (LA: 70-72; DB: 80; PU 2009: § 531, S. 227-228; VB 1984: 567).

(b.) Das Gewicht, das Wittgenstein dem Bild in diesem Zusammenhang beimisst, ist mehrmals als eine reduktionistische und daher unbefriedigende Auseinandersetzung mit religiösen Fragen interpretiert worden. Einen ersten Versuch, Wittgensteins Äußerungen so zu interpretieren, finden wir bereits am Ende des LA. Hier scheint Smythies, ein bei der Vorlesung anwesender Student, genau diesen antireduktionistischen Einwand im Sinn zu haben: “This isn’t all he [the believer] does - associate a use with a picture”. Hier ist es erhellend, Wittgensteins Antwort auf Smythies im Sinn zu haben:

“When I say he’s using a picture I’m merely making a *grammatical* remark: [What I say] can only be verified by the consequences he does or does not draw. [...] If I wished to say anything more I was merely being philosophically arrogant” (LA: 72).

Zu zeigen, wie bestimmte Bilder verwendet werden, diese Bilder zu beschreiben; darüber hinaus zu beschreiben, welche Konsequenzen bestimmte Menschen daraus ziehen: Das sind keine philosophisch arroganten Untersuchungen, aber eben auch keine reduktionistischen. Denn “all [the philosopher] wished to characterize was the conventions [the believer] wished to draw” (LA: 72), d.h. die Aufmerksamkeit auf das zu lenken, was für eine philosophische Analyse relevant sein kann.

2. Die Rolle von „Bildern“ in J. H. Newmans *Grammar of Assent*

Der Begriff „image“ spielt in Newmans Werken eine zentrale Rolle, so sehr, dass in der Literatur sogar von einer ‚Theologie der religiösen Imagination‘

gesprochen wird (Dive 2018: 447, 452). GA bildet sicherlich keine Ausnahme, denn das Thema *image* und *imagination* taucht im gesamten Buch immer wieder auf.

Trotz der Bedeutsamkeit dieses Themas stoßen wir bei der Auslegung von Newmans Schriften sofort auf zwei Schwierigkeiten, die es uns sehr schwer machen, den Gebrauch, den er von diesem Begriff macht, und die Funktionen, die er ihm zuweist, genau und vollständig zu erläutern. Die erste Schwierigkeit besteht darin, dass die Schriften Newmans nicht als Schriften der systematischen Theologie zu interpretieren sind; und auch terminologisch sind sie keine Erben der Scholastik. Vielmehr sind sie immer im Lichte eines spezifischen argumentativen Kontextes zu verstehen, der sich von Fall zu Fall ändern kann (Steeves 2022: 96-99). Die zweite Schwierigkeit besteht darin, dass „Newman did not set out a definition of ‘the imagination’“, auch nicht von „image“ (Dive 2018: 14).

Angesichts dieser Schwierigkeiten beschränken wir uns hier darauf, in den Grundzügen zu skizzieren, was er mit diesem Begriff meint, um dann die Rolle zu analysieren, die er in GA spielt.

In GA führt Newman die Rede von „Bildern“ mit der Unterscheidung zwischen „begrifflichem und realem Erfassen“ ein, die für das Verständnis des gesamten Werkes grundlegend ist.

„Apprehension [...] is simply an intelligent acceptance of the idea, or of the fact which a proposition enunciates“ (GA: 36). Dieses Erfassen wird als „begrifflich“ bezeichnet, wenn es Sätze zum Objekt hat, die aus allgemeinen Namen bestehen, die sich nicht auf individuelle, spezifische Inhalte, sondern auf allgemeine, abstrakte Inhalte beziehen. „Real“ ist es dagegen, wenn es Sätze zum Objekt hat, die aus Eigennamen bestehen, aus Bezeichnungen, die für konkrete Dinge außerhalb von uns stehen, für Einheiten und Individuen (GA: 36-48). Ein und derselbe Satz kann gleichzeitig von einer Person begrifflich und von einer anderen real aufgefasst werden: “Words which are ‘a mere common-place’, an ‘expression of abstractions’ to one person, may bring a ‘living image’ before the imagination of another person” (Ker 2009: 638).

Das Erfassen eines Gedichtes zum Beispiel ändert sich erheblich, je nachdem, ob der Leser die Wörter als allgemeine, abstrakte Bezeichnungen erfährt oder

ob er stattdessen „lebendige Bilder“ mit ihnen assoziiert, die spezifisch und individuell sind und mit seiner persönlichen Erfahrung zu tun haben. Es ist diese Art des Erfassens, die das Lesen nicht nur zu einer intellektuellen Erfahrung macht, sondern zu einer ganzheitlichen, affektiven und involvierenden Erfahrung. „Image“ wird von Newman also in einer Rolle eingeführt, die der von „Notion“ entgegengesetzt ist. Während es sich bei letzterem um eine abstrakte, allgemeine Idee handelt, ist das Bild etwas Individuelles, Einzigartiges und Spezifisches. Diese Bilder sind dem Geist durch die Erinnerung oder die Vorstellungskraft gegeben. Das Begreifen wird konkret, gerade weil der Inhalt eines Satzes, der abstrakt und allgemein ist, durch Bilder lebendig gemacht wird (GA: 36-48).

Zwei Missverständnisse, die für Newman in dieser Hinsicht unbedingt vermieden werden müssen, sind: (a.) die Betrachtung von Bildern als etwas Abstraktes, als allgemeine Begriffe, und (b.) die Betrachtung von Bildern als etwas, das sich immer und nur auf die Sinne bezieht, als ob sie verblasste visuelle Wahrnehmungen wären. (a.) Bilder sind keine Abstraktion, da sie mit der Erfahrung verbunden sind und immer etwas Individuelles, Spezifisches im Leben des Menschen, der sie hat, repräsentieren (GA: 40). Gleichzeitig ist es aber (b.) ein Missverständnis zu glauben, dass deshalb jedes Bild immer nur mit den Sinnen zu tun hat und dass es immer eine Art leicht verblasster Wahrnehmung ist. Ian Ker unterlässt es nicht, darauf hinzuweisen: “‘A mere common-place, a terse expression of abstractions’ may be ‘the record of experiences, a sovereign dogma, a grand aspiration, inflaming the imagination, piercing the heart’. In other words, it may indeed be the object of a real assent. For in this case the general idea becomes ‘a living image’ – though there is no mental image of a visual or sensory kind” (Ker 2009: 641).

Untersuchen wir nun genauer die Rolle, die das Bild für Newman speziell in Bezug auf den religiösen Kontext spielt. Zunächst ist zu bemerken, dass sein allgemeines Interesse auf ein reales und nicht auf ein begriffliches Erfassen und Zustimmung zu religiösen Inhalten gerichtet ist, denn das reale Erfassen ist das, worauf es im religiösen Fall ankommt. In GA versucht er nämlich zu verdeutlichen und so plausibel wie möglich zu machen, dass es Menschen gibt, die religiöse Aussagen real erfassen und ihnen daher auch real zustimmen. In

diesem Zusammenhang spielen Bild und Imaginationskraft als wesentliche Elemente des realen Begreifens (und Zustimmens) eine zentrale Rolle im gesamten Werk.

“The heart is commonly reached, not through the reason, but through the imagination, by means of direct impressions, by the testimony of facts and events, by history, by description. (GA: 89) “But is found, through His preachers, to have imprinted the Image or idea of Himself in the minds of his subjects individually; [...] It is the Image of Him who fulfils the one great need of human nature, the Healer of its wounds, the Physician of the soul, this Image it is which both creates faith, and then rewards it” (GA: 359).

Für Newman hat das Christentum sein eigentliches Leben in dem realen Erfassen des Bildes von Jesus als Christus (Dive 2018: 1).

Eine weitere grundlegende Funktion, die das Bild erfüllt, ist die, das moralische Leben des Gläubigen und der Gemeinschaft der Gläubigen zu beleben, zu regeln und zu motivieren.

“And that Image, apprehended and worshipped in individual mind, becomes a principle of association, and a real bond of those subjects on with another, who are thus united to the body by being united to that Image; and moreover that image [...] is their moral life” (GA: 359). “These images, when assented-to, have an influence both on the individual and on society, which mere notions cannot exert” (GA: 60).

Oder, obwohl er das Wort „image“ hier nicht verwendet, meint er aber eindeutig:

“Catholic Religion is true, because its objects, as present to my mind, control and influence my conduct as nothing else does” (GA: 174).

3. Gemeinsame philosophische Sensibilitäten und spezifische Treffpunkte

Offensichtlich gibt es Unterschiede in der Verwendung des Bildbegriffs bei Wittgenstein und Newman. Es genügt z. B., zu berücksichtigen, dass dieser Begriff in GA mit der Untersuchung verschiedener Arten von kognitiven Akten (Erfassen und Zustimmungen) verbunden ist, was offensichtlich wenig mit Wittgensteinschem Denken zu tun hat. Außerdem versteht Newman „Bilder“

als Objekte, die mit dem Verstand oder der Vorstellungskraft erfasst werden können, eine Position, die eindeutig unvereinbar ist mit den Ansichten Wittgensteins, der sich stattdessen auf den Gebrauch konzentriert, den wir von einem Bild machen (ob es nun mental ist oder nicht). Auch in der Verwendung des englischen Wortes „picture“ durch Wittgenstein anstelle des Wortes „image“ unterscheidet er sich von Newman.

Sehr klar ist allerdings, dass es eine starke Verwandtschaft der beiden über dieses Thema gibt, die (a.) eine gewisse gemeinsame philosophische Sensibilität im Umgang mit dem Thema Religion und darüber hinaus (b.) einige spezifische gemeinsame Überlegungen von philosophischem Interesse zeigen.

(a.) Wir haben bereits gesehen, wie Wittgenstein am Ende der LA von Smythies des Reduktionismus bezichtigt wird. Wir könnten uns einen realistischen Religionsphilosophen vorstellen, der mit der informellen Art und Weise, wie Newman seine Forschung betreibt, und mit der zentralen Wichtigkeit, die er der Imaginationskraft und dem Bild beimisst, nicht zufrieden ist. Um einen solchen Vorwurf noch plausibler zu machen, braucht man sich nur das eben zitierte Zitat aus GA zu vergegenwärtigen, in dem der Begriff der Wahrheit mit der Fähigkeit eines Bildes, das Verhalten der Gläubigen zu regulieren oder zu beeinflussen, verknüpft wird. Oder man beachte, wie Newman am Ende seines Buches die Ergebnisse seiner allgemeinen Untersuchung der Gewissheit auf den spezifischen Fall der Religion anwendet. Die Argumente, die er hier vorbringt, sind informell, nicht demonstrativ im strengen Sinne, sondern eher narrativ. Das Gewicht seiner gesamten Argumentation liegt gerade auf der Erzählung und der Macht, die bestimmte Bilder haben, um unser Leben zu beschreiben und zu regeln. Er scheint also den religiösen Inhalt auf eine narrative Ebene verlagern zu wollen. Newman antwortet auf ein solches Anliegen, indem er unsere Aufmerksamkeit auf das lenkt, was für seine Analyse hier tatsächlich relevant ist: “This is why science has so little of a religious tendency; deductions have no power of persuasion. The heart is commonly reached, not thorough the reason, but through the imagination [...] Many a man will live and die upon a dogma: no man will be a martyr for a conclusion” (GA: 98). Niemand leugnet hier irgendetwas, auch wenn man diesen Eindruck hat. Es ist unsere Analyse, die auf das ausgerichtet werden muss, was hier am relevantesten ist (PU 2009: § 304). Unter diesem

Gesichtspunkt erkennt auch Newman wie Wittgenstein an, dass es im Fall des religiösen Glaubens relevanter und überzeugender ist, sich mit der Kraft, den Implikationen und der Rolle zu befassen, die bestimmte Bilder im Leben und tatsächlichen Verhalten der Gläubigen haben, als sich in metaphysische Spekulationen zu vertiefen. In diesem Sinne erweist sich Newman also auch als philosophisch nicht arrogant. Beide vermeiden es, das religiöse Thema zu einer „wissenschaftlichen“ oder metaphysischen Angelegenheit zu machen, wobei sie allerdings das Risiko in Betracht ziehen, deswegen als Reduktionisten missverstanden zu werden.

(b.) Wenn wir nun auf die einzelnen Gemeinsamkeiten von Wittgenstein und Newman blicken, stellen wir zunächst fest, dass das religiöse Bild für beide eine grundlegende Rolle bei der Belebung, Regulierung und Beeinflussung des moralischen Lebens des Gläubigen spielt. Das Bild des Jüngsten Gerichts wie das von Christus oder das des allsehenden Auges Gottes usw. orientieren ein Wertesystem und beeinflussen ständig die Handlungen des Gläubigen. Bestimmte Bilder haben eine Lebendigkeit und Anziehungskraft, die sie zu einem lebendigen und regulierenden Prinzip für das gesamte ethische Leben des Gläubigen machen. Die Analyse einer religiösen Überzeugung kann also zu einer Beschreibung der Funktion und Rolle werden, die ein bestimmtes religiöses Bild innerhalb einer bestimmten Lebensweise spielt. Natürlich muss auch hier, wie bereits erwähnt, jeder relativistische oder reduktionistische Interpretationsversuch zurückgewiesen werden. Wesentlich bleibt jedoch, dass die ganze Betonung gerade auf ein bestimmtes Bild fallen kann, ein Bild, das unersetzlich ist, „irreduzibel“ zu einem anderen Ausdruck oder Bild (Schönbaumsfeld 2023: 46-53). Um Newmans Jargon zu verwenden: jedes Bild ist einzigartig, spezifisch. Verschiedene Bilder laden dazu ein, auf unterschiedliche Weise verwendet zu werden, das Leben auf unterschiedliche Weise zu regeln. So wird es manchmal der Fall sein, dass einige von ihnen nicht austauschbar sind.

Und schließlich betone ich ein weiteres und wichtiger Aspekt der Gemeinsamkeit. Der richtige Gebrauch und das richtige Verständnis dieser Bilder sind von einer bestimmten Lebensweise abhängig. Das „lebendige Bild“ eines Erlösers, des Erwarteten, kann z.B. von demjenigen zu Recht verstanden und verwendet werden, der sich wirklich als Sünder erkennt und bereits in Erwartung der Erlösung lebt (in GA: Kap. X, deutet er die Naturreligion der

Antike als eine Erwartung des Christusbildes). Eine bestimmte Lebensweise macht, könnten wir mit Wittgenstein sagen, bestimmte Bilder akzeptabel oder sogar notwendig (DB: 75, 80).

References

Diamond, Cora (2005), *Wittgenstein on Religious Belief: The Gulfs between Us*, in *Religion and Wittgenstein's legacy: a systematic study*, (ed.) D.Z. Phillips and Mario von der Ruhr, Oxon: Routledge.

Dive, Bernard (2018), *John Henry Newman and the Imagination*, London: Bloomsbury.

Ker, Ian (2009), *John Henry Newman*, Oxford: Oxford University Press.

Kfourir, André (2023), *Further Affinities between Wittgenstein and Newman*. In: Alois Pichler, Esther Heinrich-Ramharter, Friedrich Stadler, in cooperation with Joseph Wang-Kathrein (Hg.): *100 Years of Tractatus Logico-Philosophicus – 70 Years after Wittgenstein's Death. A Critical Assessment. Beiträge der Österreichischen Ludwig Wittgenstein Gesellschaft / Contributions of the Austrian Ludwig Wittgenstein Society. Band / Vol. XXIX. Kirchberg/W.*, 390–397.

Kienzler, Wolfgang (2006), *Wittgenstein and John Henry Newman on Certainty*, in *Grazer Philosophische Studien* 71, 117–138.

Newman, John, H. (2003), *An Essay in Aid of a Grammar of Assent*, Notre Dame USA: University of Notre Dame Press.

Pichler, Alois (2024), *Glaube und Aberglaube nach Wittgenstein: Zu nonkognitiven Deutungen der Grammatik des religiösen Glaubens*, in *Religionsphilosophie nach Wittgenstein Sprachen und Gewissheiten des Glaubens*, Esther Heinrich-Ramharter (Hrsg.), Berlin: Springer Nature.

Pichler, Alois and Sunday Grève, Sebastian (2024), *Cognitivism about Religious Belief in Later Wittgenstein*, *International Journal for Philosophy of Religion* 97 (1).

Pritchard, Duncan (2015) *Wittgenstein on Faith and Reason: The Influence of Newman*, in: Mirosław Szatkowski (ed.), *God, Truth and other Enigmas*, Berlin: De Gruyter.

Pritchard, Duncan (forthcoming), *Quasi-fideism and virtuous anti-evidentialism: Wittgenstein and Newman on knowledge and certainty*, in *Newman and Contemporary Philosophy*, (eds.) J. Milburn & F. Aquino, London: Routledge.

Putnam, Hilary (1992), *Wittgenstein on religious belief*, in *Renewing Philosophy*, Harvard University Press, 134–57.

Schönbaumsfeld, Genia (2007), *A confusion of Spheres. Kierkegaard and Wittgenstein on Philosophy and Religion*, Oxford: Oxford University Press.

Schönbaumsfeld, Genia (2023), *Wittgenstein on Religious belief*, in *Elements in the Philosophy of Ludwig Wittgenstein*, Cambridge: Cambridge University Press.

Steeves, Nicolas (2022), *Reaching Hearts and Minds through the Imagination: Newman and the Reality of God*, in *John Henry Newman. Welt Gottes und Wahrheit des Menschen*, Becker, P., Schlosser, M., Wodrazka, P., B. (Hrsg.), Freiburg im Breisgau: Herder Verlag, 90-110.

Wittgenstein, Ludwig (PU 2009), *Philosophical Investigations / Philosophische Untersuchungen*. Hg. Von P. M. S. Hacker und Joachim Schulte, übersetzt von G. E. M. Anscombe, P. M. S. Hacker und Joachim Schulte. New York: Wiley.

Wittgenstein, L. (ÜG 1984), *Über Gewißheit* In: Werkausgabe Bd. 8, S. 113–257. Hg. Von G. H. von Wright. Unter Mitarbeit von Heikki Nyman. Frankfurt am Main: Suhrkamp.

Wittgenstein, L. (VB 1984), *Vermischte Bemerkungen*. In: Werkausgabe Bd. 8, S. 445–575. Hg. Von G. H. von Wright. Unter Mitarbeit von Heikki Nyman. Frankfurt am Main: Suhrkamp.

Wittgenstein, L. (DB 1997), *Denkbewegungen. Tagebücher 1930–1932, 1936–1937* (MS 183). Hg. Von Ilse Somavilla. Teil 1: Normalisierte Fassung. Innsbruck: Haymon.

Wittgenstein, L. (LA 1966), *Lectures and Conversations*. Hg. von Cyril Barrett. Oxford: Basil Blackwell.

Acknowledgement

Ich bedanke mich bei Professor Wolfgang Kienzler, mit dem ich die Gelegenheit hatte, mehrmals über Newman und Wittgenstein zu sprechen. Weiterhin danke ich meinem Kollegen und Freund Vincent Nitzsche für die geduldige Korrektur der deutschen Sprache in diesem Artikel.

A Relativist Ameliorative Approach to Gender Terms

Dan Zeman (Porto, Portugal)

Abstract

Gender terms play a crucial role in our lives, in that they allow us to categorize ourselves and others as of a certain gender, which in turn has important social, moral and legal implications. In contemporary philosophy of language, the discussion of the meaning of gender terms has been mostly conducted within ameliorative projects (Haslanger (2000)). Roughly, an ameliorative project investigates what the meaning of words should be, in opposition to a descriptive project which investigates how meanings are. The main idea is that philosophy of language should strive to provide accounts of “loaded” terms that are sensitive to social, moral, and political ideals: e.g., avoid discrimination, bring about more justice etc.

In this paper, I propose a novel ameliorative account of gender terms. By focusing entirely on the case of trans people (who didn’t undergo gender-affirming medical procedures), I show that one of the main views in the semantics of natural language expressions (relativism) can be successfully appealed to in order to yield the ameliorative account sought after. However, since an orthodox version of the view is unsuitable, I propose a *flexible* version of relativism, one that is independently motivated by various considerations in the literature on perspectival expressions such as predicates of taste, moral and aesthetic terms, epistemic modals (e.g., problematic cases of truth-assessment, data about retraction, etc.). I base the flexibility in question on the general notion of *the importance of the asserter’s situation*, which in application to gender terms is replaced with the notion of *the situation of the subject of the gender attribution*. With this latter notion in hand, all the problematic cases for the orthodox relativist view (and for the other views on the market) are handled satisfactorily.

Gender terms play a crucial role in our lives, in that they allow us to categorize ourselves and others as of a certain gender, which in turn has important social, moral and legal implications. In contemporary philosophy of language, the discussion of the meaning of gender terms has been mostly conducted within ameliorative projects (Haslanger (2000)). Roughly, an ameliorative project investigates what the meaning of words should be, in opposition to a descriptive project which investigates how meanings are. The main idea is that philosophy of language should strive to provide accounts of “loaded” terms that are sensitive to social, moral, and political ideals: e.g., avoid discrimination, bring about more justice etc.

In this paper, I will focus almost entirely on the case of trans people (who didn’t undergo gender-affirming medical procedures). In general terms, an ameliorative project about gender terms in relation to trans people should fulfil at least the following two desiderata: i) to allow trans people to use the gender terms of their choice and ii) to assure that this choice be respected by

others. More specifically, I cash out this project as allowing utterances of sentences like “(G) John is a man.” (or their first-person equivalents “I am a man.”), where John (or the speaker) is a trans man who hasn’t undergone gender-affirming medical procedures, to come out as true at least in some contexts in which a) trans people attribute themselves a certain gender; b) other people attribute trans people a certain gender.

1 A quick survey of the main views and their shortcomings

When it comes to the semantics of gender terms, several views can be found in the literature, among which invariantism, contextualism and relativism. As I will show in this section, none of these (at least in their simple versions) are suitable for the type of ameliorative project I aim to undertake.

Two versions of invariantism have been popular in relation to gender terms: *biological invariantism*, according to which gender is reducible to biological traits ((G) is true iff John possesses biological traits X, Y, Z ...) and *social invariantism*, according to which gender is reducible to social roles ((G) is true iff John plays social role R1, R2, R3 ...). Both versions are problematic from an ameliorative point of view: due to its focus on biology, the former cannot help with the inclusion problem, as (G) will come out false in all contexts; due to its focus on social elements, the latter yields the wrong results if John doesn’t play any of the roles envisaged by the theory and thus (G) comes out false in some contexts in which it shouldn’t.

Contextualism has entered the debate as a way to solve the problems faced by invariantism, and also motivated by variation in truth-values of sentences like (G) across contexts (perceived to be true in, say, activism contexts, when self-identification matters, but false in medical contexts, when female biology is relevant). Building on previous views, Saul (2012) discusses a contextualist view that provides the following truth-conditions: “[(G)] is true in a context C iff X is human and relevantly similar (according to the standards at work in C) to most of those possessing all of the biological markers of [male] sex.” Generally formulated, contextualism takes the determining factor for the truth value of sentence like (G) to be found in the context of utterance. But, as Saul (2012) herself and many critics (e.g., Bettcher (2017), Díaz-León (2016), Jenkins (2016)) acknowledge, this type of contextualism doesn’t help with the inclusion problem either, since it allows (G) to come out false in the mouth of

transphobes (it is their context that determines its truth-value). Another version of contextualism, proposed by Díaz-León (2016) has it that the factor determining the truth-value of (G) is the situation of the subject of the gender attribution; this “subject-contextualism” fares better than “speaker-contextualism”, but it has been criticized on various counts (by, e.g., Laskowski (2020), Zeman (2020b), Chen (2021), etc.).

Relativism has not been applied to gender terms until recently. Generally formulated, relativism takes the determining factor for the truth value of sentence like (G) to be found not in the context of utterance, but in that of assessment (thus, treating gender terms as *assessment-sensitive*). Zeman (2020a) has explored how the view fares vis-à-vis the ameliorative project assumed here, and his findings were mixed: while it gets some good results (e.g., accounts for cases in which (G) is uttered by a transphobe but assessed by a trans person or an ally or makes sense of trans people taking themselves to have been wrong about their gender before transition), it also fails to improve on contextualism (it cannot account for cases in which (G) is uttered by a trans person or an ally, but assessed by a transphobe; or for the simple case in which a transphobe would utter (G): since the context of utterance and that of assessment are the same, it will come out as false in their mouth).

The conclusion is bleak: three of the main semantic views are not suitable for the type of ameliorative project pursued. In the remainder of the paper, however, I will show that a suitably modified version of relativism might give us what we want.

2 Motivating Flexible Relativism

Relativism has come to prominence in connection to the semantics of perspectival expressions – terms like “tasty”, “beautiful”, “good” and their negative counterparts, (epistemic) “might”, “knows”, etc. Among the phenomena that relativism has been taken to offer a better treatment than rival views are: eavesdropping (Egan, Hawthorne & Weatherson (2005)); faultless disagreement (Kölbel (2004), Lasersohn (2016)); retraction (MacFarlane (2014)). However, it has become clear in the debate that a rigid, orthodox version of relativism is not tenable anymore. The main challenges that relativism faces come from certain cases of truth-value assessment in which relativism yields the wrong results and from recent arguments against

the support relativism has been taken to receive from retraction. In relation to the first, Dietz (2008) and others have presented cases of so-called “knowledgeable assessors” involving certain exchanges with epistemic modals (see the scenarios about Mastermind in von Fintel & Gillies 2008) where certain utterances are felicitous, yet according to relativism they shouldn’t be. In connection to the second, in addition to many critical considerations “from the armchair” (von Fintel & Gillies (2008), Dowell (2011), Rafmann (2016), Marques (2018)), several recent experimental studies show either that the connection between retraction and truth-value assessment is not as tight as the relativist assumes (Knobe & Yalcin (2014)) or that people in fact don’t retract (Kneer (2021)).

Both challenges can be met if relativism is construed as a flexible position. While the view has been usually conceived of as a rigid position, according to which the relevant truth-determinative factor, always initialized in the context of the assessor, is the one relevant in that very context. But flexible versions of relativism are also available (see, for example, Zeman (2010, 2024), MacFarlane (2011, 2014), Khoo (2015), Beddor & Egan (2018) etc.); according to such versions, the assessor can “step out” of their context of assessment and sometimes adopt the perspective that is relevant in the context of utterance, when the initial assertion was made (regardless of whether the initial assertion was made by themselves or by a different person). The flexibility inherent in such versions of relativism solves the challenges mentioned (see MacFarlane (2011, 2014) in relation to “knowledgeable assessors” scenarios; see Beddor & Egan (2018) in relation to Knobe & Yalcin’s (2014) results; see Zeman (2024) in relation to Kneer’s results).

But now another challenge arises – that of finding a non-ad-hoc principle underlying the shift from the context of assessment to that of utterance (see Dowell (2011), Kneer (2021)). One ready answer is the literature is that of Beddor & Egan (2018), who take the main factor that drives flexibility to be the question under discussion. But they also agree that this cannot be the only principle, and they mention charity as another option. Other factors can be envisaged: for example, in explaining someone’s beliefs or actions in a certain context, appeal to the body of information of that person at the time of believing/acting seems the obvious choice. Given this, it seems to be better to postulate some kind overarching principle: in Zeman (2024) I propose a meta-

principle based on the notion of *the importance in a context of assessment of the situation of the person making the assertion assessed* (“the importance of the asserter’s situation”, for short). This allows for some context-sensitivity in what exactly the flexibility-driving factor is, but it addresses the ad-hocness challenge.

3 Application to gender terms

Having motivated a flexible version of relativism and having proposed a solution to the ad-hocness problem, in this section I apply the view to gender terms. At least two questions arise in this connection: what does importance amount to in this case and how does the view account for the problematic cases.

In looking for an answer to the first question, we should follow closely the goals of the ameliorative endeavor. Since these are to allow trans people to use the gender terms of their choice and this choice to be respected, *self-identification* should be part of the picture. So, the proposal is to include self-identification as one of the factors that makes the situation of the asserter important in a context of assessment. Granted, this seems to be a different aspect of importance, in comparison with those mentioned in the previous section. But my proposal was purposely *pluralistic*: various elements could be important, and arguably how one conceives of themselves is such an element.

In relation to the second issue, the view proposed accounts for the problematic cases, but not without (a slight) modification. First, the view yields the right result in the cases in which (G) (or, rather, the first-person equivalent) is uttered by John himself (in non-medical contexts) and assessed by a transphobe: since John’s self-identification makes his situation important in the context of assessment, it is the truth-determinative factor in his context that matters for the truth of (G). A complication that arises here, however: (G) could be uttered by someone else, who could be either a trans man or not, a transphobe or not. In such cases, the asserter’s situation is *not* the one relevant for the truth of (G); instead, I propose we should take that to be *the situation of the subject of the gender attribution*. (This is similar to Díaz-León’s (2016) subject-contextualist proposal, which seems to be equally available to the

relativist.) Going this way will also solve the cases in which (G) is uttered (and assessed) by a transphobe by making John's self-identification as a man what determines whether an assertion is true or not.

Further questions might arise. For example, can the view account for the perceived variation in truth-value of sentences like (G) across contexts, which would avoid the view's collapse into invariantism? It seems so: since it is the importance of the situation of the subject of the gender attribution that matters for truth-evaluation, and since self-identification is the crucial factor in the situation of the subject, it all comes down to how the subject self-identifies. So, if John self-identifies as a man in a context of activism, then (G), as uttered in such a context and assessed from any other, comes out as true. But if John doesn't self-identify as a man in a medical context, then (G), again as uttered in such a context and assessed from any other, comes out as false. We thus get the relevant variation in truth-value across contexts.

Thus, by embracing a flexible version of relativism and allowing the notion of importance to comprise not only the situation of the asserter, but that of the subject of the gender ascription as well, we arrive at a solution to the cases that proved to be problematic for the orthodox version, as well as for other views on the market. Taking gender terms to be assessment-sensitive (in the way suggested) is thus one way in which an ameliorative account of gender terms can proceed.

References

- Beddor, Bob; Egan, Andy (2018) "Might Do Better: Flexible Relativism and the QUD", *Semantics & Pragmatics* 11(7).
- Bettcher, Talia Mae (2017) "Trans feminism: Recent philosophical developments", *Philosophy Compass* 12(11).
- Chen, Hsiang-Yun (2021) "Contextualism and the Semantics of 'Woman'", *Ergo* 8: 20.
- Díaz-León, Esa (2016) "Woman as a politically significant term: A solution to the puzzle", *Hypatia* 31(2), 245-258.
- Dietz, Richard (2008) "Epistemic Modals and Correct Disagreement", in: Manuel Garcia-Carpintero and Max Kölbel (eds.), *Relative Truth*, Oxford University Press, 239-263.
- Dowell, Janice (2011) "A Flexible Contextualist Account of Epistemic Modals", *Philosophers' Imprint* 11(14).

Egan, Andy; Hawthorne, John; Weatherson, Brian (2005) "Epistemic modals in context", in: Gerhard Preyer and Georg Peter (eds.), *Contextualism in Philosophy: Knowledge, Meaning, and Truth*, Clarendon Press, 131-170.

von Fintel, Kai; Gillies, Anthony (2008) "CIA Leaks", *Philosophical Review* 117(1), 77-98.

Haslanger, Sally (2000) "Gender and race: (What) are they? (What) do we want them to be?", *Noûs* 34(1), 31-55.

Jenkins, Katharine (2016) "Amelioration and inclusion: Gender identity and the concept of woman", *Ethics* 126(2), 394-421.

Khoo, Justin (2015) "Modal Disagreements", *Inquiry* 58(5), 511-534.

Kneer, Markus (2021) "Predicates of Personal Taste: Empirical Data", *Synthese* 199, 6455-6471.

Knobe, Joshua; Yalcin, Seth (2014) "Epistemic Modals and Context: Experimental Data", *Semantics & Pragmatics* 7, 1-21.

Kölbel, Max (2004) "Faultless Disagreement", *Proceedings of the Aristotelian Society* 104, 53-73.

Lasersohn, Peter (2016) *Subjectivity and Perspective in Truth-Theoretic Semantics*, Oxford University Press.

Laskowski, N. G. (2020) "Moral constraints on gender concepts", *Ethical Theory & Moral Practice* 23(1), 39-51.

MacFarlane, John (2011) "Epistemic Modals are Assessment-Sensitive", in: Andy Egan and Brian Weatherson (eds.), *Epistemic Modality*, Oxford University Press, 144-178.

MacFarlane, John (2014) *Assessment Sensitivity: Relative Truth and Its Applications*, Oxford University Press.

Marques, Teresa (2018) "Retractions", *Synthese* 195, 3335-3359.

Raffman, Diana (2016) "Relativism, Retraction, and Evidence", *Philosophy and Phenomenological Research* XCII (1), 171-178.

Saul, Jennifer (2012) "Politically significant terms and philosophy of language: Methodological issues", in: Sharon L. Crasnow and Anita M. Superson (eds.), *Out from the Shadows*, Oxford University Press, 195-216.

Zeman, Dan (2010) "Knowledge attributions and relevant standards", in: François Recanati, Isidora Stojanovic and Neftali Villanueva (eds.), *Context-dependence, perspective and relativity*, Mouton de Gruyter, 225-250.

Zeman, Dan (2020a) "Invariantist, Contextualist, and Relativist Accounts of Gender Terms", *EurAmerica* 50(4), 739-781.

Zeman, Dan (2020b) "Subject-contextualism and the meaning of gender terms", *Journal of Social Ontology* 6(1), 69-83.

Zeman, Dan (2024). "Relativism and Retraction: The Case Is Not Yet Lost", in: Dan Zeman and Mihai Hîncu (eds.), *Retraction Matters. New Developments in the Philosophy of Language*, Springer, 71-98.

Planetary and inter-scalar entanglements of feminist new materialisms, exemplified by Astrida Neimanis's *Bodies of Water: Posthuman Feminist Phenomenology*

Wer Ziemian (Toruń, Poland)

Abstract

Feminist new materialisms (FNMs) offer tools for analysing global forms of oppression. Using their critical approach, it is possible to identify the hegemonic power structures that the Global North imposes on marginalised communities, and to propose new ways of thinking about living together on the planet. FNMs challenge dominant axioms, such as anthropocentrism, atomism or dualistic ontologies, which limit the possibilities of solving contemporary environmental and social crises and, as some scholars point out, these axioms may be at the root of current problems. In their place, FNMs introduce concepts of openness, responsibility and entanglement, exemplified by Karen Barad's ethico-onto-epistemology (Barad, 2007). I will present the theoretical framework of FNMs illustrating this approach by Astrida Neimanis's *Bodies of Water: Posthuman Feminist Phenomenology* (Neimanis, 2017). I will focus on the example given by Neimanis, in which she refers to the toxins substances produced in the centres of Global North that flow out of its areas. The example highlights the planetary entanglement of human bodies, more-than-human bodies, environment and power relations. The aim of the paper is to present potential of FNMs in exploring the interconnections and developing models of responsibility and coexistence in times of climate crisis.

How do we develop new models of responsibility and coexistence in times of climate crisis? How do we explore interconnections when we get rid of the conviction of fixed scales constituting reality? What new figurations can help us do this? We can look for answers to these questions within the feminist new materialisms (FNMs) movement. The work of those forming or attributed to this current can serve as a fascinating study to inspire next Arts of Noticing (Tsing, 2015) in processes of knowledge production, indicating how we can reinterpret materiality, influencing its shape and form. In order to present the potential of FNMs, I use Astrida Neimanis' work "Bodies of water. Posthuman feminist phenomenology", in which the assumptions of the FNMs resound, particularly those of entanglement and interdependence. In this paper I draw on English-language texts that I have also read in Polish. The citations I provide are from the texts in English (original or translated). In the case of works by Polish researchers, I present my own translation from Polish into English.

1. About Astrida Neimanis

Astrida Neimanis in the Polish version of her book is described as “a cultural theorist conducting research at the intersection of feminism and environmental humanities. Her interests focus on corporeality, water and atmospheric phenomena, through which she attempts to re-imagine relationality, justice, care and agency in times of climate crisis” (Neimanis 2024). Andrzej Marzec, Polish researcher in environmental humanities, writes of her work: „In her reflections, water becomes not only the guarantor of material, corporeal changeability, but above all of an always surprising, unfettered naturocultural creativity” (Marzec 2021: 133). A work that particularly indicates Neimanis's interests is the introduction to *Feminist, Queer, Anticolonial Propositions for Hacking the Anthropocene: Archive* titled *An Archive of an Epoch that Almost Was* (Neimanis 2021), in which, as Mateusz Borowski notes, Neimanis states that “we should ultimately accept that the scientific conception of the Anthropocene sustains not only the division between nature and society, but also the resulting practices of environmental and human resource exploitation” (Borowski, 2023).

2. About her work: *Bodies of Water. Posthuman Feminist Phenomenology*

One of the main questions of her book *Bodies of Water. Posthuman Feminist Phenomenology* is: „how to think our commonality as water bodies alongside, rather than against, a more specific politics of location” (Neimanis 2017: 21). How does the concept developed by Neimanis answer the question thus posed? In order to illuminate Neimanis' project, it is useful to understand what lies behind the title thus formulated. The second part of the title may cause confusion. What is this new assemblage of theories?

The posthumanist aspects of Neimanis's work include going beyond the human and not reducing the whole (of reality) to the activity or idea of the human. It is also consist of flat ontology (which means there are no hierarchical distinction of entities: the human/subjectivity is no longer “above” the object (f. ex. nature), but everything is entangled in the dependency networks of reality. A human can no longer be seen as an atomic being, but is “porous”, “not fully closed” and dependent on other causal instances. As Neimanis notes: “a more-than-human hydrocommons thus present a challenge to anthropocentrism, and the privileging of the human as the sole or primary site of embodiment. (...) we have never been (only) human” (Neimanis 2017: 2).

Neimanis's work methodologically draws from the phenomenology of Merleau-Ponty (Merleau-Ponty 2001), in which a key assumption was that subjectivity is embodied and exists in the world. The subject, as a body, has access to reality through phenomenal experiences, i.e., sensory and bodily experiences. It points to the instrumental and prosthetic nature of cognition, as also indicated by Neimanis, referring to Donna Haraway: "all knowledge is mediated. (...) phenomenology always comes from somewhere" (Neimanis 2017: 61). Merleau-Ponty pointed out the existence of only the corporeal sphere. Additionally, within the paradigm of being-in-the-world, he emphasized the active nature of philosophical work and his method as a form of description. Neimanis adopts assumptions of embodiment and situatedness, indicating that by starting from a phenomenological description, it is possible to capture multiple instances of agency and interconnections that transcend scales: "(...) if phenomenology asks us to understand the world as lived, we can only begin from a situated politics of location – albeit one whose spatiotemporal scale is torqued through posthuman relationalities and becomings" (Neimanis 2017: 25). Like Merleau-Ponty, she emphasizes the performative nature of the content created within the act of description.

Neimanis's phenomenology is also feminist because it draws from feminist thought – mainly referencing Luce Irigaray, whose water figuration interweaves feminist themes in a creative way (Neimanis 2017: 78). Neimanis also engages in dialogue with various scholars working within feminist new materialisms (such as the aforementioned Donna Haraway, Rosi Braidotti, Karen Barad) and refers to feminist reinterpretations of evolutionary theories (citing Grosz, Le Guin). She draws attention to forms of oppression and power relations (such as patriarchy, colonialism, racism, species chauvinism), which were not addressed in Merleau-Ponty's phenomenological project.

Her theory points to the continuous reconfiguration and differentiation of reality. Her work is queer and post-gender (Neimanis 2017: 129). It demonstrates sensitivity to the "otherness", which does not fit into the phallogocentric logic characterized by static and "given" forms as well as dualistic and hierarchical ontology. Water, in Neimanis's perspective, is a figure of openness, boundlessness, overflowing, spilling, and life. By using "water," she emphasizes our constitutive entanglement and porosity. Neimanis writes: "changing how we think about bodies means changing how we think

about water" (Neimanis 2017: 19), highlighting the necessity of developing new models of thinking about water that do not replicate the aforementioned theoretical positions. She references researcher Linton, who formulated the concept of modern water (Neimanis 2017: 19; Linton 2010). According to him, this is a perspective that emerged in the latter part of the 20th century, treating water as something abstract, neutral, chemically distinct (H₂O), quantifiable, and existing "outside" as a representation. Global water is an extension of modern water: now represented through depictions of its circulation in nature. In this way, it remains quantifiable, measurable, and abstract. Neimanis notes that this abstraction is linked to the portrayal of water as pure and transparent. She also refers to the issue of bottled water, which intertwines concerns of ownership, consumption, capitalist and colonial logic (Neimanis 2017: 178). In the perspective developed by Neimanis, the socialization of water is crucial. This premise underscores the universality of water – meaning that water is everywhere. She speaks of hydrocommons arising from its multi-scalarity (Neimanis 2017: 86). Again, the figure of water breaks the dualism between the abstract human and the abstract water "beyond them". Watery bodies are permeable on different levels and scales. Neimanis' work counters the notions of modern, global, and abstract water. By using the figure of water, she creates new connections, emphasizing community, relationality, dependence, and our entanglement in the world – the environment.

3. About Feminist New Materialisms

To describe FNMs, the "ontological map of the world (from the perspective of a new materialist feminist)" proposed by Monika Rogowska-Stangret proves useful (Rogowska-Stangret 2021: 123-124). The researcher lists several key points that outline the assumptions of FNMs. These include flat ontology, the absence of an external point of reference, the lack of an "objective" perspective, distributed agency, and the implosion of the centre.

I have already mentioned flat ontology when discussing the "posthumanist" aspect of the title. The next two points refer to the concept of situated knowledge—the idea that knowledge is always formulated from a specific standpoint and is inherently perspectival. There is no neutral, external vantage point from which reality can be "observed" and "objectively" described, which aligns with the ontology of immanence. The point on

"distributed agency" challenges the idea that agency belongs exclusively to humans or specific individual agents. Agency is understood broadly because reality is a network of interdependencies. Rather than being attributed to isolated entities, agency should be considered in terms of intensities and material entanglements. This connects to the final point, which concerns the absence of a fixed centre from which causality, knowledge, or scale emanate. This idea can be understood through the Deleuzian-Guattarian contrast between a tree and a rhizome. Deleuze and Guattari argue that some books resemble trees, branching out from a single point into successive "branches," while others resemble rhizomes—decentralized, dynamic networks of connections and references (Deleuze and Guattari 1987). Similarly, the "implosion of the center" can be illustrated through Marder's work, where he contrasts the image of a human as having a "center" in the head—localized and privileged—with the image of plants as entangled wholes, entirely dependent on their surroundings (Marder 2013: 64).

An important concept developed within feminist new materialisms is "ethico-onto-epistemology", that was introduced by Karen Barad (Barad 2007). It emphasizes the necessary entanglement of ethical, ontological, and epistemological aspects, particularly in scientific practice. Scientific "objects" are not pre-existing, passively awaiting "discovery and description" by a subject; rather, the very choice of research apparatus, methodology, material conditions, ideology, etc., influences what is observed and described. The researcher and the object of study are not independent of each other—scientific practice is a form of co-participation in the world. Barad articulates this stance as follows: "[M]aking knowledge is not simply about making facts but about making worlds, or rather, it is about making specific worldly configuration (...) in the sense of materially engaging as part of the world in giving it specific material form" (Barad 2007: 91).

Neimanis refers to other FNMs scholars, stating: "We could say, using the language of Haraway or Karen Barad, that these various bodily interfaces – biology and mood and culture and context – are always co-worlding the phenomenon we come [to] know as our bodies. Rather than two separate entities interacting, they intra-act; they become what they are only in relation.

Co-worlding is always a collaborative process, and always emergent. The thing called 'the body' (...) is a congeries of other bodies, and always on the move" (Neimanis 2017: 34).

Within the framework of ethico-onto-epistemological entanglement and the aforementioned "implosion of the centre," the world can no longer be perceived as divided into successive emergent levels of constitution. An example of such a hierarchical vision would be one where logic and mathematics form the foundation, followed by physics, chemistry, biology, and then culture. Monika Rogowska-Stangret discusses the problem of scales in an article on epigenetics (Rogowska-Stangret 2023), also referring to Noela Davis's work on this topic (Davis 2017). Similarly, traditional scales such as "local" and "global" are also abandoned, as Neimanis illustrates in her work: "It is about paying attention to the complication of scale, where a familiar deictics of 'here' and 'there', 'mine' and 'ours', even 'local' and 'global', or 'now' and 'then', which might have once seemed relatively securable, are now queerly torqued. Time, place, and bodies are all caught in the warp and woof of planetary colonialities that are naturalcultural and diffracted, but still racialized and gendered, all the same" (Neimanis 2017: 37). Inter-scalar entanglements are made visible in Neimanis's work through examples related to toxicity.

4. Toxicity, becoming-milieu

Neimanis refers to a case documented in the *Mother's Milk* project, led by midwife and activist Katsi Cook. The project investigated toxic sludge discharged by General Motors (a major automotive industry player), which ultimately seeped into wells located on properties adjacent to the Akwesasne Mohawk reservation. The study revealed that: "[w]omen living on the reservation and eating fish from the St Lawrence River had a 200 per cent greater concentration of PCBs in their breast milk in comparison to the general population" (Neimanis 2017: 35). This example illustrates colonial practices of displacing waste from "civilization" to what is considered "beyond". However, it does not solely depict the impact of human activities: "The permeability of the ground, the path of the river, the appetites of the fish, all become caught in these currents" (Neimanis 2017: 35).

Neimanis also discusses another case that highlights the planetary scale of interconnected (including colonial) power relations through the displacement

of waste, as well as the global environmental consequences of mass industry. Ocean currents carry pollutants to the Arctic, contaminating fish and plankton, which are then consumed by marine mammals. These toxins accumulate in the fat of marine mammals, meaning that they may become more “contaminated” than the surrounding environment. The consumption of fish and marine mammals by Inuit communities leads to contamination of Inuit women’s breast milk. As Neimanis writes: “[a]s a result, the breast milk of Inuit women in the Canadian Arctic contains two to ten times the amount of organochlorine concentrations of samples from white women hundreds of kilometres to the south; PCB levels are also alarmingly high” (Neimanis 2017: 35-36).

Neimanis asserts that: “when an Inuit woman nurses her young, her transcorporeal gift is laced with a specific colonial politics of location – but one where the effects of colonial incursion do not require direct proximity to a colonizer” (Neimanis 2017: 36). The toxicity illustrated in the examples above is what breaks through scales and directs attention to the ontology of entanglements proposed within FNMs. It sensitizes us to mutual constitution and the impossibility of separation: “These bodies are all caught up in one another’s currents – as they are with the whale’s body, the body of the rain cloud, and the body of the increasingly toxic sea. As bodies of water, we are always, at some level, implicated” (Neimanis 2017: 38). These flows of biomagnified toxins in breast milk also remind us that: “Bodies are both nature and culture, both science and soul, both matter and meaning” (Neimanis 2017: 33).

By employing water as a figuration, Neimanis moves away from a language that defines “things” as given and fixed. Drawing on the works of Gilles Deleuze and Félix Guattari, as well as their feminist new materialist interpretations—including those of Haraway—she operates with the concept of *becoming*, emphasizing the dynamic nature of such an ontology and highlighting the potential for change through the ongoing transformation of material assemblages. As she writes: “[T]o become milieu as womb, we know, is not just a question of biological proclivity; it is also bound up in questions of which bodies get to birth children, and under what circumstances. (...) As

bodies of water, our capacity to proliferate the not-yet is also always bound up in our capacity to bear toxic messages. Gestation is always non-innocent” (Neimanis 2017: 103).

5. Figurations

One might wonder what exactly these figurations are. The term resonates with the activity described by Gilles Deleuze and Claire Parnet: “Thinking in things, among things – this is producing a rhizome and not a root, producing the line and not the point” (Deleuze and Parnet 1977: 26) and what Rogowska-Stangret points out: “[T]he ontology—the ontological story we continue to tell—is already at once a political action when we acknowledge the existence of such and not other categories, such and not other relations, such and not other hopes for the future” (Rogowska-Stangret 2021: 96).

Explaining the figurative approach of her work, Neimanis writes: “[F]igurations are keys for imagining and living otherwise, but unlike a concept unfettered by the world we actually live in or as, figurations are importantly grounded in our material reality” (Neimanis 2017: 5). Figurativity is based on principles also proposed by Barad and other scholars working within FNMs. These: “imaginaries, concepts, discourses, and figures all function as invitations to imagine otherwise. (...) all of these notions are material-discursive phenomena, tethered to materialities they might purportedly only describe” (Neimanis 2017: 184).

Neimanis’s figurations align with the approach articulated by Françoise Vergès: “The most enlightening and productive analyses in recent decades have been those that have drawn the greatest number of threads together to highlight the concrete and subjective networks of oppression that weave the web of exploitation and discrimination” (Vergès 2021: 22). Vergès provides an example from her own work, in which she uses the figure of the banana to map colonialism, agricultural practices, and the global supply system (Vergès 2021: 21). Within Feminist New Materialisms, aside from Neimanis’s water figuration, one can also mention Tsing’s mushrooms (Tsing 2015) and Haraway’s cyborg and humus (Haraway 1991; Haraway 2016).

6. Conclusions

In summary: “[B]odies of water, as figuration, invite us to amplify a relational aqueous embodiment that we already incorporate, and trans-corporate. Bodies of water ask us to imagine these corporeal waters as part of a hydrocommons that we make, and that makes us in turn” (Neimanis 2017: 169). I would also like to recall the words of Alexis Pauline Gumbs from her book *Undrowned: Black Feminist Lessons from Marine Mammals* (Gumbs 2020), which resonate with Neimanis’s work: “Marine mammals live in a volatile substance whose temperature is changing for reasons not of their own making. Their skin is always exposed, they are surrounded on all sides by depth. What could enable us to live more porously, more mindful of the infinite changeability of our context, more open to each other and to our own needs?” (Gumbs 2020: 28).

Ultimately, the subversive potential of Neimanis’s work is expressed in her words: “I’m not interested in narratives that put The Family back in the place, but kinship in all of its queer and wondrous forms is still something I seek. This is the difficulty, of course: embracing queer stories of evolution as a ground of connection and responsibility without reinstating old myths of Dark Continents, Woman-as-Womb, and power plays of family ties. But no story is innocent, and this is the kind of trouble I’d like to - that we need to - stay with” (Neimanis 2017: 151).

References

Barad, Karen (2007) *Meeting the Universe Halfway: Quantum Physics and the Entanglement of Matter and Meaning*, Durham and London: Duke University Press.

Borowski Mateusz (2023) “Splątania w imię przetrwania. Spekulatywne projekty wspólnot więcej-niż-ludzkich w obliczu kryzysu klimatycznego”, in: Sugiera, Małgorzata (ed.) *Hakowanie antropocenu*, Kraków: Wydawnictwo Uniwersytetu Jagiellońskiego, 93-128.

Davis, Noela (2017) “Material Culture: Epigenetics and the Molecularisation of the Social”, in: Kirby, Vicky (ed.) *What if Culture was Nature all Along*, Edinburgh: Edinburgh University Press, 110-133.

Deleuze, Gilles; Guattari, Felix (1987) *A Thousand Plateaus*, Minneapolis: University of Minnesota Press.

Deleuze, Gilles; Parnet, Claire (1987) *Dialogues*, New York: Columbia University Press.

Gumbs, Alexis Undrowned (2020) *Black feminist lessons from marine mammals*, Edinburgh: AK Press.

Haraway, Donna (1991) *Simians, Cyborgs, and Women: The Reinvention of Nature*, New York: Routledge.

Haraway, Donna (2016) *Staying with the Trouble: Making Kin in the Chthulucene*, Durham and London: Duke University Press.

Linton, Jamie (2010) *What is Water?*, Vancouver and Toronto: UBC Press.

Marder, Michael (2013) *Plant-Thinking. A Philosophy of Vegetal Life*, New York: Columbia University Press.

Marzec, Andrzej (2021) *Antropocień: filozofia i estetyka po końcu świata*, Warszawa: Wydawnictwo Naukowe PWN S.A.

Merleau-Ponty, Maurice (2001) *Fenomenologia percepcji*, Warszawa: Fundacja Aletheia.

Neimanis Astrida (2017), *Bodies of water. Posthuman feminist phenomenology*, London and New York: Bloomsbury Publishing Plc.

Neimanis Astrida (2021) „An Archive of an Epoch that Almost Was” in: Hamilton, Jennifer; Reid, Susan; van Gelder Pia; Neimanis, Astrida (eds.) *Feminist, Queer, Anticolonial Propositions for Hacking the Anthropocene: Archive*, Open Humanities Press.

Neimanis Astrida (2023) *Ciała wodne. Posthumanistyczna fenomenologia feministyczna*, Warszawa: Pamoja Press.

Rogowska-Stangret, Monika (2021) *Być-ze-świata*, Gdańsk: wydawnictwo słowo/obraz terytoria.

Rogowska-Stangret, Monika (2023) „Myślenie (z) gąszczem. O relacyjnych narracjach w czasach antropocenu na przykładzie epigenetyki. Perspektywa feministycznego nowego materializmu”, *Postscriptum Polonistyczne* 2 (32), 67-87.

Tsing, Anna (2015) *Mushroom at the End of the World: On the Possibility of Life in Capitalist Ruins*, Princeton: Princeton University Press.

Vergès, Françoise (2021) *A Decolonial Feminism*, London: Pluto Press.

