9th National Meeting in Analytic Philosophy

9

10-12 July 2024

School of Social Sciences and Humanities NOVA University Lisbon (NOVA FCSH)

KEYNOTE SPEAKERS

Teresa Marques
University of Barcelona

Graham Priest
CUNY / Disputatio Lecturer

Ernest Sosa
Rutgers University

ORGANIZING COMMITTEE

Nuno Venturinha
IFILNOVA
NOVA University Lisbon
President of SPFA

Pedro Abreu
IFILNOVA
NOVA University Lisbon

Sofia Miguens
IF, University of Porto

Ricardo Santos
CFUL, University of Lisbon

Célia Teixeira
CFUL, Federal University
of Rio de Janeiro

enfa

Abstracts

Keynote Speakers

Teresa Marques

University of Barcelona

The Responsibility of Individuals

Should we displace the moral responsibility from the individual to the social in accounts of oppression, discrimination, and injustice? Here, I argue that individuals and their attitudes should have explanatory priority in the assessment of moral responsibility. This rests on evidence from recent research in the social sciences aiming at explaining cross-cultural commonalities that can't be properly predicted nor explained by purely structural explanations. I argue that individual attitudes and mental states are necessary for the explanation of social phenomena, and that structural explanations are not sufficient for such explanations, since they cannot, by themselves, explain the common features of social events across socio-historical contexts. Crucially, only individuals can be morally responsible for their intentions to participate in the kind of phenomena considered here: from ideological choices to dangerous speech and mass violence.

Graham Priest
The City University of New York

How Not to See Pierre Making Sense of Absences

There is good reason to suppose that there are absences. For example, it would appear that they can be seen. But the supposition that there are absences faces both metaphysical and epistemological problems. This talk will deploy the tools of formal mereology to show how these problems can be solved.

Ernest Sosa
Rutgers University

Competence, Control, and Credit

What They Are and How It Matters for Epistemology and Ethics.

Abstracts
Parallel Sessions

Rory Aird
University of Glasgow

A Virtue-Theoretic Approach to the Epistemic Normativity of Objecting

We object to others' asserted falsehoods all the time. We frequently have our own asserted falsehoods objected to as well. Nevertheless, it is fair to say that despite the apparent ubiquity of the actual practice of objecting to falsehoods, there is a distinct dearth of discussion of the phenomenon in the (social) epistemology literature, with the small amount of extant scholarship focusing primarily on the *prescriptive* – viz., if an epistemic duty to object is plausible, when do agents have such a duty, what does such a duty look like, and so on. The spotlight being placed so, however, has elided a key area of research which, to my mind, comes before any prescriptive discussion – namely, *evaluative* norms about what makes for a good objection. After all, knowing *when* one epistemically ought to object seems to lack utility if one doesn't also have an idea of how to do it well. Thus, the three main contributions in this paper are as follows: first, I develop evaluative norms of objecting based on a Sosan telic virtue epistemology; from there, however, I make a turn to the prescriptive, and generate prescriptive norms of objecting that work synchronously with the evaluative; finally, I advocate for a prescriptive epistemic duty to *quiesce* (i.e., *not object*) that works symmetrically with the objecting norm.

I begin by discussing a couple of extant definitions of objecting from Lackey (2020; 2021), and critique them through some standard objecting cases, culminating in a more developed definition of objecting which views it as a performative response that constitutively aims at getting an agent to no longer hold their asserted false belief. From there I turn to a simple evaluative norm where a good objection is just one that is successful and show through three different cases how it is clearly inadequate. It is so because, as the cases illustrate, there appear to be more dimensions to what makes an objection good or not than mere success.

Accordingly, the first main contribution of this paper is made: drawing on a performance-normative, virtue-theoretic framework (Sosa 2007; 2021), I generate evaluative norms of objecting based on **Success**, **Competence**, and **Aptness**. These norms are developed in detail, particularly the **Competence** norm with its **Seat**, **Shape**, and **Situation** structure, wherein I carefully delineate precisely what an objecting competence looks like. I then show that these norms give correct diagnoses of each of the so-called problem cases from earlier, resolving any worries previously flagged by them. Such evaluative norms on objecting are already noteworthy insofar as this is hitherto uncharted territory in the literature but, from here, a further step is taken from the evaluative to the *prescriptive*.

I draw on Sosa's recent work on *suspending* and *forbearance*, and show that we can use this work to generate *prescriptive* norms of objecting that work in tandem with the evaluative. Central to this idea is Sosa's contention about the proper aim of attempts: "*to make an attempt on that target if and only if the attempt would succeed aptly.*" (2021: 66) In gnoseology, this gives the *overarching* epistemic aim of "answering one's question aptly, with an apt alethic affirmation" along with the *subordinate* aim of "affirming alethically iff that affirming would be apt (otherwise suspend)" (67) The analogue aims pertaining to *objecting* (call this area *sermology*) could then look like:

> *Overarching aim*: getting a target to jettison a false belief with an apt objection.

> *Subordinate aim*: object to the target assertion iff that objecting would be apt (otherwise quiesce).

The foregoing essentially boils down to the (intuitive) idea that if one's objection would be apt, then one (epistemically) ought to object. A surprising upshot of this framework, however, is that if one's objection *wouldn't* be apt, then one (epistemically) *ought not* to object. An epistemic duty to *not object* may naturally give pause, so the remainder of the paper and the third and final contribution made is the development and motivation of a prescriptive epistemic duty to *quiesce*.

I argue that once we grant that inapt objections can actually be wildly epistemically deleterious, an epistemic duty to quiesce is a feature and not a bug of the account. I show that by objecting inaptly, one takes themselves further away from the overarching and subordinate aims earlier outlined, resulting in a norm that directs one to quiesce being justified and intelligible. I further demonstrate that this will directly track the **Competence** norm discussed in detail earlier, meaning the action guiding virtues of the dual duties of objecting or quiescence are satisfactory.

I close by outlining some benefits this prescriptive account has that no competitors can offer – namely explanatory power and allowing for the possibility of pernicious objections. I finally discuss how this framework fits in the new areas of *sermology* and the *ethics of social discourse*, my analogues of Sosa's *gnoseology* and *intellectual ethics*. Overall, I develop a comprehensive picture of the entire epistemic normativity of objecting – from the evaluative to the prescriptive.

References

Lackey, Jennifer. 2020. 'The Duty to Object'. *Philosophy and Phenomenological Research* 101 (1): 35–60. https://doi.org/10.1111/phpr.12563.
———. 2021. 'When Should We Disagree about Politics?' *Political Epistemology*, 280.
Sosa, Ernest. 2007. *A Virtue Epistemology: Apt Belief and Reflective Knowledge, Volume I*. Vol. 1. OUP Oxford.
———. 2021. *Epistemic Explanations: A Theory of Telic Normativity, and What It Explains*. Oxford University Press.

Gloria Andrada
NOVA University Lisbon

with

Richard Menary
Macquarie University

Cognitive Injustice and the Extended Mind

In this paper, we contribute to the study of disharmonious and non-cooperative human-technology interactions by examining what we have called Cognitive Injustice from an Extended Mind perspective.

Cognitive injustice is a type of injustice caused when the enculturation of cognition leads to detrimental effects on the cognitive abilities of agents, and that negatively affects the capacity of that agent to live a flourishing life, while at the same time contributes to perpetuating social injustice.

In our view, cognitive injustice can take (at least) three forms:

1. *Attenuation*: Enculturation is detrimentally affected by an unjust cultural environment where learning is not possible, such that cognitive abilities that are key for an agent's cognitive flourishing are not acquired.

2. *Warping*: Agents acquire abilities that are detrimental to their cognitive flourishing.

3. *Absence*: Agents lack the opportunity to fully exercise their cognitive abilities, and consequently, they are prevented from living flourishing cognitive lives.

On the other hand, in the extended mind literature, there are two main ways in which cognitive extension can be understood. The first way is Artefact Extension (AE): an artefact gets integrated into the cognitive system of an agent through the right kind of causal coupling (Clark and Chalmers 1998; Clark 2008, 2010). This puts artefacts on a functional par with "in-the-head" cognitive processes. The second way understands cognitive extension as the transformation of human cognitive abilities by a cognitive species of cultural practices (cognitive practices), some of which include artefacts of different sorts (Menary 2007, 2010, 2018). Cognition is extended because, on this account, cognitive practices are themselves part of the cycle of cognitive processing as they are key to maintaining and updating the information stored in their artefacts and their use, and not merely

causally supportive of in-the-head processes. We call Extended Mind-Enculturated style (EnC).

Since the publication of the classic article by Clark and Chalmers (1998), the literature on the extended mind has been quite fertile. However, it has not been until recently that extended cognition advocates have been accused of falling prey to what has been called the "dogma of harmony", namely the fact that they have painted an overly idealized picture of human–technology relations in which all entities are presumed to cooperate and collaborate (Aagard 2022). To avoid this picture, there has been an increasing attempt to examine non-harmonious and non-cooperative interactions between humans and technologies.

We believe that extended mind theorists should in fact depart from such an idealized picture, and examine cases in which cognitive extensions are disruptive or do not improve the agent's cognitive capacities. However, care should be taken not to pursue these examinations from an AE perspective, as it will only give us a partial and picture. This is mainly because AE is not concerned with the actual practices by which human agents manipulate artefacts, but focuses instead on the fact that such artefacts (e.g., notebooks and smartphones) are part of human cognitive systems via causal coupling. If there is disharmony, the idea is therefore that it is caused by specific artefactual operations or designs, or by the agent's causal contribution.

In our view this is just part of the picture, and focusing simply on artefacts and the causal connection between agent and technologies will give us an insufficient understanding of what happens when cognitive extension goes wrong, and does not improve an agent's cognitive standing.

On the contrary, EnC is fully equipped to provide a multidimensional analysis of the different ways in which interactions with technologies can be non-cooperative and disharmonious. This is mostly because EnC focuses not only on the causal connection between human agent and artefact, but also on the cultural practices that govern such interactions, and on the cognitive transformations that such interactions elicit in the agents.

Focusing on the phenomenon of cognitive injustice, our plan is threefold.

First, we focus on *how* cognitive injustice happens. In particular, we examine how and what features of the cultural environment causally influence cognitive development. We assess how cognitive flourishing can be impaired either by attenuating abilities, warping cognitive development or precluding exercise of cognitive abilities. What extended mind EnC-Style brings to the picture is that given that cognitive abilities are constituted by neural and other embodied features, cognitive tools and cognitive practices, cognitive injustice can happen in virtue of affecting any one (or a combination) of them.

Second, we illustrate what cognitive injustice is by looking into cases in which attention is attenuated, warped or cannot be exercised due to features of the cultural environment.

Finally, we spell out some ideas concerning how cognitive injustice can be resisted. Our analysis of the phenomenon of cognitive injustice under an extended mind EnC-style understanding of human cognition provides us with a helpful framework from which to understand the kind of cognitive agents we become in real-life scenarios, i.e., in scenarios of injustice. An extended mind EnC-style approach also helps us see that overcoming cognitive injustice will be a collective enterprise as it requires changes in social practices, and in the material world we inhabit and build.

Consequently, this paper not only expands the literature on extended mind by providing a framework to study disharmonious human-technology interactions and their effect on cognitive development, but it also widens our understanding of the ways in which unfair social systems can harm or promote human cognition.

References

Andrada, G. And Menary, R. (ms) Cognitive Injustice.

Aagaard, J. (2021). 4E cognition and the dogma of harmony. *Philosophical Psychology* 34 (2): 165-181.

Alfano, M. and Skorburg, J. A. (2018). Implications for virtue epistemology from psychological science: Intelligence as an interactionist virtue. In Heather Battaly (ed.), *Handbook of Virtue Epistemology*. New York, NY, USA: Routledge: 433-445.

Andrada, G. (2020). Cognition as an Enculturated and Extended Social Skill. *Tandf: Australasian Philosophical Review* 3 (1):71-75.

Andrada, G. and Menary, R. (submitted). Cognitive Injustice.

Clark, A. (2011). *Supersizing the mind: Embodiment, action, and cognitive extension*. Oxford University Press.

Clark, A., & Chalmers, D. (1998). The extended mind. *Analysis*, *58*(1), 7–19. https://doi.org/ 10.1093/analys/58.1.7

Menary, R. (2007). Cognitive Integration: Mind and Cognition Unbounded. New York. Palgrave Macmillian

Menary, R. (2012). Cognitive practices and cognitive character. *Philosophical Exploratlons* 15 (2):147 - 164.

Menary, R. (2018). Cognitive integration: How culture transforms us and extends our cognitive capabilities. In A. Newen, L. de Bruin, & S. Gallagher (Eds.), *Oxford handbook of 4E Cognition*. Oxford, England: Oxford University Press: 187–215.

Álvaro Domínguez Armas

NOVA University Lisbon

Protests as Group Speech Acts

In this paper, I argue that protests can be described in speech act terms. Protests are extra-institutional political actions going from complex actions like the organisation of rallies or civil disobedience, to simple acts such as the carrying of protest signs with slogans written on them. In this paper I focus on slogans written in protest signs, as they are means by which protesters demonstrate and clarify their ideas in the public sphere when protesting (Van De Velde, 2022).

The study of protests has been carried out mostly within the field of political theory. These studies have focused on what protesters are doing 'by protesting' (Jasper, 2014; Zheng, 2022). However, there is a question that lurks within this literature: what are protesters doing 'in protesting'? In this paper I propose to embrace speech act theory as a framework for answering this question, as it studies the communicative context of speakers in order to understand what they do with their words (Austin, 1962).

Speech acts have three aspects. First, the 'locutionary aspect' in which the propositional content is predicated. Secondly, the 'illocutionary aspect' which refers to the socially recognisable act that speakers perform with their words. Thirdly, the 'perlocutionary aspect' that refers to the consequential effects that the act may have when performed. This paper aims to disclose the illocutionary aspect of protest. As a case in point, I analyse the following two slogans written on protest signs carried by feminist protesters in the demonstration of the 8th March 2020 in Lisbon:

(1) 'Our voice matters'

(2) 'Together we can decide. We demand an end to violence against women'

I examine two proposals to describe what illocutionary act constitutes (1) and (2). The first proposal is the characterisation of (1) and (2) as expressions of opposition and petitions for change (Searle, 1976). In this view, protesters would be expressing disapproval of contexts in which women's voices do not matter and calling the target of protests to take action to make women's voices matter (in (1)) and expressing disapproval to the fact that women are victims of systemic violence and calling the target to take action to eliminate violence toward women (in (2)). I argue, however, that the characterisation of protests as constituted by petitions for change is inappropriate as it undermines the directive force of protests. Petitions have a similar force to that of a request, where speakers try to get the listener to do

something as a favour (Lance & Kukla, 2013), in contrast with other directives acts (e.g., orders, commands, or demands) where speakers try to impose an obligation on the listener to do something.

The second proposal that I will evaluate is the characterisation of (1) and (2) as constituted by expressions of opposition and prescriptions for change (Chrisman & Hubbs, 2021). Chrisman and Hubbs (2021) argue that the expression of disapproval and the petition for change are one among the multiple ways in which a protest can convey opposition to the object and prescribe for change to the target. Importantly, prescriptions are acts that aim at

getting something done that is the prerogative of the hearer. In the case of protests, for instance, social actors in positions of power with the authority to act on what is protested against. I agree with their perspective, but I observe that two aspects must be amended.

The first aspect is that the relationship between expressive and prescriptive acts in protests is not clear. I will survey two possibilities: (a) to consider that there is a hierarchical relationship between both acts (i.e., the prescription is performed *by means of* the expression of opposition) (Searle, 1979) and (b) to consider that both acts are performed at the same time with one and the same locution (Lewiński, 2021). I argue that (b) is more appropriate for protests.

The second aspect is that their characterisation overlooks the collectiveness of protesters. Protesters, when protesting, represent a large group—namely, a social movement. This is perceivable in slogans like (1) and (2) where the plural pronoun 'We' plays a central role. Consequently, I propose to introduce the notion of 'group speech act' to refine the characterisation of protests in speech act terms. Group speech acts are coordinated efforts of more than one individual to do something via the words of a single speaker, such as when a political party returns a verdict to parliament via a designated spokesperson. While it is a single spokesperson who utters the words expressing a verdict in front of parliament, this utterance is understood to count as the verdict issued by the party as a whole. Accordingly, I argue that protests are actions that protesters perform on behalf of the social movement. In the last part of the paper, I argue that characterising protests as group speech acts is advantageous because it allows us to explain other social phenomena, such as what happens when social movements disallow protests and what happens when social movements split into antagonistic groups.

References

Austin, J. L. (1962). *How to Do Things with Words*. Oxford University Press.

Chrisman, M., & Hubbs, G. (2021). Protest and speech act theory. In R. K. Sterken & J. Khoo (Eds.), *The Routledge Handbook of Social and Political Philosophy of Language* (pp. 179– 192). Routledge.

Jasper, J. M. (2014). *Protest: A Cultural Introduction to Social Movements*. Polity.

Lance, M., & Kukla, R. (2013). Leave the gun; take the cannoli! the pragmatic topography of second-person calls. *Ethics*, *123*(3), 456–478.

Lewiński, M. (2021). Illocutionary pluralism. *Synthese*, *199*(1), 6687–6714.

Searle, J. R. (1976). A taxonomy of illocutionary acts. In K. Günderson (Ed.), *Language, mind, and knowledge. Minnesota studies in the philosophy of science* (pp. 344–369). University of Minnesota Press.

Searle, J. R. (1979). *Expression and Meaning: Studies in the Theory of Speech Acts*. Cambridge University Press.

Van De Velde, C. (2022). The power of slogans: using protest writings in social movement research. *Social Movement Studies*. https://doi.org/10.1080/14742837.2022.2084065

Zheng, R. (2022). Theorizing social change. *Philosophy Compass*, *17*(4), 1–14.

Rafael Augusto
University of Edinburgh

Personal-level Control: conscious states as voices, conscious thought as a conversation

There is a persistent and problematic gap in our understanding of cognition: on the one side, we know that cognition is fundamentally nonconscious and automatic (e.g., Strawson, 2003; Lieberman, 2007; Wegner, 2002; Sklar et al., 2021; Nisbett & Wilson, 1977; Bargh & Chartland, 1999). On the other, "It is undeniable that deliberation contains many elements that are straightforwardly intentionally controlled" (Vierkant, 2022: 68. See also Wu, 2013; Jennings, 2022; Brent, 2023) without which – and an explanation of which – the notion of agency is undermined (e.g., Wu, 2013; Schroeter, 2004). This creates a dissonance between the observation that most of our cognitive processing is automatic and the requirement that we explain how intentional, controlled cognitive processing is possible – and how it might emerge from its (automatic) foundations. As Wu (2023) puts it, "the difficult challenge facing cognitive science is how to bridge these 'lower' levels of analyses with the subject level we care about in deciding how to live" (Wu, 2023: 3).

Most solutions to this problem are themselves problematic. The most common one consists in assigning the abilities to coordinate voluntary control of behaviour and thought to one (or several) top-down systems or processes – usually, Working Memory or systems that use Working Memory (e.g., in Executive Function models and Dual-Process Theories. See, for instance, Baddeley & Hitch, 1974; Braver et al., 2007 or Carruthers, 2015). But these subpersonal level accounts, by attributing control to a subpersonal process or mechanism, either fall into the homuncular fallacy or fail to explain how a subpersonal mechanism can be *consciously* controlled in a meaningful way. Even distributed approaches rely on mechanisms that are present in perceptive, intuitive and nonreflective cases, thus failing to isolate the kind of control we aim at. The conceptual mystery persists, and so does the gap, independently of the mechanisms or subpersonal explanations we place "behind" control, because there is no present way of cashing out "control" in a subpersonal sense. And whatever this system is, despite the fact of being supposed to inherit the powers of the subject, and thus *exert* the control, it can still be – and is – described as *under* the control of the subject. Personal level accounts, on the other hand, by relying on a central, intentional Self as causally relevant, fail to explain what such Self is and how it makes its decisions.

I offer an account of personal level control that tackles the conceptual and the pragmatic problems. The former is solved by filling the gap between automaticity and control. The latter is solved by explaining how Executive Control (unique to high order cognition) is distinguished from Executive Function (present in nonreflective action and perception). Starting with the conceptual problem, I build a picture according to which control is relative to a level. Without relying on any controlling entity – at the personal or at the subpersonal level, I start with the claim that, while automatic processes happen at the subpersonal level, they deliver outputs at the personal level – as conscious states or events. If we generalize this principle, which is taken for granted in automatic processing (perception, intuition), we have that deliberation consists of a loop where those outputs work as inputs to the same systems. At the personal level, this loop creates a "virtual" hub in which a personal level "process" emerges. This process, although its contents are automatically generated, is relatively controll*ed* – the sequence of states, *from the personal level perspective*, can be *stopped* or *interrupted* – and controll*ing* – each state, being a state of the person, not the personal level system, controls the generation of the next state by working as an input. Metaphorically, if we take the suprapersonal (social) level into account, and imagine a conversation between people, each voice is a suprapersonal level state – that only exists as such –, and, from the perspective such level (e.g., a room), the emerging conversation is controlled while each person is an autonomous black box which receives inputs and delivers outputs. We are the room.

Regarding Executive Control, the proposed architecture explains it by defining it in the following terms: an organism has the ability for Executive Control if it can react to a stimulus by reacting to an internally generated state directly or indirectly caused by the stimulus, but which does not constitute a direct perception of the environment. In other words, it is the ability to react to the external environment by reacting to an internally simulated re-presentation. According to this definition, although a trained dog has inhibitory powers, it does not *necessarily* have this ability because both stimuli (the food and the owner telling him to wait) are environmental ones. But, if Executive Control is in place, the owners' voice saying "wait" can be mental – that is, internally generated or simulated, absent in the immediate environment. Importantly, this only includes cases where the representation causes the desire (I inhibit the desire to smoke because of a thought), and not cases where the desire causes the representation (I want to smoke, so I start imagining ways of finding a cigarette). So, instead of having certain *additional* brain mechanisms which explain the ability *not to smoke* a cigarette, what happens is that a conflict, which could happen between two competing actions triggered by a stimulus, might happen between actual stimulus and internally generated one. In this personal level, "controlled" virtual hub, there is thus a "conversation", an exchange of outputs-inputs that, at some point, end up in a conscious decision.

Being able to explain the available data, this model explains controlled cognition and how it might emerge from its automatic, subpersonal foundations, without relying on questionable centres, selves or psychological constructs. And it does so in a way that aligns with our knowledge of the suprapersonal level: in a social environment,

relative control emerges from the same principles – and states that depend on their *shared* nature *only* exist at their level – like conscious states.

Matteo Baggio

IUSS Pavia/University of Turin

The logical epistemology of logical anti-exceptionalism

Logical anti-exceptionalism (AEL) is the mainstream view in the epistemology of logic (cf. Maddy 2002; Russell 2015; Priest 2016; Hjortland 2017; Williamson 2017; Martin & Hjortland 2021). In general, AEL is an empiricist proposal that aims at explaining knowledge (warrant) of *logical theories* (cf. Ferrari, Martin, Sforza Fogliani 2023). AEL has been conceived in two different ways. That is, (i) as the idea that logical inquiry is somehow continuous with scientific inquiry (cf. da Costa & Arenhart 2018); or (ii) as the idea that the traditional features attributed to logical claims do not hold, e.g., *apriority* and *foundationality* (cf. Williamson 2017).

Although AEL can be articulated in these two ways, (i) represents too vague a thesis (cf. Rossberg & Shapiro 2021). Indeed, there are several ways to spell out this continuity link. For example, limiting ourselves to considering only how theories get selected, the relationship of continuity binding logic and science could be specified both in terms of evidence and methodology. After quickly presenting the shortcomings that follow from the vagueness of (i), I examine (ii).

In presenting (ii), I follow the current literature distinguishing between Evidential AEL and Methodological AEL. EAEL is the thesis that the evidence for logic is the same as for other sciences, i.e., empirical evidence. MAEL is the thesis that the methodology of logic is the same as for other sciences, i.e., non-deductive and inferential. This difference allows MAEL to admit *a priori* evidence and, consequently, the possibility of having *a priori* knowledge (warrant) for logic. My goal will be to show that each version of EAEL and MAEL are defective from an epistemic point of view.

There are at least two ways to account for EAEL: (a) to deny the *a priori-a posteriori* distinction, or (b) to say that the analysis of logical concepts does not give privileged access to logical truths. Regarding (a), I show that proponents of this strategy struggle to exclude *a priori* evidence. For example, although Williamson denies the naturalness of the *a priori-a posteriori* distinction, he seems to admit the existence of innate *a priori* knowledge on which much logical knowledge depends (Williamson 2013, 427). Regarding (b), I show that proponents of this strategy wrongly assume that *a priori* and analyticity are coextensional concepts. For example, in the Kantian conception of the *a priori*, analyticity is neither necessary nor sufficient for *a priori* knowledge (cf. Casullo 2003). Thus, this alleged co-extensionality must be further motivated or abandoned.

The main focus of my paper will be on MAEL. According to MAEL proponents, knowledge of logic can be *a priori* (cf. Biggs & Wilson 2022; Martin & Hjortland 2022). MAEL's main proposals are *abductivism* and *predictivism*. While logical abductivism is the thesisthat theory selection is based on abductive inferences

based on criteria such as *simplicity*, *adequacy to the data*, and *consistency*, predictivism equates logical methodology with the *hypothetico-deductive* method.

Abductivism faces two problems. First, theories fed into the abductive process must be minimally warranted for the process to begin (at least, this seems to apply to the principles that form the original core of a theory). After all, we would not compare theories we are not minimally warranted in believing. However, abductivism is silent on how this minimal warrant is granted to our theories. In this sense, rather than explaining the epistemology of logic, abductivism presupposes other methods by which we explain how we know (warrant) logical principles. Second, even assuming that the first problem can be solved, there are technical details that vitiate this proposal. For example, theories with different ontological commitments cannot be compared with abductive inferences (cf. Kramer 2014), and infinite loop cases can vitiate the possibility of selecting a theory (cf. Priest 2016; Woods 2018).

Predictivism arises as a theory that can solve the problems of abductivism. The hypothetical part of the process ensures the minimal warrant of a logical theory. In contrast, the deductive part evaluates its predictions. However, predictivism is entangled with expert epistemology. To construct and evaluate the correctness of a logical theory, the predictivist uses expert judgments (cf. Martin & Hjortland 2021). This fact prevents *a priori* knowledge (warrant) for logic, as it collapses predictivism (MAEL) into EAEL. This fact depends on two constitutive facts of the predictivist strategy. First, the evidence used to evaluate a logical theory is empirical, i.e., evidence consisting of observations about the logical principles used by experts. Second, the evidence gathered to construct the theory is also empirical, i.e., evidence conveyed through testimony or observations of experts' inferential behavior.

These two factors make the evidential basis with which the predictivist investigator knows (warrants) a logical theory dependent on experience in an evidential way. Therefore, the supposed compatibility of predictivism with logical apriorism seems precluded.

AEL represents orthodoxy in the epistemology of logic. I have argued that each version of AEL clashes with well-established epistemological results. I conclude that orthodoxy should be reconsidered.

Biggs, S. & Wilson, J. M. (2022). Does anti-exceptionalism about logic entail that logic is a posteriori? *Synthese,* 200: 209.

Casullo, A. (2003). *A Priori Justification*. Oxford University Press.

da Costa, N., & Becker Arenhart, J., (2018). Full-blooded anti-exceptionalism about logic. *Australasian Journal of Logic*, 15(2):362-380.

Hjortland, O. T. (2017). Anti-exceptionalism about Logic. *Philosophical Studies*, 174:631–658. Krämer, S. (2014). Implicit commitment in theory choice. *Synthese*, 191(10): 2147–2165.

Maddy, P. (2002). A Naturalistic Look at Logic. *Proceedings and Addresses of the American Philosophical Association*, 76(2):61-90.

Martin, B. & Hjortland, O. (2021). Logical Predictivism. *Journal of Philosophical Logic*, 50:285–318. Martin, B. & Hjortland, O. (2022). Anti-Exceptionalism about Logic as Tradition Rejection. *Synthese* 200: 148. Russell, G. (2015). The Justification of the Basic Laws of Logic. *Journal of Philosophical Logic*, 44(6):793-803. Priest, G. (2016). Logical disputes and the a priori. *Logique et Analyse*, 59(236), 347-366.

Williamson, T. (2013). *Modal logic as metaphysics*. Oxford University Press.

Williamson, T. (2017). Semantic Paradoxes and Abductive Methodology. In B. Armour-Garb (Ed.), *Reflections on the Liar* (1ᵗʰed., p. 325-346). Oxford University Press.

Woods, J. (2018). Logical Partisanhood. *Philosophical Studies*, 187, 321-341.

Zsolt Bátori
Kodolányi János University

AI-generated photo-based images: their ontological status and interpretation

In a recent paper in *Science*, Epstein, Hertzmann, et al. announce that "generative AI is not the harbinger of art's demise, but rather is a new medium with its own distinct affordances" (Epstein, Hertzmann, et al., 2023, p. 1110). The authors recall that throughout the history of art, several stages emerged when new technologies seemed to threaten some, if not all, artistic practices. For instance, the invention of photography as a new technology for mechanically recording light values and producing depictions of scenes in front of the camera was perceived by many as bringing about the end of painting and drawing. Especially in Western art and culture, where realistic portrayal of scenes, relying on perspectival visual representation, was important at that time, the automatic and mind-independent nature of photography was easily viewed as "superior" to handmade, mind-dependent images in terms of realistic depiction. However, as Epstein, Hertzmann, et al. also remind us, photography did not by any means replace painting and drawing. Although portraiture did indeed become largely a photographic genre from that time onwards, photography also liberated painting from realism. Many view developments in painting, such as Impressionism, for instance, as the most welcome effect of this liberation.

Building on this analogy, Epstein, Hertzmann, et al. argue that artists and audiences need not perceive generative AI as a threat to current art practices. While it will inevitably bring about changes, artists will reformulate current practices by using generative AI as a new tool in their creative endeavours. Automatisms offered by generative AI should be seen as the rise of a new medium, providing new ways for creative artistic work done by humans. As Epstein, Hertzmann, et al. point out, one potentially misleading aspect of the perception and reception of works produced with generative AI lies in the term we use for it. The term "artificial intelligence" might misleadingly imply human-like intentions, agency, or even consciousness or self-awareness. Removing some misconceptions about generative AI may also foster its acceptance as a new technological tool in the hands of human artists.

The controversy surrounding the ontological status of AI-generated artworks often revolves around the question of authorship. Besides the human beings involved in the process, such as programmers and artists utilising the program, generative AI programs themselves have been suggested as possible candidates for being considered authors (see Elgammal, 2019, and Mazzone and Elgammal, 2019, for instance). However, others argue that, in the absence of consciousness and conscience, AI-generated artworks should be considered as artworks produced by human persons and mediated by a generative AI program, rather than

being artworks produced by the AI program itself. In other words, the ontological status of artworks is derived from the connection between an artwork and consciousness and conscience (see Linson, 2016, for instance). Another sceptical argument about AI authorship is that, according to our current understanding, art is created by social agents. Therefore, until this understanding is changed, generative AI cannot be credited with the authorship of art (see Hertzmann, 2018).

In my paper I argue that the ontological status of AI-generated photo-based images, whether they are artworks or other photo-based images, is better understood in terms of their contextual interpretation rather than in terms of their connection to consciousness, conscience, or social agency. This also means that I do not consider artistic creativity and non-artistic image-making creativity to be fundamentally distinct from the point of view of ascribing authorship of such images to generative AI. (However, they are distinct in terms of interpreting them as artworks or non-artistic images.) For my arguments, I rely on the theory of pictorial illocutionary acts developed by Kjørup (1974, 1978) and Novitz (1975, 1977), as well as on the theory of photographic illocutionary acts proposed by Bátori (2015, 2018).

According to the theory of pictorial illocutionary acts (Kjørup, 1974, 1978, Novitz 1975, 1977), the production and presentation of images themselves are to be understood and interpreted as pictorial locutionary acts, similar to verbal locutionary acts, such as uttering words and sentences. At the locutionary act level, only the literal semantic pictorial meaning of the image is interpreted. This meaning is based on our visual recognition abilities, such as object recognition, face recognition, recognising spatial relations, arrangements, and perspective. Currie (1995) refers to this pictorial semantic content as 'natural' pictorial meaning because it is not learned, unlike the learned symbolic semantic content of words and other morphological meaning units in natural languages. At the level of pictorial locutionary acts, contextual information is not utilised. It is only at the level of pictorial illocutionary acts that we interpret the image in the context of its presentation and use. For instance, at the pictorial locutionary act level we merely recognise the visual characteristics of the picture of a human head in the barber shop window, while at the pictorial illocutionary act level, we interpret it as a possible statement (pictorial proposition) about the skills of the barber or as a promise of getting a similarly skilful haircut in that barber shop. As Bátori (2015, 2018) further elaborates, photographic illocutionary acts constitute a specific type of pictorial illocutionary act in which the interpretation process at the illocutionary act level necessarily includes interpreting the images as indexical photographs, as opposed to non-indexical, hand-made images.

When interpreting photo-based artworks and other photo-based images produced using generative AI, the locutionary and illocutionary acts involve the following components. At the locutionary act level, audiences identify the literal semantic pictorial content of the images, utilising their visual recognition capacities. This process yields pictorial mental representations of the image content for the mental processing of the audiences. At the illocutionary act level, audiences utilise their contextual knowledge that the image they are considering is a generative AI rendering of a photo-based image or images. They also take into account that the rendering was created using a) the algorithms of the programmer and b) the ideas of the person (artist, creative professional, etc.) instructing the generative AI

program. This means that the image as a whole will not be interpreted as an indexical depiction of a scene captured by the camera at the time of exposure, as the interpreter knows that the image has been altered. The role the original indexicality plays in the interpretation depends on the specific modifications and the extent of the interpreter's knowledge about them. However, in terms of their ontological status, AI-generated photo-based images will not be treated and interpreted as indexical photographs. However, it is not clear whether this implies the emergence of a distinct genre in the process, as suggested by Epstein, Hertzmann, et al. Alternatively, there might simply be new technological means of producing composite images.

With regards to authorship, the interpretation at the illocutionary act level attributes authorship to the person (artist, creative professional, etc.) using the program, not to the programmer or the generative AI program. This is because the person utilising the generative AI program is the one who produces and presents the image (locutionary act) with the assistance of the generative AI program as an image manipulation tool. In the production and presentation process of the image, the programmer is attributed a role similar to that of camera and darkroom equipment constructors, or image manipulation software engineers. Meanwhile, the generative AI is regarded as a complex technical tool for rearranging parts of one or more indexical photographic images into a new, non-indexical image as a whole. Attributing authorship to generative AI is no more a part of the illocutionary act than attributing authorship to image manipulation software used to adjust the contrast or saturation of an image or even to rearrange an indexical photograph or photographs into a new composite, non-indexical image.

Furthermore, the differentiation between "traditional" (non-AI-generated) images and AI generated ones draws a parallel to the contrast between handmade and mass-produced items (like shoes, tableware, etc.). In the instance of "traditional" production, the object's creator retains complete control over all the encompassing processes, whereas the designer of a mass produced item only creates the distinctive facets of the product, without direct involvement in each stage of production.

Based on the interpretation process described at the illocutionary act level, it can be concluded that audiences come to have true beliefs about the nature of photo-based images produced using generative AI, as long as the image's nature is readable form it or deducible from the context. Audiences are not deceived in such cases. However, if the image's nature is neither deducible from the context nor readable from the image itself, they might be deceived into interpreting it as an indexical photograph of a scene captured by a camera. During my talk, I will present examples of both deceptive and non-deceptive photo-based images produced using generative AI.

Keywords

generative AI, photo-based images, pictorial and photographic illocutionary acts

References

Bátori, Zsolt. 'Photographic Manipulation and Photographic Deception.' *Aisthesis* 11:2 (2018):35-47. doi: 10.13128/Aisthesis-23863

Bátori, Zsolt. 'Photographic Deception.' *Proceedings of the European Society for Aesthetics* 7 (2015):68-78.

Currie, Gregory. *Image and mind: Film, philosophy and cognitive science.* Cambridge: Cambridge University Press, 1995.

Elgammal, Ahmed. 'AI Is Blurring the Definition of Artist: Advanced algorithms are using machine learning to create art autonomously.' *American Scientist* 107:1 (2019):18-21. doi: 10.1511/2019.107.1.18

Epstein, Ziv, Hertzmann, Aaron, et al. 'Art and the science of generative AI.' Science 380 2023):1110-1111. doi: 10.1126/science.adh4451

Hertzmann, Aaron. 'Can Computers Create Art?'Arts 7:2 (2018):18. doi:10.3390/arts7020018 Kjørup, Søren. 'George Inness and the Battle at Hastings, or Doing Things with Pictures.' *The Monist* 58:2 (1974):216-235.

Kjørup, Søren. 'Pictorial Speech Acts.' *Erkenntnis* 12 (1978):55-71.

Linson, Adam. 'Machine Art or Machine Artists?: Dennett, Danto, and the Expressive Stance.' In V.C. Müller (ed.), *Fundamental Issues of Artificial Intelligence.* Switzerland: Springer International Publishing Switzerland, 2016, pp. 443-458. doi: 10.1007/978-3-319-26485- 1_26

Mazzone, Marian, Elgammal, Ahmed. 'Art, Creativity, and the Potential of Artificial Intelligence.' Arts. 8:1 (2019):26. doi:10.3390/arts8010026

Novitz, David. 'Picturing.' *Journal of Aesthetics and Art Criticism* 34:2 (1975):145-155. Novitz, David. *Pictures and their Use in Communication: A Philosophical Essay,* The Hague: Martinus Nijhoff, 1977.

Javier Belastegui
UPV/EHU

Classical Concepts, Definitions and Analyticity in Conceptual Spaces

The *conceptual spaces* approach, first introduced by Peter Gärdenfors (Gärdenfors, 2000, 2014), aims at representing concepts (e.g. about colours, actions) as geometric regions in a space. The approach has been used to illuminate philosophically interesting phenomena such as natural concepts, vagueness, inductive and analogical inference, and so on. In particular, Gärdenfors conjectured that natural concepts correspond to convex regions of the (Euclidean) space.

Different theories of concepts ascribe different structures to concepts (see (Machery, 2009)). A *prototype concept* is one structured around a prototype, which is a set of those properties that are most typically found among the instances of the concept. An *exemplar concept* is one structured around a set of typical instances. In contrast, a *classical concept* is one with a definition, a set of necessary and sufficient conditions that an object has to satisfy to fall under it.

Usually, the conceptual spaces approach is combined with the basic assumptions of the prototype and exemplar theories of concepts (see (Mormann, 2021), (Belastegui, 2022) for formal approaches). Several reasons may be given for it. On the one hand, during the last decades, psychologists have accumulated evidence against the classical theories of concepts and in favor of the prototype theories (Machery, 2009). So one may contend that, if most of the concepts of daily life have prototype structure, we should stop considering classical concepts. But arguably, some concepts we theorize about, such as concepts from mathematics (e.g. CONTINUOUS FUNCTION), law (e.g. MURDER), genealogy (e.g. UNCLE), and so on, have classical structure. Formal semantic approaches in linguistics also seem to assume that the meanings of the predicates of natural language have classical structure. Moreover, since Socrates, philosophers have thought that the phenomena they were interested in were represented by classical concepts (e.g. KNOWLEDGE, JUSTICE). On the other hand, classical concepts can already be studied using *logic*. So, one may contend that the spatial approach has nothing new to contribute to their study. However, the aim of this talk is precisely to show that this is not the case. Conceptual spaces can also be successfully applied to learn new things about the structure of classical concepts.

The space I choose for this task is the *Cantor space,* which is, by a famous result of the mathematician Marshall Stone (Stone, 1936), the space of classical propositional logic. We start from a countably infinite set Q of basic qualities (e.g. being wise, being round, and so on, think about them as propositional variables). As the points of the space, we take all the subsets of this set, i.e. all the *bundles of*

qualities (basically the classical valuations). Each such a bundle represents the qualities that an object may have in a given context. On this set, which is 2^Q, we define the standard Cantor topology, i.e. a subbasic contains all the bundles that share (lack) a given quality. Each basic set contains all those bundles that are similar with respect to finitely many qualities. The crucial step is new: each classical concept is represented as the region of those bundles that share exactly the qualities defining the concept. I will argue that this spatial model of classical concepts has three philosophically interesting consequences.

First, the regions are shown to satisfy plausible adequacy conditions linking *similarity* and concepts, including Goodman's criterion for Perfect Communities (see (Carnap, 1928), (Goodman, 1951)) and Douven-Gärdenfors's Well-Formedness Design Principle (Douven, Gärdenfors, 2019).

Second, following the conceptual spaces approach (Gärdenfors, 2000), these regions are shown to be *geometrically well-behaved*, for they are topologically closed and geometrically convex. Since prototype concepts are usually vague, regions representing prototype concepts are taken to have boundary instances and thus are represented as open sets (see (Rumfitt, 2015), (Mormann, 2021)). In contrast, most of the regions in this talk are shown to have empty boundaries (they are clopen), as one would expect from classical concepts.

Finally, I show that, since these regions form a Boolean algebra, they satisfy a *compositionality* condition. I use this to consider some contemporary discussions on Kantian-style conceptions of *analyticity* as content containment. For instance, some contemporary approaches to hyperintensionality (e.g. Fine's truthmaker semantics (Fine, 2017)) reject classically valid principles such as the Introduction of Disjunction and the Absorption Laws, due to their being in tension with the conception of content containment. E.g. although "the ball is red" classically entails "the ball is red or the sun is shiny", the content of the latter is not contained in the content for the former, for the former says nothing about the sun. I will show that, once we interpret these principles according to the definitional structure of concepts, they become uncontroversial.

References

(i)   Belastegui, J. (2022): A Qualitative Approach to Conceptual Spaces: Prototypes as Qualitative Atoms. *Erkenntnis*.

(ii)   Carnap, R. (1928): *The Logical Structure of the World*. University of California Press. (iii) Douven, I.; Gärdenfors, P. (2019): What are Natural Concepts? A Design Perspective. *Mind Lang.*, 1–22.

(iv)   Fine, K. (2017): A Theory of Truthmaker Content I: Conjunction, Disjunction and Negation. *Journal of Philosophical Logic*, **46**, 625–674

(v)   Gärdenfors, P. (2000): *Conceptual Spaces: The Geometry of Thought*. MIT Press. (vi) Gärdenfors, P. (2014): *The Geometry of Meaning: Semantics based on Conceptual Spaces*. MIT Press. (vii) Goodman, N. (1951): *The Structure of Appearance*. Harvard University Press. (viii) Machery, E. (2009): *Doing without Concepts*. Oxford University Press.

(ix)   Mormann, T. (2021): Prototypes, Poles and Tessellations. Towards a Topological Theory of Conceptual Spaces. *Synthese*, 1–35.

(x)   Rumfitt, I. (2015): *The Boundary Stones of Thought: Essays in the Philosophy of Logic.* Oxford University Press.

(xi)   Stone, M. H. (1936): The Theory of Representation for Boolean Algebras. *Transactions of the American Mathematical Society*, **40**, 1, 37–111.

Miloud Belkoniene
University of Zurich

Moral Intuitionism: Between Reasons and Inclinations

Although moral intuitionism somewhat fell into disrepute during the second half of the past century, there has been a renewal of interest in this view with philosophers such as Audi (1997, 2009), Huemer (2005) and Roeser (2011) offering new discussions of the challenges that were raised against it. Moral intuitionism, as I shall understand it here, is the view that one can secure direct knowledge of moral truths or direct justification for believing moral propositions. The notion of directness at play here can be cashed out in terms of non-inferentiality and the claim that one can have direct moral knowledge or direct justification for believing moral propositions can therefore be understood as the claim that one can know or justifiably believe moral propositions non-inferentially.

Of course, even when those clarifications have been made, moral intuitionism, as just formulated, leaves quite a lot of room for interpretation. Does moral intuitionism amount to the view that one can know or justifiably believe moral propositions without inferring them from other propositions, or does it amount to the view that moral knowledge and moral justification are achievable even if the target moral propositions cannot be inferred from other propositions? In response to a new challenge raised by Sinnott-Armstrong (2006a, 2006b, 2008) for moral intuitionism, Tropman (2011) tackles the task of providing a precise answer to this question and argues that moral intuitionism is best understood as the claim that one can be justified in believing a moral proposition without believing that proposition on the basis of reasons.

Tropman's preferred construal of moral intuitionism has several advantages. However it raises an interesting question pertaining to the nature of doxastic justification and, more specifically, the justification one can have for believing moral propositions: how can a belief in a moral proposition be justified without being based on the reasons one has for believing that proposition? There is a widespread intuition among philosophers interested in doxastic justification that in order to be doxastically justified in believing that *p* a subject's belief that *p* must, in addition to being supported by the reasons she possesses, depend on those reasons.
Yet, moral intuitionism, as construed by Tropman, runs counter to that intuition and, as a result, requires a specific explanation of why and how some moral beliefs can be doxastically justified without being based on reasons.

In the present paper, my aim is to show that given a certain conception of the bases of intuitively justified moral beliefs, such an explanation can be offered. To that end,

I first examine Tropman's preferred construal of moral intuitionism. I then address the question as to

what are the bases, if not reasons, of intuitively justified moral beliefs. After having examined the view that those bases are intellectual appearances understood as the vehicles of a subject's direct awareness of moral facts, I argue that intuitively justified moral beliefs are best understood as resulting from inclinations to believe certain moral propositions. Next, I address the question as to how some moral beliefs can be justified while resulting from doxastic inclinations. I argue that, in certain circumstances, a subject's inclination to believe a moral proposition is intelligible to her in light of the fact that certain considerations support believing the object of her inclination and that this suffices for her intuitive belief to count as justified.

References

Audi, R. (1997). *Moral Knowledge and Ethical Character*. Oxford: Oxford University Press.

Audi, R. (2009). *The Good in the Right: A Theory of Intuition and Intrinsic Value*. Princeton University Press.

Huemer, M. (2005). *Ethical Intuitionism*. New York: Palgrave Macmillan.

Roeser, S. (2011). *Moral Emotions and Intuitions*. New York: Palgrave Macmillan. Sinnott-Armstrong, W. (2006a). Moral Skepticisms. New York: Oxford University Press.

Sinnott-Armstrong, W. (2006b). Moral Intuitionism Meets Empirical Psychology. In T. Horgan & M. Timmons (Eds.). *Metaethics after Moore*. Oxford: Clarendon Press, 339–365.

Sinnott-Armstrong, W. (2008). Framing Moral Intuitions. In W. Sinnott-Armstrong (Ed.). *Moral Psychology: The Cognitive Science of Morality: Intuition and Diversity*. Cambridge: The MIT Press, 47–76.

Tropman, E. (2011). Non-Inferential Moral Knowledge. *Acta Analytica,* 26, 355–366.

Ángela Villamuelas Castro

University of Santiago de Compostela

The Risk of Social Movements in the Neoliberalism Era

Isaiah Berlin makes a distinction between negative freedom and positive freedom: (i) negative freedom is defined as the absence of coercion by external bodies, (ii) while positive freedom refers to the freedom that each person should have to carry out his or her will and to make decisions (Berlin 2002). Negative freedom would be linked to non-interference, while positive freedom has to do with autonomy and the capacity for self-realisation. Thus, according to the negative conception, a person is free if and only if he or she is free from the interference of external agents in carrying out the designs of his or her will. In this sense, poverty or lack of economic resources are not seen as an absence of freedom, but as the impossibility of exercising a freedom that formally belongs to the person. Therefore, poverty would be considered, within the liberal perspective of freedom, as a limitation on the same level as other intrinsic limitations such as lack of wit, lack of intelligence or illness (Nozick 1974; Friedman 1989).

In contrast, alternative conceptions of freedom, such as effective freedom, have been put forward from non-liberal positions (Cohen, 2000). Those who defend the position of freedom understood as effective freedom will argue that it is trivial to make a distinction between negative freedom (the non-interference of external agents to fulfill the designs of the will) and the possession of material resources that make it possible to carry out what is desired. In this way, I will argue that there is no formal or normative difference between the interference of external bodies and the absence of material goods, such as money, to do what is desired. For example, not being able to afford a good or service will, for all practical purposes, imply the interference of external agents: if a person decides to take a packet of lentils from a supermarket without paying for them because they do not have enough money, and is discovered, the supermarket's security personnel will become an external agent that limits the supposed freedom of action of this person, who at all times would have the capacity to take the packet of lentils if it were not for the security personnel, so it would be an extrinsic limitation and not an intrinsic one as when we referred to the lack of intelligence.

This argument would already imply admitting that the non-interference of external agents is a priority in order to guarantee people's freedom over the fact of having the conditions of existence that transform a certain action into a real option. Likewise, when the alternative to leaving a disadvantageous situation is going to cause greater harm to the person, from the point of view of effective freedom, it will be understood that there is no real freechoice.

Neoliberalism, on the other hand, has as its main objective the halting of interventionist policies and the end of the welfare state (Hayek, 2020). The neoliberal economic model is characterized, among other things, by an absolute primacy of negative freedom, the understanding of the individual as an autonomous ethical agent who is governed by his or her personal interests (Rand, 2019).

Likewise, in relation to the ideological sphere, neoliberalism has managed to establish in the collective imaginary the idea that there is no viable alternative to its principles, so that both supporters and opponents have to adapt to its rules (Anderson, 1999; Centeno and Cohen, 2012). From this premise, I will argue that, if neoliberalism has achieved a great impact at the political level, inevitably the social movements that have developed within it must also have been affected by the great political charge they present. The neoliberal model tries to apply the idea of individual freedom that emanates from the logic of the free market to the different spheres of society. Consequently, certain branches of the different social movements end up adopting a neoliberal stance, which means that their analysis of their own oppression is biased by the interests of the system itself and they only present superficial demands without ever reaching a root solution. In other words, the system itself infiltrates the groups that initially fight against the oppression that it itself perpetuates, only to end up offering certain alternatives that ultimately do not provide any real solution to the problem (Gradin, 2017; Turner, 2011).

In this way, by creating an increasingly individualistic culture due to the influence of its ideology, neoliberalism is achieving a certain destructuring of social movements. All of this ends up leading to a sectorialisation of the social and political demands made by the different oppressed collectives. Likewise, the idea of freedom that is developed within the neoliberal rationale implies a certain rejection of the social which, ultimately, results in the denial of what social movements identify as privileges (gender, race, class), given that they deny that this position of power is the result of a social product.

By denying the social origin of the construction of inequalities and oppressions, neoliberalism manages to eclipse social awareness and the consequences for vulnerable groups of the existence and construction of these systems of power (Brown, 2019). The denial of the social component on which social hierarchies and power relations are built can also entail the denial of the existence of oppression itself. Thus, if there is no oppression, there is no reason for struggle or demands: social movements end up lacking a purpose and are reduced to mere absurdity. This is the main danger that the logic of the free market will spread to the different spheres of society, affecting both public and private life.

Consequently, I will argue that this progressive neoliberal colonization is a risk factor for the future of the social movements, because neoliberal reason has succeeded in legitimising inequility. If all demands start to be framed within an individual perspective that complies with the discourse of negative freedom, there will be no real and tangible change in social injustices.

Bibliography

Anderson, P. (1999). Neoliberalismo: un balance provisorio. En P. Gentili, & E. Sader, *La trama del neoliberalismo: mercado, crisis y exclusión social* (págs. 1-18). Buenos Aires: EUDEBA.

Berlin, I. (2002). *Liberty: Incorporating Four Essays on Liberty.* Oxford: Oxford University Press.

Brown, W. (2019). *En las ruinas del neoliberalismo.* Madrid: Traficantes de sueños.

Centeno, M., & Cohen, J. (2012). The Arc of Neoliberalism. *Annual Review of Sociology, 38*, 317-340.

Friedman, D. (1989). *The Machinery of Freedom: Guide to a Radical Capitalism.* Chicago: Open Court Publishing Company.

Gradin, A. (2017). Los movimientos sociales en en neoliberalismo tardío. En A. Gradin, & D. García, *El neoliberalismo tardío: teoría y praxis* (págs. 215-226). Buenos Aires: FLACSO.

Hayek, F. (2020). *Camino de servidumbre.* Madrid : Alianza Editorial.

Nozick, R. (1974). *Anarchy, State and Utopia.* New Jersey: Blackwell.

Rand, A. (2019). *La rebelión de Atlas.* Barcelona: Ediciones Deusto.

Turner, R. (2011). *Neo-liberal Ideology: History, Concepts and Policies.* Edinburgh: Edinburgh University Press.

Robert Clowes
NOVA University Lisbon

Paul Smart
University of Southampton

Exploring The Creative Landscape of Deep Learning Neural Networks and the Challenge it may pose to Human Creativity

This paper explores how our concepts of human cognitive activities may be being reshaped by our interaction with deep-learning neural networks, with a particular emphasis on our notions of creativity and creative authorship. The investigation begins with real-world instances that challenge traditional notions of human creativity. Jason Stanley, a games designer, won an art contest using an AI image production system, while German artist Boris Eldagsen also was awarded a prize in an international photography competition before withdrawing his entry. In 2024 a novelist won a writing competition in China with a story developed with a Large Language Model (LLM). These recent developments showcase the significant impact of deep learning neural networks on digital art, writing, and generative tasks.

These cases, and our everyday interactions with deep learning systems and especially Large Language Models (LLMs) have transformed our understanding of AI's capabilities and our conceptual model of mentalistic concepts such as creativity. But should they also prompt a reassessment of what creativity is and whether AI is capable of it?

On one view the real creative potential of deep learning systems is a mirage. LLMs may appear to do creative work but they are really just stochastic parrots (Bender et al. 2021). But while it is true that LLMs follow statistical distributions in the datasets – such as the huge amounts of digitized language – they are exposed to (Shanahan 2022), does this mean that their output are not really creative? How are we to assess claims that these systems produce genuinely creative outputs?

To analyse this question our paper is structured in three parts.

1 – Internalist Creativity and Virtual Philosophers: examines some attempts to use LLMs in a case of creative authorship close to our hearts, namely the production of philosophical articles by "DigiDennett" a large language model of the philosopher Daniel Dennett (Schwitzgebel, Schwitzgebel, and Strasser 2023). We examine this question further by looking at one of the

most respected philosophical accounts of creativity, namely (Boden 1990). Using Boden's analysis we examine whether DigiDennett meets the criteria for creativity in both the historical and psychological senses Boden describes. We argue that while it is successful in meeting Boden's criteria it also presents a rather inadequate picture of the cognitive basis of creativity. In particular, it presents an unrealistic internalist and snapshot like view of the creative process that is worth challenging.

2 – Against Internalist Accounts of Creativity and Extended Agency Interleaving: So in the second part we look at the question of creativity again from a 4E perspective, noting how much real-world creativity involves circular patterns activity involving the skilled agent and the proximal environment. We do in fact find a similar pattern in the use of deep learning systems. To make this argument we examine Elvis Deane's use of generative AI in comic creation. Drawing on Enactivist Literature, we question the internalist perspective in some analytic thinking about creativity and authorship and introduce the concept of "agency interleaving," to help explain creativity in the context of deep learning systems. We thus argue against overly internalist models of creativity assumed by Schwitzgebel et al is better rethought through our own interleaved agency model and that this gives a better and more realistic account of creative authorship.

3 – The Future of Human Creativity against a Deep Learning Background: The paper concludes by examining future possibilities for creativity within the evolving landscape of deep learning, presenting a nuanced perspective that integrates AI and human creative processes. But it presents a puzzle around the reflective opacity of generative AI (Andrada, Clowes, and Smart 2022). It discusses traceability as a key feature and evaluates the reflective opacity of current systems in academic contexts. It addresses the challenges of writing with AI systems and introduces a new LLM-based model with an extended memory, challenging more traditional internalist forms of accounting for creative activity. We argue traceability may be a central feature of whether "extended creativity" is really possible in many contexts.

Citations

Andrada, Gloria, Robert W Clowes, and Paul R Smart. 2022. "Varieties of transparency: exploring agency within AI systems." *AI & society*:1-11.

Bender, Emily M, Timnit Gebru, Angelina McMillan-Major, and Shmargaret Shmitchell. 2021. "On the dangers of stochastic parrots: Can language models be too big?" Proceedings of the 2021 ACM conference on fairness, accountability, and transparency.

Boden, Margaret A. 1990. *The Creative Mind: Myths and Mechanisms*. London: Sphere Books Ltd.

Schwitzgebel, Eric, David Schwitzgebel, and Anna Strasser. 2023. "Creating a large language model of a philosopher." *arXiv preprint arXiv:2302.01339*.

Shanahan, Murray. 2022. "Talking about large language models." *arXiv preprint arXiv:2212.03551*.

Giada Coleschi
University of Genoa

Don't Stop Believin': A Journey Into Language Logicality

According to the logicality of language hypothesis (hereinafter, LLH), logical considerations are relevant for syntactic formation to the point that they are needed to explain some ungrammaticalities (Gajewski 2002, 2009; Fox & Hackl 2006; Chierchia 2013, 2021; Abrusán 2014; Del Pinal 2019, 2021). In other words, the syntactic unacceptability of some linguistic constructions is traced back to their logical status, suggesting that speakers judge a word-sequence as ungrammatical or ill-formed when it is always false or always true. The aim of this talk is to examine the recent literature that attempts to offer alternative explanations and test its ability to account for the new examples, involving propositional attitudes, that we will submit.

Chief evidence in support of LLH comes from word-sequences, such as (1)–(4), judged ungrammatical *qua* contradictory.

(1) ˙Some students but John passed the exam (von Fintel 1993)

(2) ˙There are any cookies left (Chierchia 2013)

(3) ˙How fast didn't you drive? (Fox & Hackl 2006)

(4) ˙How did John regret that he behaved at the party? (Abrusán 2007, 2014)

Other word-sequences, such as (5)–(7), are judged ungrammatical *qua* tautological.

(5) ˙There is every fly in my soup (Barwise & Cooper 1981)

(6) ˙Mary is taller than no student is (Gajewski 2008)

(7) ˙At least zero students smoked (Haida & Trinh 2020)

However, in general, contradictions and tautologies are not ungrammatical.

(8) It is raining and it is not raining

(9) It is raining or it is not raining

Within LLH, the syntactic acceptability of (8)–(9) is not questioned. As a result, the asymmetry between the ungrammaticality of (1)–(7), due to their logical status, and

the grammaticality of (8)–(9), notwithstanding their logical status, has to be accounted for. This is the so-called «analyticity puzzle».

According to an early articulation of LLH (Gajewski 2002, 2009; Fox & Hackl 2006; Chierchia 2013), one must assume a "natural" (i.e., purely linguistic) logic and associate linguistic constructions with particularly austere representations, which reveal truly contradictory or tautological contents only in the case of constructions later judged ungrammatical. According to an alternative articulation, on the other hand, contradictions and tautologies are ungrammatical insofar as it is not possible to transform them into informatively adequate contributions, by modulating their lexical content (Del Pinal 2019, 2021; Sauerland 2014; Pistoia-Reda & Sauerland 2021; Pistoia-Reda & San Mauro 2021). There have also been attempts to combine the one and the other articulation. For example, a recent proposal by Chierchia (2021) assumes a natural logic, as in the early articulation, but modulates traditional logical forms, as in the alternative one.

And there have been attempts to explain the rejection of (1)–(7) typologically, i.e., in terms of a failure of composition (Abrusán, Asher & Van de Cruys 2021). To that extent, (1)–(7) would be similar to examples of semantic anomaly, such as (10)–(11).

    (10) #Tigers are Zermelo-Fraenkel sets

    (11) #My toothbrush is pregnant

According to LLH, (1)–(7) are only superficially uninterpretable: when analysed, they receive an interpretation, which is that of being either contradictory or tautological. According to Abrusán, Asher & Van de Cruys (2021), (1)–(7) are instead uninterpretable for, in building up their semantic representation, an insuperable semantic problem is encountered. Semantic anomaly would be the result of a type presupposition that cannot be satisfied (cf., Asher 2011). As far as (10)–(11) are concerned, the type is so high in the type-hierarchy that it has no neighbours that share the same syntactic or semantic dependencies, therefore there is nowhere in the space for its meaning to shift. As far as (1)–(7) are concerned, the type denotes a context-invariant logical meaning that is simply invisible for the distributional methods of distributional semantics, i.e., the computational method Abrusán, Asher & Van de Cruys (2021) resort to. Furthermore, whereas in (10)–(11) type conflict arises at the level of predicate argument composition, in (1)–(7) type conflict arises at a more global level, and therefore involves more linguistic elements (and types, as a consequence). These diversities should explain the intuitive difference between (10)–(11), which are semantically anomalous but grammatical, and (1)–(7), which are «ungrammatical for semantic reasons» (Abrusán, Asher & Van de Cruys 2021: 276).

Let's consider the following:

    (12) Donald is so crazy, he believes [his toothbrush is pregnant]φ

(13) ·Donald is so crazy, he believes [some students but John passed the exam]ψ

The embedded construction ϕ is equivalent to (11), whilst the embedded construction ψ is equivalent to (1). The semantic anomaly of ϕ does not transfer to the compound construction, whilst the ungrammaticality of ψ does. From this, we will argue that:

(i) Against classical discussion (e.g., Mellor 1954), belief ascriptions seem to be sensitive to logical considerations (as already suggested by Varnier & Pistoia-Reda 2022).

(ii) Let's assume, with Abrusán, Asher & Van de Cruys (2021), that the semantic anomaly of (11) and the ungrammaticality of (1) are the result of an irreparable type mismatch. Why then the irreparable type mismatch jeopardises the grammaticality of (13), whereas it does not make (12) semantically anomalous? Note that the point is not that (12)–(13) mimic the intuitive difference between (10)–(11) and (1)–(7). Indeed, (1) is ungrammatical, so it is (13); (11) is semantically anomalous, but (12) is not so. The syntactic unacceptability of (1)–(7) might not be a matter of type mismatch after all.

(iii) Let's concede that the syntactic unacceptability of (1)–(7) is a matter of type mismatch. As Abrusán, Asher & Van de Cruys (2021) highlight, the type mismatch in (1)–(7) is of a very different nature from that in (10)–(11). However, such a difference is motivated on a logically relevant dimension: (1)–(7) involve types denoting a context-invariant logical meaning. As a result, logical considerations are relevant for syntactic formation after all.

References

Abrusán M., 2007, *Contradiction and Grammar: The Case of Weak Islands*, PhD thesis, MIT.

Abrusán M., 2014, *Weak Island Semantics*, Oxford, Oxford University Press.

Abrusán M., N. Asher & T. Van de Cruys, 2021, «Grammaticality and Meaning Shift» in G. Sagi, J. Woods (eds.), *The Semantic Conception of Logic*, Cambridge (MA), Cambridge University Press, pp. 249–276.

Asher N., 2011, *Lexical Meaning in Context: A Web of Words*, Cambridge, Cambridge University Press.

Barwise J. & R. Cooper, 1981, «Generalized Quantifiers and Natural Language», *Linguistics and Philosophy*, 4, 2, pp. 159–219.

Chierchia G., 2013, *Logic in Grammar: Polarity, Free Choice, and Intervention*, Oxford, Oxford University Press.

Chierchia G., 2021, «On Being Trivial: Grammar vs. Logic» in G. Sagi, J. Woods (eds.), *The Semantic Conception of Logic*, Cambridge (MA), Cambridge University Press, pp. 227–248.

Del Pinal G., 2019, «The Logicality of Language: A New Take on Triviality, "Ungrammaticality", and Logical Form», *Noûs*, 53, 4, pp. 785–818.

Del Pinal G., 2021, «The Logicality of Language: Contextualism vs. Semantic Minimalism», Manuscript, University of Illinois.

Fox D. & M. Hackl, 2006, «The Universal Density of Measurement», *Linguistics and Philosophy*, 29, 5, pp. 537–586.

Gajewski J., 2002, «L-analyticity and Natural Language», Manuscript, MIT.

Gajewski J., 2008, «More on Quantifiers in Comparative Clauses» in J. Gajewski, T. Friedman, S. Ito (eds.), *Proceedings of the 18th Semantics and Linguistic Theory Conference (SALT 18)*, Ithaca, Cornell University, pp. 340–357.

Gajewski J., 2009, «L-triviality and Grammar», Handout, UConn Logic Group. Haida A. & T. Trinh, 2020, «Zero and Triviality», *Glossa: A Journal of General Linguistics*, 5, 1, 116, pp. 1– 14.

Mellor W. W., 1954, «Believing the Meaningless», *Analysis*, 15, 2, 41–43.

Pistoia-Reda S. & L. San Mauro, 2021, «On Logicality and Natural Logic», *Natural Language Semantics*, 29, 3, pp. 501–506.

Pistoia-Reda S. & U. Sauerland, 2021, «Analyticity and Modulation: Broadening the Rescale Perspective on Language Logicality», *International Review of Pragmatics*, 13, 1, pp. 1–13.

Sauerland U., 2014, «Making Fuzzy Logic Work for Language», *Semantics and Philosophy in Europe*, 7.

Varnier G. & S. Pistoia-Reda, 2022, «Believing the Formless?», Argumenta, pp. 85–91. von Fintel K., 1993, «Exceptive Constructions», *Natural Language Semantics*, 1, 2, pp. 123–148.

Violeta Conde
University of Santiago de Compostela

with

Concha Martínez-Vidal
University of Santiago de Compostela

To Understand Modality by Understanding Essences

The core question nowadays in the epistemology of modality might be how we can gain modal knowledge. Many responses have come to light in recent years in this regard, many of them based on the distinction popularized by Bob Hale (2003) between possibility-first approaches and necessity-first approaches.

Our aim in this paper is to take Williamson's (2007) counterfactual theory, an apparently neutral position concerning asymmetry (see Mallozzi, 2023), as a starting point to tackle some problems of epistemology of modality related to essences. Therefore, in the first part of the talk, we will introduce briefly Williamson's counterfactual theory and, then, we will explore some critiques it has received, focusing on the problem of *epistemic friction* identified by Vaidya and Wallner (2018); hereafter, we will explore the essence-based account proposed by Vaidya (2010) and we will criticize his method of *variation in imagination* adducing that it is closed under logical consequence and that it provides us with a consequentialist notion of essence instead of a constitutive one (Fine, 1995). Then we will argue that we can obtain constitutive essences starting from consequentialist essences using the notions of *grounding* (Fine, 2012; Schnieder, 2018) and *relevance* (Hirèche, 2023).

According to Williamson (2007), we can gain modal knowledge through analyzing our use of counterfactuals in natural language. Moreover, we can approach metaphysical necessity by understanding it as a limiting case because necessities could be seen as logically equivalent to some counterfactual constructions (Williamson 2007; Vetter, 2016).

But his view is not accepted by some authors (see Jenkins, 2008; Vetter, 2016 for examples). However, our target here is point out what Vaidya and Wallner (2008) have called the problem of *epistemic friction* since, they argue, it applies to a counterfactual development of epistemology of modality.

But what is this problem about? There is an implicit and held background according to which we express counterfactuals. As Williamson says (2007), our background knowledge should be *cotenable* with our imagination when it comes to counterfactuals. In this way, some knowledge about the nature entities involved in counterfactuals is presupposed. We can approach counterfactuals by using our

imagination, but not in a way totally unrestricted. When we imagine what would happen if the vase had fallen, we imagine it could have broken; but we do not imagine the vase violating the laws of physics by coming back to the table. What creates the *epistemic friction* are the essentialist theses that underpin those theories, so the task for the philosopher should be first to articulate an epistemology of essences and then switch from the debate "necessity-first vs. possibility-first" to the debate "essences-first vs. possibility-first".

But how can we know essences? It seems that to know something there should be a connection between the subject and the truth-maker and it does not seem we have that connection in the case of essences. Vaidya's move (2010) elaborates on the difference between "knowing modal claims" and "understanding modal claims": even if both verbs, "to know" and "to understand" are viewed as factive, the first one is incompatible with epistemic luck, whereas the second is not. The basic idea in Vaidya's proposal is that "we can make a judgement about the essential properties of an object by varying properties of the objects in imagination and seeing which vary and which do not" (Vaidya, 2010: 820). The variant properties would be the accidental ones, whereas the invariant would be the essential ones. Vaidya (2010) names the method as variation in imagination and he aims to show that the method is reliable.

To illustrate the method of Vaidya, let's consider Aristotle in first instance: in a first moment, namely, t_1, we can think of him having certain properties, such as "being human", "being a philosopher", "having two legs", "being born in Stagira" etc., in a second moment, t_2 we can think of him as "being human", "being a musician", "having one leg", "being born in Athens", etc. In any case, we cannot think of him as not being human, so the fact of being human will be in the essence of Aristotle (and so will be every property that remains invariant through the process of variation in imagination). This procedure is closed under logical consequence for every disjunction such that one of his members is "being human", for example, could be considered as forming part of the essence of Aristotle. We want to avoid conclusions such a "being human or the moon is made up of cheese" being in the essence of Aristotle".

Fine (1995) is going to distinguish between two approaches to the concept of essence. On the one hand, "an essential property of an object is a constitutive part of the essence of that object if it is not had in virtue of being a consequence of some more basic essential properties of the object; and otherwise, it is a consequential part of the essence" (Fine, 1995b: 57). So, it is constitutively essential to Aristotle to be human, but it is just consequentially essential to him to be a human, or the moon is made up of cheese.

Fine (2012) claims that we should start from properties that are consequentially essential to somebody or something and then depurate this notion to obtain the properties that are constitutively essential. So, when we observe through the method of variation in imagination that the property "Aristotle is human or the moon is made of cheese" remains invariant, we should filtered out that claim in order to obtain the property of Aristotle that form parts of his constitutive essence. But, how to filter that out? Through the notion of grounding and assessing which properties ground the others and remain invariant. Aristotle is human or the

moon is made of cheese is ground in the fact that Aristotle is human (a property that remains invariant). How should be this notion of grounding? We will argue, following Schnieder (2018) and Hirèche (2023) that it should be a relevant one.

References

Fine, K. (1995). "Senses of Essence". In Sinnott-Armstrong, W., Raffman, D. & Asher, N. (Eds.). *Modality, Morality and Belief. Essays in Honor of Ruth Barcan Marcus*. Cambridge: Cambridge University Press, 53 – 73.

Fine, K. (2012). "Guide to ground". In Correia, F. & Schnieder, B. (Eds.). *Metaphysical Grounding. Understanding the Structure of Reality.* Cambridge: Cambridge University Press: 77 – 95.

Hale, B. (2003). "The Presidential Address: Knowledge of Possibility and of Necessity", *Proceedings of the Aristotelian Society*, 103: 1 – 20. doi: 10.1111/j.0066-7372.2003.00061.x

Hirèche, S. (2023). "Grounding, necessity, and relevance", *Philosophical Studies*. doi: 10.1007/s11098-023-01968

Mallozzi, A. (2023). "The Epistemology of Modality". In Zalta, E.N. (Ed.), *The Stanford Encyclopedia of Philosophy* (Fall 2021 Ed.). URL = < https://plato.stanford.edu/entries/modality-epistemology/#exFE>

Schnieder, B. (2018). "On Ground and Consequence", *Synthese* (198): 1335 – 1363. doi: 10.1007/s11229-018-02012-9

Vaidya, A. J. (2010). "Understanding and Essence", *Philosophia* (38): 811 – 833. doi: 10.1007/s11406-010-9243-z

Vaidya, A. J. & Wallner, M. (2021). "The epistemology of modality and the problem of modal epistemic friction", *Synthese* (Suppl 8): 1909 – 1935.

Vetter, B. (2016): "Williamsonian modal epistemology, possibility-based", *Canadian Journal of Philosophy*, 46 (Issue 4 – 5: Williamson on Modality): 766 – 795. doi: 10.1080/00455091.2016.1170652

Williamson, T. (2007). *The Philosophy of Philosophy*. Oxford: Blackwell Publishing.

Marc Lara Crosas

University of Barcelona

Unmasking Safety

Safety-based accounts of knowledge defend that a belief being *safe* is at least a necessary condition, or even sufficient, for knowledge (Sosa 1999; Williamson 2000; Pritchard 2009; Mortini 2023). A belief is safe if and only if it could not have been easily false, where this is usually unpacked in modal terms: roughly, a belief is safe if and only if in all close possible worlds where the subject forms the same belief as in the actual world following the same belief-forming method, the belief is also true. Despite its initial appeal as an anti-skeptical condition (Sosa 1999) or as an anti-luck condition (Pritchard 2009), safety has faced objections regarding both its necessity and sufficiency for knowledge.

Beddor and Pavese (2020) have argued that certain counterexamples to the sufficiency of safety are analogous to cases of mimicking, and "this suggests that the problem is not specific of knowledge (...). Rather, it is a problem that arises for dispositions more generally" (Beddor and Pavese 2020: 69). Following their lead, in this talk, I will argue that some prominent counterexamples to the necessity of safety (Baumann 2008; Kelp 2009; Bogardus 2014) are cases of masking. If safety is understood in dispositional terms, this suggests a promising way to answer these counterexamples: one should simply apply to these counterexamples one's preferred solution to cases of masking (e.g., Manley and Wasserman 2008).

Paradigmatic cases of dispositions include the disposition of a glass to break if hit or the disposition of salt to dissolve in water. It is widely accepted that dispositions are closely connected to counterfactual conditionals (e.g., "if the glass were hit, then it would break"), in which the stimulus conditions of the disposition are placed as the antecedent and its manifestation as the consequent. For this reason, it might be tempting to endorse the Simple Conditional Analysis of dispositions (henceforth "SCA"):

> SCA: An object x possesses a disposition to M when S if and only if,
> were S to obtain, x would M (where "M" stands for the manifestation
> of the disposition and "S" for its stimulus conditions)

SCA is almost universally rejected as a reductive analysis of dispositions since it falls prey to straightforward counterexamples, such as cases of masking. In cases of masking,
an extrinsic object would causally interfere between the stimulus conditions and the manifestation of the disposition, if the stimulus conditions obtained, likely preventing the manifestation of the disposition. For instance, imagine a glass that is carefully protected with styrofoam so it does not break. This is a case of masking and a counterexample to SCA because the glass retains its disposition to break

when hit, but the conditional "if the glass were hit, it would break" is false. In other words, the analysandum of SCA is true while the analysans is false.

The relevant dispositions for safety would have the formation of a belief as its stimulus conditions and the truth of that belief as its manifestation. After all, even if safety need not be formulated in these terms, it has been formulated as a counterfactual: if S were to believe that p, then p would be the case (Sosa 1999). Since safety is a modal condition that could be formulated as a counterfactual, and dispositions are modal properties closely connected to counterfactual conditionals; it should not be surprising to analyze safety in dispositional terms. Assuming a dispositional analysis of safety, I will argue that at least some of the counterexamples to the necessity of safety for knowledge (Baumann 2008; Kelp 2009; Bogardus 2014) are analogous to cases of masking. In these counterexamples, there is also an extrinsic entity that would causally interfere between the stimulus conditions of a disposition and its manifestation, were the stimulus to obtain, likely preventing the manifestation.

One might object that there is a clear disanalogy between these counterexamples (Baumann 2008; Kelp 2009; Bogardus 2014) and cases of masking: whereas dispositions are not manifested in cases of masking, the relevant dispositions are manifested in these counterexamples because true beliefs are formed. I will answer this objection by relying on Turyn's (2021) analysis of masks: dispositions may still manifest even if masked, the manifestation is just less likely to occur. For instance, a glass carefully protected with styrofoam can still break when hit and, arguably, its disposition to break when hit would still be masked despite actually breaking.

To sum up, analyzing safety in dispositional terms advances the debate since it suggests well-motivated answers to at least some of the objections that safety has received both to its necessity and sufficiency for knowledge. At least some of the counterexamples to the necessity and sufficiency of safety for knowledge are analogous to cases of masking and mimicking, respectively. This suggests these are not problems of safety in particular, but of dispositions in general, and that one should look for solutions in the metaphysics of dispositions. In addition, analyzing safety in dispositional terms provides an ontological basis for this well-known modal epistemic condition: safety is grounded in the dispositions to form true beliefs of epistemic agents.

References

Baumann, Peter. 2008. «Is Knowledge Safe?» *American Philosophical Quarterly (Oxford)* 45 (1): 19-30.

Beddor, Bob, i Carlotta Pavese. 2020. «Modal Virtue Epistemology». *Philosophy and Phenomenological Research* 101 (1): 61-79. https://doi.org/10.1111/phpr.12562.

Bogardus, Tomas. 2014. «Knowledge Under Threat». *Philosophy and Phenomenological Research* 88 (2): 289-313. https://doi.org/10.1111/j.1933-1592.2011.00564.x.

Kelp, Christoph. 2009. «Knowledge and Safety». *Journal of Philosophical Research* 34 (July): 21-31. https://doi.org/10.5840/jpr_2009_1.

Manley, David, i Ryan Wasserman. 2008. «On Linking Dispositions and Conditionals». *Mind* 117 (465): 59-84. https://doi.org/10.1093/mind/fzn003.

Mortini, Dario. 2023. «Etiological Proper Function and the Safety Condition». *Synthese (Dordrecht)* 202 (6): 203-. https://doi.org/10.1007/s11229-023-04412-y.

Pritchard, Duncan. 2009. «Safety-Based Epistemology: Whither Now?» *Journal of Philosophical Research* 34 (July): 33-45. https://doi.org/10.5840/jpr_2009_2.

Sosa, Ernest. 1999. «How to Defeat Opposition to Moore». *Noûs (Bloomington, Indiana)* 33 (s13): 141-53. https://doi.org/10.1111/0029-4624.33.s13.7.

Turyn, Gus. 2021. «On Dispositional Masks». *Synthese (Dordrecht)* 199 (5-6): 11865-86. https://doi.org/10.1007/s11229-021-03315-0.

Williamson, Timothy. 2000. *Knowledge and Its Limits*. Oxford; Oxford University Press.

Rui Vieira da Cunha

University of Porto/Catholic University of Porto

Parfit's Embodied Person view -The evolution of Parfitian brains and their (un)importance for personal ontology

PARFIT's original position on the nature of persons (1971, 1984) did not hold a special place for their brains. His final word on personal ontology (2012), however, did make brains into the most important element of his account. This paper aims to explain why and how that evolution took place and what it means to Parfitian reductionism about persons.

According to PARFIT's original views, his constitutive reductionism about persons, beings like you and I were a bundle of experiences, either mental or physical, but undifferentiated. Persons were thus to be understood as various interrelated physical and mental events, dependent on their brains and bodies for instantiation, yes, but not for their persistence, at least not necessarily. Besides, personal identity was not what mattered. The importance of personal identity, as it had been traditionally conceived, was replaced by the importance of Relation-R, understood as psychological connectedness and continuity, always a matter of relations between mental or psychological events. Of Lockean inspiration, this view was heavily influenced by SHOEMAKER but also by GRICE. The relevant element in this original notion of personhood was neither the brain nor the body but the bundle of experiences - hence, the common association of Parfitianism with HUME and with Buddhist views – and a person's brain was only relevant as a carrier of information, of psychological continuity, or Relation-R. It was the experiences that mattered, the series of experiences and the causal connections between them. In our ordinary life up to this point, these causal connections depended on the brain but nothing of necessity turned on it.

During his lifetime, though, and in replying to different critiques, especially from the animalist camp (OLSON 1997) but also from the neo-Kantian side (CASSAM 1989, 1993), PARFIT changed his reductionist account significantly. It began to evolve in 1999, by conceding that we are essentially animals and may exist before becoming persons and continue to exist after ceasing to be persons, thus treating person as a phase sortal. Its final stage, however, is the so-called "Embodied

Person View", according to which PARFIT holds that beings like you and I are simultaneously a person in the Lockean sense and the conscious thinking part of a human animal - its cerebrum. By contrast with the original view, the brain is now more than a mere carrier of psychological connectedness and continuity, it actually seems to be the person, thus standing in direct opposition to the original constitutive reductionism with its bundle view.

The paper finishes with an appreciation of the ways in which PARFIT (2012) can be said to be a denial or a withdrawal of the original reductionist view into a veiled form of the (essentialist and anti-reductionist) brain criterion of personal identity, similar to Nagel's view in 1971 that Parfit earlier rejected, and what its implications for the rational and moral importance of personal identity may be, namely, whether PARFIT can truly still claim that personal identity is not what matters.

Pablo Valencia Equy

University of Santiago de Compostela

Truthmaking Without Truthmakers and Arithmetic Without Numbers

An important question arises in the semantics of arithmetic: *What kind of entity*—if there is one—*explains the truth of an arithmetical statement?* The classical-realist answer of truthmaker theory is that arithmetical statements can only be made true by worldly entities (Mulligan et al. 1984; Fox 1987; Bigelow 1988; Martin 1996; Armstrong 1997, 2003, 2004, 2007; Cox 1997; Molnar 2000; Rodriguez-Pereyra 2005; Heil 2006; Cameron 2008; Pendlebury 2010; Jago 2018; Asay 2022; etc). Truthmaker theorists then conclude that the truth of arithmetical statements requires substantive commitments from the spacetime world. In the first part of the talk, I would argue that this view implies a clear contradiction when it comes to making true naive arithmetical statements, such as $7+5 = 12$. In short, my argument is supported by the assumption—made explicit by this realist version of truthmaker theory—that each arithmetical statement must have only one truthmaker (Armstrong 1997, 2004). Moreover, to explain the truth of arithmetical statements, in one step of their argument (which would be discussed in the talk), the realist truthmaker theorist resorts to the entailment principle (Bigelow 1988; Restall 1996; Armstrong 1997, 2003, 2004, 2007; Read 2000; Rodriguez-Pereyra 2006; Cameron 2008; Tałasiewicz et al. 2013; Jago 2018, etc):

If τ is a truthmaker of p and $p \to q$, then τ is also a truthmaker of q (**EP**)

Now, $7+5 = 12$ is a necessary truth; therefore, $3+1 = 4 \to 7+5 = 12$. Let α be the truthmaker of $3+1 = 4$ and β the truthmaker of $7+5 = 12$. By (**EP**), α must also be the truthmaker of $7+5 = 12$. However, according to the realist approach to truthmaker theory, $\alpha \neq \beta$, since $7+5 = 12$ and $3+1 = 4$ are different arithmetical truths; thus, they must be made true by different worldly entities, or, at least, by a different number of them (Bigelow 1988; Armstrong 1997, 2004). But since $7+5 = 12$ has only one truthmaker, and by (**EP**), α is a truthmaker of $7+5 = 12$, then it must be the case that $\alpha = \beta$. I would also argue that this contradiction can be avoided if the truthmakers of arithmetical statements are not conceived of as worldly entities.

I would then continue with my own proposal, according to which it is not necessary to postulate entities of any kind to guarantee the truth

of arithmetical statements—not even mathematical facts or numbers conceived as abstract objects of some kind. Thus, I would conclude that the truth of arithmetical statements does not necessarily require a committment to entities of any kind, whether concrete or abstract.

To support this view, I would first argue for a structuralist approach—one informed particularly by the contributions made by Hellman in 1991 and in 1998—in which the place of the truthmaker in the truthmaking relation could be occupied by placeholders linked by functional relations and not necessarily by numbers as individuated objects (see also Reck and Price 2000; Florio 2018). Hence, we can make arithmetical statements true without committing ourselves to numbers or other mathematical facts as truthmakers,[1] instead relying solely on the following truthmaking principle:[2]

$$\text{If } PA^2(\mathbb{N}, s, 0, +, \cdot) \models \Lambda(\mathbb{N}, s, 0, +, \cdot), \text{ then } PA^2(\mathbb{N}, s, 0, +, \cdot) \models \alpha \quad (\models\mathbf{PA})$$

where $\models$ denotes the truthmaking relation (e.g., $X \models \varphi$ means that φ is made true by X), PA^2 is the conjunction of the second-order Peano axioms, and $PA^2(\mathbb{N}, s, 0, +, \cdot)$ denotes the non-logical terms used in the axioms. Finally, $\Lambda(\mathbb{N}, s, 0, +, \cdot)$ is the formalization in Peano arithmetic of a naive arithmetical statement α—e.g., $+[s(s(0)), s(s(s(0)))] = s(s(s(s(s(0)))))$ is the formalization of $2 + 3 = 5$.

Therefore, what is ontologically relevant in this truthmaking statement are the relations among the items and not their alleged nature. One might object, however, that the truthmaking relation itself requires substantive ontological committments in the sense that it reflects some kind of ontological dependence between the truthbearer and the truthmaker (Lewis 1986a, 1986b; Fine 2001; Schaffer 2009, 2010, 2016, 2017; Sider 2011; Daly 2012; Hofweber 2016; Kovacs 2017; McDaniel 2017; Wilson 2018a; 2018b). This would imply that, even if it were not necessary to postulate abstract entities such as mathematical facts as truthmakers, it would still be mandatory to postulate a metaphysical structure in asserting statements such as ($\models\mathbf{PA}$). But this structure would presuppose an ontological priority of the truthmaker over the truthbearer, thus making explicit a substantive committment to a hierarchical picture of reality.

To argue that ($\models\mathbf{PA}$) does not engage with any kind of ontological dependence relation—assuming some kind of metaphysical structure, whether abstract or concrete—I would show that this principle can be derived from other principles that are empty of ontological content. For instance, ($\models\mathbf{PA}$) can be obtained from ($\triangleright\mathbf{PA}$):

$$\text{If } T[\Lambda(\mathbb{N}, s, 0, +, \cdot)], \text{ then } T[\Lambda(\mathbb{N}, s, 0, +, \cdot)] \triangleright T\alpha^3 \quad (\triangleright\mathbf{PA})$$

---

[1] Of course, this argument only establishes that the truth of arithmetical statements can be explained without recourse to mathematical facts as truthmakers, but this does not imply that there are no independent reasons to believe in their existence.

[2] See, for similar principles, Fine 2016; Donaldson 2017; Shumener 2017, 2020; Florio 2018; Correia and Skiles 2019; Florio and Carrara 2020.

[3] As usual, "T" stands for the truth predicate.

where $\rhd$ expresses a grounding relation between $T[\Lambda(\mathbb{N}, s, 0, +, \cdot)]$ (the ground) and $T\alpha$ (what is grounded). However, in this principle, $\rhd$ can be understood as a primitive sentential operator, not necessarily one that relates entities of any kind (Melia 2005; Schnieder 2006, 2011; deRosset 2013; Thompson 2016; Dasgupta 2017; Maurin 2019). In this sense, ($\rhd\mathbf{PA}$) would express a trivial dependency relation that does not imply a committment to a hierarchical structure of reality. Assuming additional truthmaking and grounding principles, I would then prove the following theorem:

Theorem 1. ($\models\mathbf{PA}$) *follows from* ($\rhd\mathbf{PA}$).

Now, if ($\models\mathbf{PA}$) is obtained from a principle empty of ontological content, and in the proof of the theorem nothing is added that requires ontological committments, then the $\models$ in ($\models\mathbf{PA}$) can be interpreted as a primitive sentential operator that does not necessarily relate truthmakers to truthbearers. This truthmaking principle would then reflect a trivial dependence relation that does not reify reality in any sense.

In conclusion, in this talk I would argue that the answer to the opening question is that explaining the truth of arithmetical statements does not require a committment to entities of any kind, whether abstract or concrete. On the contrary, this deflationary approach suggests that the truth of such statements is grounded in principles, such as ($\models\mathbf{PA}$), that do not engage with either substantive entities or a hierarchical structure of reality.[4] Rather, these principles are expressions of a trivial dependency relationship, denoted by $\models$, which is in no way a reification of reality. Therefore, if arithmetical truths are founded on such principles, it only remains to say that their truth conditions are, in some sense, trivial.[5]

References

Armstrong, D. (1997). *A World of States of Affairs*. Cambridge University Press.

Armstrong, D. (2003). Truthmakers for modal truths. In Lillehammer, H. and Rodriguez-Pereyra, G., editors, *Real Metaphysics: Essays in Honour of D. H. Mellor*, pages 12–24. Routledge.

Armstrong, D. (2004). *Truth and Truthmakers*. Cambridge University Press.

Armstrong, D. (2007). Truthmakers for negative truths, and for truths of mere possibility. In Monnoyer, J., editor, *Metaphysics and Truthmakers*, page 99. Ontos Verlag.

---

[4]See footnote 2 for references to other principles similar to ($\models\mathbf{PA}$). I would consider some of them in the talk, but I would also briefly argue that ($\models\mathbf{PA}$) captures the truth of arithmetical statements better than these other principles.

[5]It is easy to see that my proposal is similar to A. Rayo's mathematical trivialism (2003, 2008, 2009, 2013). However, the arguments provided are very different. For instance, while Rayo uses "just is" statements, I employ truthmaking principles.

Asay, J. (2022). *A Theory of Truthmaking*. Cambridge University Press.

Bigelow, J. (1988). *The Reality of Numbers*. Clarendon Press.

Cameron, R. (2008). Truthmakers and ontological commitment: or how to deal with complex objects and mathematical ontology without getting into trouble. *Philosophical Studies*, 140:1–18.

Correia, F. and Skiles, A. (2019). Grounding, essence, and identity. *Philosophy and Phenomenological Research*, 98(3):642–670.

Cox, D. (1997). The trouble with truth-makers. *Pacific Philosophical Quarterly*, 78:45–62.

Daly, C. (2012). Scepticism about grounding. In Correia, F. and Schnieder, B., editors, *Metaphysical Grounding: Understanding the Structure of Reality*, pages 81–100. Cambridge University Press.

Dasgupta, S. (2017). Constitutive explanation. *Philosophical Issues*, 27(1):74–97.

deRosset, L. (2013). Grounding explanations. *Philosophers' Imprint*, 13.

Donaldson, T. (2017). The (metaphysical) foundations of arithmetic? *Noûs*, 51(4):775–801.

Fine, K. (2001). The question of realism. *Philosophers' Imprint*, 1:1–30.

Fine, K. (2016). Identity criteria and ground. *Philosophical Studies*, 173(1):1–19.

Florio, C. D. (2018). On grounding arithmetic. In Giordani, A. and de Florio, C., editors, *From Arithmetic to Metaphysics: A Path Through Philosophical Logic*, pages 103–118. De Gruyter.

Florio, C. D. and Carrara, M. (2020). Identity criteria: An epistemic path to conceptual grounding. *Synthese*, 197(7):3151–3169.

Fox, J. (1987). Truthmaker. *Australasian Journal of Philosophy*, 65(2):188–207.

Heil, J. (2006). The legacy of linguisticism. *Australasian Journal of Philosophy*, 2(84):233–244.

Hellman, G. (1991). Mathematics without numbers. towards a modal-structural interpretation. *Tijdschrift Voor Filosofie*, 53(4):726–727.

Hellman, G. (1998). Mathematical constructivism in spacetime. *British Journal for the Philosophy of Science*, 49(3):425–450.

Hofweber, T. (2016). *Ontology and the Ambitions of Metaphysics*. Oxford University Press UK, Oxford, England.

Jago, M. (2018). *What Truth Is*. Oxford University Press.

Kovacs, D. M. (2017). Grounding and the argument from explanatoriness. *Philosophical Studies*, 174(12):2927–2952.

Lewis, D. (1986a). Causal explanation. In Lewis, D., editor, *Philosophical Papers Vol. Ii*, pages 214–240. Oxford University Press.

Lewis, D. (1986b). *Philosophical Papers, Volume Ii.* Oxford University Press, New York, US.

Martin, C. (1996). Entities, absences and voids. *Australasian Journal of Philosophy*, 74(1):57–65.

Maurin, A. (2019). Grounding and metaphysical explanation: It?s complicated. *Philosophical Studies*, 176(6):1573–1594.

McDaniel, K. (2017). *The Fragmentation of Being.* Oxford University Press, New York.

Melia, J. (2005). Truthmaking without truthmakers. In Beebee, H. and Dodd, J., editors, *Truthmakers: The Contemporary Debate*, pages 67–84. Clarendon Press.

Molnar, G. (2000). Truthmakers for negative truths. *Australasian Journal of Philosophy*, 78(1):72–86.

Mulligan, K., Simons, P., and Smith, B. (1984). Truth-makers. *Philosophy and Phenomenological Research*, 44(3):287–321.

Pendlebury, M. (2010). Facts and truth-making. *Topoi*, 29(2):137–145.

Rayo, A. (2003). Success by default? *Philosophia Mathematica*, 11(3):305–322.

Rayo, A. (2008). On specifying truth-conditions. *Philosophical Review*, 117(3):385–443.

Rayo, A. (2009). Towards a trivialist account of mathematics. In Bueno and Linnebo, editors, *New Waves in Philosophy of Mathematics.* Palgrave-Macmillan.

Rayo, A. (2013). *The Construction of Logical Space.* Oxford University Press, Oxford, England.

Read, S. (2000). Truthmakers and the disjunction thesis. *Mind*, 109(432):67–79.

Reck, E. H. and Price, M. P. (2000). Structures and structuralism in contemporary philosophy of mathematics. *Synthese*, 125(3):341–383.

Restall, G. (1996). Truth-makers, entailment and necessity. *Australasian Journal of Philosophy*, 74:331–340.

Rodriguez-Pereyra, G. (2005). Why truthmakers? In Beebee, H. and Dodd, J., editors, *Truthmakers: the contemporary debate*, pages 17–31. Oxford University Press.

Rodriguez-Pereyra, G. (2006). Truthmaking, entailment, and the conjunction thesis. *Mind*, 115(460):957–982.

Schaffer, J. (2009). On what grounds what. In Manley, D., Chalmers, D. J., and Wasserman, R., editors, *Metametaphysics: New Essays on the Foundations of Ontology*, pages 347–383. Oxford University Press.

Schaffer, J. (2010). The internal relatedness of all things. *Mind*, 119(474):341–376.

Schaffer, J. (2016). Grounding in the image of causation. *Philosophical Studies*, 173(1):49–100.

Schaffer, J. (2017). The ground between the gaps. *Philosophers' Imprint*, 17.

Schnieder, B. (2006). Truth-making without truth-makers. *Synthese*, 152(1):21–46.

Schnieder, B. (2011). A logic for because. *Review of Symbolic Logic*, 3(4):445–465.

Shumener, E. (2017). The metaphysics of identity: Is identity fundamental? *Philosophy Compass*, 12(1):1–13.

Shumener, E. (2020). Identity. In Raven, M. J., editor, *Routledge Handbook of Metaphysical Grounding*, pages 413–424. Routledge.

Sider, T. (2011). *Writing the Book of the World*. Oxford University Press, Oxford, England.

Tałasiewicz, M., Odrowka.z-Sypniewska, J., Wciórka, W., and Wilkin, P. (2013). Do we need a new theory of truthmaking? some comments on disjunction thesis, conjunction thesis, entailment principle and explanation. *Philosophical Studies*, 165(2):591–604.

Thompson, N. (2016). Grounding and metaphysical explanation. *Proceedings of the Aristotelian Society*, 116(3):395–402.

Wilson, A. (2018a). Grounding entails counterpossible non-triviality. *Philosophy and Phenomenological Research*, 96(3):716–728.

Wilson, A. (2018b). Metaphysical causation. *Noûs*, 52(4):723–751.

Domingos Faria
University of Porto

Evaluative and Prescriptive Norms of Group Inquiry

In this talk, we want to extend the norms of inquiry for groups or collectivities, examining whether or not they are epistemically responsible for such norms. To achieve this goal, we will first start with the norms of inquiry for individual cases.

In a first step, following McHugh (2012), Simion, Kelp, and Ghijsen (2016), and Simion (2019), we can draw a useful distinction in normativity theory between evaluative and prescriptive norms. On the one hand, evaluative norms are primarily about what is good or valuable and do not presuppose accountability or blameworthiness. On the other hand, prescriptive norms are mainly about what one ought to do, and we are accountable to them in the sense that violating them is likely to leave us open to blame. Despite being distinct norms, there is a relation between evaluative norms and prescriptive norms. Evaluative norms often have implications for prescriptive norms; we can derive a prescriptive norm from an evaluative norm precisely because prescriptive norms make it likely that evaluative norms will be met. Based on this framework and inspired by a knowledge-first epistemology (cf. Williamson (2021) and Lasonen-Aarnio (2021)), we can argue for the following norms of inquiry:

- **Evaluative norm:** One's inquiry into a question Q at t is a good inquire only if one knows at t that Q has a true answer.
- **Prescriptive norm**: One must inquiry into a question Q at t only if one has good cognitive dispositions in believing at t that Q has a true answer.

In this account, the evaluative norm comes first and determines what counts as a good inquiry; namely, it states that a good inquiry qualifies as knowledge. From this, we derive the prescriptive norm, which determines what one should do; namely, it prescribes that one should inquire only if one has virtuous beliefs that Q has a true answer, resulting from good cognitive dispositions (i.e., good dispositions about belief formation and retention). Such a prescriptive norm gives a general guide for action, making compliance with the evaluative norm likely. That is so because, in general, good cognitive dispositions typically generate knowledge, and bad cognitive dispositions usually do not generate knowledge. More specifically, cognitive dispositions are classified as "good" or "bad" according to the quality of the epistemic states they tend to manifest across normal counterfactual cases (in which knowledge is the best quality status). Here, we understand normal counterfactual cases as non-deviant cases relative to the subject of evaluation, but where the type of situation and the disposition manifested are very similar (to the actual case).

Since the function of the prescriptive norm is to reinforce behavior conducive to compliance with the evaluative norm by giving us some action guide, we have responsibility for prescriptive norms in the sense that in typical circumstances, we can be blamed for violating them or we can be praised for complying with them. However, we also concede that developing virtuous dispositions is not always accessible from the subject's standpoint. For example, one could argue that the prescriptive norm has no force for subjects not in a position to develop virtuous dispositions. In such cases, they are blameless for violating the prescriptive norm. Thus, we can specify that the proposed norms of inquiry only apply in circumstances or communities conducive to the development of virtuous dispositions. Outside of these favorable circumstances, the subject can be excused for violating the prescriptive norm of inquiry.

In a second step, building on this previous theoretical framework for norms of inquiry for individual cases, we argue that such norms can also apply to groups or collectivities in a non-summative way. In this case, we maintain that there is a prescriptive norm for which groups can be held accountable. Namely, a group G must inquire into a question Q at t only if G has good cognitive dispositions in believing at t that Q has a true answer. In a nutshell, groups can also be responsible for their inquiries by being blameworthy or praiseworthy in their compliance with the prescriptive norm.

References

Lasonen-Aarnio, Maria. 2021. "Perspectives and Good Dispositions." Philosophy and Phenomenological Research.

McHugh, Conor. 2012. "The Truth Norm of Belief." Pacific Philosophical Quarterly 93 (1): 8–30.

Simion, Mona, Christoph Kelp, and Harmen Ghijsen. 2016. "Norms of Belief." Philosophical Issues 26 (1): 374–92.

Simion, Mona. 2019. "Knowledge-First Functionalism." Philosophical Issues 29 (1): 254–67.

Williamson, Timothy. 2021. "E=K, but What about R?" In The Routledge Handbook of Evidence, edited by Maria Lasonen-Aarnio and Clayton Littlejohn. London and New York: Routledge.

Manuel Ferrer
NOVA University Lisbon

What is for a disagreement to be rationally resolvable?

As a first approach, a disagreement is rationally resolvable when there is a rational path that each party can take to agreement (Matheson 2021), being a rational path a way of exchanging arguments that lead the parties to converge on the issue at hand. Disagreements where this is not the case are called ratio nally irresolvable disagreements (RID in what follows). RID has been used to argue for important theses in different areas of philosophy, especially in the liter ature on deep disagreements (Fogelin 1985; Ranalli and Lagewaard 2022), where these are claimed to be rationally irresolvable, but also in epistemology, where it has been defended that the existence of RID supports epistemic relativism (Pritchard 2011). Despite its importance, few efforts have been devoted to de veloping a true and substantive definition of rational resolvable disagreements (RRD in what follows). In most cases we encounter correct but too schematic definitions of RRD (Fogelin 1985; Matheson 2021; Ranalli and Lagewaard 2022). The only exception to this trend is that of Melchior (2023), who provides a sub stantive definition of RID. However, his account is subject to several important counterexamples, regarding his notion of subjective rationality, that render it incorrect.

In this talk, I will provide a theory of rational resolvability. To do this, I will first present the definition of RRD and, afterward, I will give an account for each of the key terms that are part of it. The definition is as follows:

A disagreement on *p* between agents *a* and *b* is a RRD *iff* there is one *dialectically correct argumentative interaction* that is *available* for *a* and *b* such that, after this interaction, *a* and *b* have the same position on whether *p*.

Accordingly, the building blocks of this theory are:

1. A definition of argumentative interactions.

2. An account of what makes an argumentative interaction available, in the relevant sense, for a pair of agents.

3. A definition of what makes an argumentative interaction dialectically virtuous.

Regarding the first building block, I define an argumentative interactions between two agents (*a* and *b*) as any series of argumentative actions performed by these agents that results from iterations of the following process:

1. One of the two agents presents an argument to the other party.

2. The attacked party responds to that argument.

Where a response to an argument can be either a rejection or an acceptance of that argument. Here I defend that anything that is a move in an argumen tative interaction can be captured as either the presentation of an argument, a rejection, or an acceptance. It is also important to note that the notion of argumentative action is taken to capture whatever counts as a meaningful act inside an argumentative interaction, in an independent way on whether it ful fills the relevant normative standard (i.e. of being a good argument, a justified response, etc.).

Concerning the second building block, the notion of available argumentative interaction is constructed from that of available argumentative action, in the sense that the former consists of a series of argumentative actions available to the parties. Regarding the sense of "availability" that is at work here, it is not the one implied by the ordinary use of "can" (i.e. implied by the biconditional: ϕ is available for *a iff a* can, in the ordinary sense, ϕ). That the ordinary sense of "can" is not the one that we are looking for in a theory of RRD is clear when we consider scenarios where a disagreement is RR but the agents can't (in the ordinary sense) perform the relevant argumentative actions to solve it.

In the ordinary use of "can", "*S* can ϕ" is true *iff S* has ϕ related capacities and has an opportunity to exercise them (Wedgwood 2013). But in the RR sense of "can", agents *can* perform argumentative actions even if it turns out that they will never have an opportunity to exercise them. This particular sense of "can" will be made precise by providing a semantics for it: one that can be obtained by making some small adjustments to the ordinary semantics for this term (Mandelkern et al. 2017). Also, I'll defend that the main reason why we need this special sense of "can" is that RR involves judgments about the argumentative capacities of agents and not about the real possibility that they resolve their disagreement through argumentation.

Regarding the last building block of the theory –what makes an argumen tative interaction dialectically virtuous– it is necessary because mere convergence through argumentation doesn't amount to resolvability. Only convergence through correct argumentative interaction counts, or what is the same, convergence through dialectically virtuous argumentative interactions. These are argumentative interactions composed of dialectically virtuous argumentative actions. For argumentative actions to be such, they have to satisfy a bidi mensional normative standard: one objectual, which focuses on the result of the action, and another procedural, which focuses on how this action is performed. Following this, a presentation of an argument is dialectically virtuous when the argument is one of the best ones that the agent *can* provide (in the earlier determined special sense of "can") and also is presented in a dialectically virtuous manner (i.e.: in a clear way, disposed to rephrase it or provide clarifications if needed). The same can be said regarding dialectically virtuous rejections or revisions for a given argument. For instance, a dialectically virtuous rejection of an argument is one where the agent

presents the best reasons that she *can* pro vide that effectively defeat the argument presented, and does it in a dialectically virtuous manner.

Once these building blocks are set up and with the earlier definition of RRD, we'll know what is for a disagreement to be rationally resolvable.

References

1. Fitzpatrick, S. (2014). Moral realism, moral disagreement, and moral psychology. Philosophical Papers, 43(2), 161–190.

2. Fogelin, R. J. (1985). The Logic of Deep Disagreements. Informal Logic , VII(1), 1–8.

3. Mandelkern, M., Schultheis, G., and Boylan, D. (2017). Agentive Modals. The Philosophical Review, 126(3), 301–343.

4. Matheson, J. (2021). Deep Disagreements and Rational Resolution. Topoi, 40(5), 1025–1037.

5. Melchior, G. (2023). Rationally irresolvable disagreement. Philosophical Studies, 180(4), 1277–1304.

6. Pritchard, D. (2011). Epistemic Relativism, Epistemic Incommensurability, and Wittgensteinian Epistemology. In S. D. Hales (Ed.), A Companion to Relativism (pp. 266–285). Wiley-Blackwell.

7. Ranalli, C., and Lagewaard, T. (2022). Deep Disagreement (Part 2): Epistemology of Deep Disagreement. Philosophy Compass.

8. Wedgwood, R. (2013). Rational 'Ought' Implies 'Can.' Philosophical Issues, 23, 70–92.

Carolina Flores
UC Santa Cruz

Beyond misinformation: How the platform web distorts epistemic agency

Online misinformation is often blamed for contemporary epistemic pathology—the rise of conspiracy theories, anti-science beliefs, and the like. In particular, the internet is taken to primarily play a role in contemporary epistemic pathology by spreading misinformation, i.e., false or misleading information (O'Connor and Weatherall 2019).

I argue that this story is incomplete. The internet makes another, underappreciated contribution to epistemic pathology by reshaping users' epistemic agency. Specifically, I argue that the platform web (the algorithmic, social-media-centered internet that has become dominant since the early 2010s) inculcates and habituates agents to new, and often epistemically detrimental, epistemic styles. Our online lives change *how* we interact with evidence, not just *what* evidence we have.

To argue for this, I begin by reviewing existing work on epistemic styles (Flores 2021). Epistemic styles are unified way of interacting with evidence which are often socially molded. They express epistemic preferences, values, and policies. These include things like Jamesian preferences for collecting truths vs. avoiding falsehoods, theoretical values like how much an agent values simplicity relative to goodness of fit, the agent's testimonial policies (whose testimony they trust), and their evidential threshold policies (how much evidence the agent requires to change their mind for different beliefs; Morton and Paul 2018). I suggest that, much as practical agency can be a matter of deploying different agential modes (Nguyen 2020), our epistemic agency is constituted by the set of epistemic styles that we have at our disposal. And it is here that the platform web operates on epistemic agency.

The platform web has three features that make it a powerful device for inculcating new, distinctive epistemic styles. First, it is a powerful site for social normativity, with emerging social norms on discourse that are associated with systems of social reward and punishment. These social norms imply, I show, norms on how to reason, to which agents then face social pressure to conform. Second, online platforms are designed to maximize engagement. This leads to the prioritization of content with distinctive features: content that is high in shock value or promotes outrage, encourages in-grouping and out-grouping, maximizes a sense

of clarity, among others (Nguyen 2023). Via social conformity and informational cascades (Sunstein 2017), online agents learn to reason in ways that reflect or express these features. Third, these platforms are designed for fast, superficial, repetitive use. This makes internet use a powerful tool for habituation into new epistemic styles: for acquiring the dispositions that constitute these styles, not just learning that others use them and being encouraged to conform at a superficial level.

I then offer concrete examples of epistemic parameters that tend to be reshaped by exposure to environments with these features. Specifically, I argue that agents' testimonial policies tend to become narrower and more extreme, focused on trusting only and intensely the testimony of in-group members; that agents' evidential threshold policies tend to be transformed to incorporate a higher threshold for revising a wide beliefs, as identity-protective reasoning (Kahan 2017) becomes a more important driver; and that agents' Jamesian preferences tend to change to value collecting truths over avoiding falsehood (Fraser 2020).

Here I'll sketch an example, that of changes to evidential threshold policies. Via self-sorting and algorithmic prioritization, agents tend to find themselves in communities of people with shared social identities, or communities that construct new identities. Because engagement-maximizing content tends to be promoted, and content connected to an agent's identity tends to be highly engaging, users are exposed to a lot if identitarian content. As a result, relevant social identities become more central to
individuals, and also connected to a wider range of topics. This leads to an expansion of the depth and breadth of identity-protective reasoning: agents will be more resistant to evidence for a wider range of beliefs. In other words, they will raise their evidential thresholds for many beliefs, a change to the shape of their epistemic agency.

These internet-mediated changes to epistemic style are troubling. First, I argue that these changes often facilitate the acceptance of misinformation and its evidence-resistance once accepted. In conjunction with misinformation that is tailored to the agent's newly developed epistemic preferences, this will lead to a deterioration in the quality of agents' belief sets. Second, I argue that these changes can compromise our epistemic autonomy. This is because we acquire these new epistemic styles in a way that bypasses reflection and our own pre-existing goals as epistemic agents. In this way, we risk becoming a mere vehicle for our environment, instead of autonomous epistemic agents whose interactions with information are driven by our goals.

Finally, I take a step back and consider implications of this discussion for the emerging field of hostile epistemology, the study of how our environment can exploit our cognitive vulnerabilities and weaknesses to lead us to false beliefs and worldviews (Nguyen 2023). My discussion illustrates that doing justice to the role of our environments also requires attending to how environments *can sculpt epistemic agency itself,* not just how environments exploit agents with fixed traits. Specifically, bad actors can set up environments to create new vulnerabilities in agents, which

can in turn be exploited to lead to false beliefs. If we are going to understand and address epistemic pathology, we also need to consider agency-molding features of environments, as I do in this project.

References

Flores, Carolina (2021). Epistemic Styles. *Philosophical Topics* 49 (2):35-55.

Rachel Fraser (2020). Epistemic FOMO. *The Cambridge Humanities Review*.

Kahan, Dan M., Misconceptions, Misinformation, and the Logic of Identity-Protective Cognition (May 24, 2017). *Cultural Cognition Project Working Paper Series* No. 164, Yale Law School.

Nguyen, C. Thi (2020). *Games: Agency as Art*. New York: Oxford University Press.

Nguyen, C. Thi (2023). Hostile Epistemology. *Social Philosophy Today* 39:9-32.

O'Connor, Cailin & Weatherall, James Owen (2019). *The Misinformation Age: How False Beliefs Spread*. New Haven, CT, USA: Yale University Press.

Paul, Sarah K. & Morton, Jennifer M. (2018). Grit. *Ethics* 129 (2):175-203.

João Fonseca

NOVA University Lisbon

A Proof of '1st/3rd Person Relativism'

The suggestion of something like a 'relativist solution to the Mind-Body problem' has recently been held by some scientists and philosophers; either explicitly (Galadí, 2023; Lahav & Neemeh, 2022; Ludwig, 2015) or in more implicit terms (Solms, 2018; Velmans, 2002, 2008). In this talk I provide an argument in favor of a relativist approach to the Mind-Body problem, more specifically, an argument for '1st/3rd person relativism', the claim that 'The truth value of some sentences or propositions is relative to 1st and 3rd person perspectives'.

The argument for 1st/3rd person relativism is akin to a formal or mathematical proof. It is shown that, just by assuming the 1st/3rd person distinction itself and using first order logic and set theory, '1st/3rd person relativism' follows as *a theorem*. The premises of the proof follow from the acceptance of a possible (indeed, *physically* possible) scenario where our corner of the universe is perfectly replicated in another part of the universe. The inhabitants of 'our' planet earth and the 'replica' are exactly the same with the exact same 'physical' and 'mental' lives. If we pick one such inhabitant, say 'Susan', there is one exact copy of her (in all respects of her physical and mental lives) on the other planet. We can devise a temporal ordered set of conscious mental states A corresponding to the conscious 'mental life' of both 'Susans' in the two planets: $A = \{M'(t1), M''(t2), M'''(t3),\ldots, M(tn)\}$ where 'M' stands for a specific conscious mental state. Now we can form another ordered set, just like A but substituting the mental predicates (with the form 'M(tn)') with open sentences with the general form 'Mx(tn)' (stating: 'x is in mental state M at time tn') thus getting set $B = \{M'x(t1), M''x(t2), M'''x(t3),\ldots, Mx(tn)\}$. Now, given this scenario, take the following question: 'How many ordered sets can we form by satisfying all open sentences of set B with a single object?' (in a sense, this question equates to asking: 'how many objects satisfy *all* open sentences of B?'). In order to provide an answer, two operational definitions of 1st and 3rd person perspectives are proposed for clarity and rigor's sake:

> **Def₁**: Given an object O, there is a set of properties (P_1, P_2,...,P_n) of O that exists from a *3rd person's perspective iff* those properties of O are *objectively and publicly determined*

> **Def₂**: Given an object O, there is a set of properties (P'_1, P'_2,...,P'_n) of O that exists from a *1st person's perspective* to a subject S *iff S subjectively experiences* those properties of O

Assuming the 1st/3rd person distinction in terms of **Def₁** and **Def₂**, it can be shown (through first order logic and set theory) that the answer to the question regarding

the number of sets *differs* if it is answered from a 1st or a 3rd person perspective. The 1st person perspective answers 'There is *one* set' whereas the 3rd person answers 'There are *two* sets'. A proof (by *reductio ad absurdum*) of '1st/3rd Person Relativism' easily follows: '1st/3rd Person Relativism' states: 'The determination of the truth value of some sentences or propositions is relative to 1st and 3rd person perspectives'. Now, assume the *negation* of this claim, i.e., that '1st/3rd Person Relativism' is *false*. If '1st/3rd Person Relativism' is false, we must assume that the sentences: 'There is one ordered set of mental states' *and* 'There are two ordered sets of mental states' are *both true* (or are both derivable). But, of course, this is an *inconsistent result*; we can very easily derive obvious contradictions like 'there are two ordered sets *and* there are *not* two ordered sets'. So, we must *negate* the *falsity* of '1st/3rd Person Relativism'. Therefore, '1st/3rd Person Relativism' *is true*.

The talk presents a demonstration of '1st/3rd person relativism' to some detail. It is also shown that: *1)-* the question being posed, concerning the number of ordered sets constructed from set *B, can* be answered *both* by 1st and 3rd person perspectives, *2)-* 1st and 3rd person perspectives have the *same* 'epistemic credibility' (in particular, the truth of 1st person claims is *not* dependent on introspection), and *3)-* neither perspective 'knows more than the other' to the point of precluding the other perspective's answer.

The talk concludes with a brief evaluation of some consequences of '1st/3rd person relativism' to the Mind-Body Problem (in particular to the Hard Problem of consciousness). It is shown that '1st/3rd person relativism' *predicts* the existence of an (apparent) Explanatory Gap; *explains why* the Explanatory Gap is just apparent (and the origins of such illusion); *dissolves* the Hard-Problem; provides a *possible solution* the problem of Mental Causation; *explains why* Mental Causation looks like a problem in the first place and accurately *predicts* the actual empirically found *correlation* and *covariation* between conscious experiences and brain states. This explanatory power of '1st/3rd person relativism' is particularly impressive since it was not designed as a possible solution to the Mind-Body problem in the first place.

References

Galadí, J., (2023) "What I am and What I am Not; Destruction of the Mind-Body Problem", *Philosophies*, 8: 101

Lahav, N. & Neemeh, Z (2022) "A Relativistic Theory of Consciousness" *Frontiers in Psychology*, 12: 704270

Ludwig, D. (2015) *A pluralist Theory of the Mind*, European Studies in Philosophy of Science, Springer

Solms, M. (2018) 'The Hard Problem of Consciousness and the Free Energy Principle', *Frontiers in Psychology*; 9: 2714.

Velmans, M. (2002) 'How Could Conscious Experience Affect Brains?'. *Journal of Consciousness Studies*, 9(11), 3-29

—— (2008) "Reflexive Monism", *Journal of Consciousness Studies*, 15(2): 5–50.

Manuel García-Carpintero

University of Barcelona

Fictionalism about Fictional Characters (Reloaded)

In previous work (GC 2010, 2021a, 2023) I offered a fictionalist account of apparent reference to fictional characters in fictional discourse. Here I want to set it in contrast with deflationary realist accounts like Thomasson's (2014). I will present in my own terms how I think best to understand the deflationary realist accounts I target, my alternative fictionalist proposal, and finally I'll outline the reasons to prefer it that I'll develop in my presentation.

Realists about fictional characters argue that a proper account of claims in (1) below requires interpreting names like 'Gregor Samsa' in them as referring to fictional entities:

(1)   Gregor Samsa is a fictional character.

Artifactualists like Salmon (1998), Thomasson (1999, 2003) or Schiffer (2003) think of them as having an ontological status analogous to that of the fictional works in which they occur (Thomasson 1999, 143; 2003, 220; Salmon 1998, 78-9). In contrast, my brand of fictionalism about fictional characters claims that we don't need to take referential expressions in such discourses to really refer to them. But I don't have any ontological scruples about *ficta* understood along the lines of the artifactualist proposal. In fact, to assume them is helpful to present my own view. Let me elaborate on how I understand the realist proposal.

Fictional works result in my view from the communicative acts of fiction-makers; they are social constructs, abstract created artefacts with norm-regulated functions. They have a complex structure, grounded on that of the vehicles that express them; they are in part composed of singular representations (GC 2019). It is these singular representations what I'll take fictional characters to be on the artifactualist account I'll use as a foil: terms like 'Gregor Samsa' have as semantic value a singular representation associated with that very name, which is a constituent of Kafka's *Metamorphosis*.

I'll provide a few more details on how I suggest thinking of the semantic values of *prima facie* empty names on the just outlined realist proposals. I assume a view standard in current semantics (GC 2021b). On this view, referential expressions like indexicals and proper names carry *presuppositions of acquaintance*, or *familiarity*. This is to be cashed out by assuming that contexts include *discourse referents*, which we may think of as shareable singular representations that may well not pick out anything. For proper names, the relevant discourse referents are crucially defined by naming practices (distinct ones for the 'David' that picks out Lewis and the one that picks out Hume, cf. REDACTED); typically existing ones, but in some

cases practices created with the very discourse including the name. For indexicals, they might be constituted by perceptual information, or by information available in previous discourse to which the expression is anaphorically linked.

The singular representations that fictional realism takes to be the semantic value of referential expressions like 'Gregor Samsa' are thus to be individuated by such discourse referents. There is a tradition that associates some descriptions with entities of the kind we are positing, *roles* (see Glavaničová 2021, Stokke 2021) like *the president of the USA* or *the mayor*. Roles can be understood as Carnapian individual concepts picking out their occupiers relative to worlds, to the extent that we think of them as merely partial functions. If we model the fictional world by means of standard possible worlds, the role that we are taking as the semantic value of 'Gregor Samsa' will pick out different individuals in different such worlds.

As said, I agree that we are committed to representational entities of the kind just outlined, but I reject that 'Gregor Samsa' in (1) refers to one of them. On the view of such discourse I hold (a version of Yablo's (2001) *figuralist* brand of fictionalism), the semantic referential apparatus (*de jure* directly referential expressions such as names and indexicals, quantifiers generalizing over the positions they occupy, expressions for identity) is used metaphorically in the likes of 'Gregor Samsa is a fictional character', deploying the figure of speech called *hypostatization* (REDACTED). It is a rather dead, conventionalized kind of metaphor, so, in contrast to pretense-theoretic fictionalist proposals (Walton 1990, Everett 2013), on this view utterances in metatextual discourse are straightforward assertions with truth-conditions.

This might suggest that the view is after all realist, committed to referents of some sort for singular terms in metatextual discourse, but it is not. In support of this, I follow Yablo's (2014) recent development of his views, on which the truth of metatextual sentences including fictional names and their generalizations do not really commit us to the existence of fictional characters. For this is merely pretend-presupposed and, when we look at what they are really *about* (the *truth-makers* for the claims we make with them), we do not find referents that they pick out. What we find in all is a fiction like *Metamorphosis*, referentially deploying 'Gregor Samsa'. This account thus has some similarities with artifactualist views; it relies on similar truth-makers for the likes of (1). We end up interpreting them as making genuine assertions, whose truth is grounded on the pretenses thereof in the relevant fictional work. But it doesn't need to assign a referent to fictional singular terms like 'Gregor Samsa'.

Why not endorse the deflationary realist proposal? This is not needed on the view just outlined, and there are good reasons against it. The ones I will develop are metaphilosophical in nature. If we take it seriously that fictional names refer, we have to find reasons to choose among the different candidates that realists offer for the role of *ficta* – Platonistic *abstracta*, Meinongian nonexistents, or what have you, in addition to artifacts; but there is no rational way of taking a stance on such matters, I'll argue. For all we can tell, the relevant empirical and theoretical considerations we can deploy do not select just one of them, even if they do allow us to dispose of some. The fictionalist attitude authorizes us to ignore the issue in good faith.

References

Everett, Anthony (2013): *The Nonexistent*, Oxford: Oxford University Press.

García-Carpintero, Manuel (2010): 'Fictional Entities, Theoretical Models and Figurative Truth', in Frigg, R, and Hunter, M. (eds.), *Beyond Mimesis and Convention – Representation in Art and Science*, Springer, 139-68.

García-Carpintero, Manuel (2019): "Normative Fiction-Making and the World of the Fiction", *Journal of Aesthetics and Art Criticism*, 77 (3), 267-279.

García-Carpintero, Manuel (2021a): "Models as Hypostatizations", in Alejandro Cassini & Juan Redmond (eds.): *Models and Idealizations in Science: Fictional and Artifactual Approaches*, Cham: Springer, 179-197.

García-Carpintero, Manuel (2021b): "Reference-fixing and Presuppositions", in S. Biggs & H. Geirsson (eds.), *Routledge Handbook on Linguistic Reference*, London: Routledge, 179-198.

García-Carpintero, Manuel (2023): "Semantics of Fiction", *Mind and Language* 38 (2), 604-618.

Glavaničová, Daniela (2021): "Rethinking Role Realism", *British Journal of Aesthetics* 61(1), 59-74.

Salmon, Nathan (1998): "Nonexistence," *Noûs* 32, pp. 277-319.

Schiffer, Stephen (2003): *The Things We Mean*, Oxford: Clarendon Press.

Stokke, Andreas (2021): "Fictional names and individual concepts", *Synthese* 198, 7829–7859.

Thomasson, Amie L. (1999): *Fiction and Metaphysics*, Cambridge: Cambridge U.P.

Thomasson, Amie L. (2003): "Speaking of fictional characters", *Dialectica* 57: 205-23.

Thomasson, Amie L. (2014): "Public artifacts, intentions and norms", in M. Franssen, P. Kroes, T. Reydon, P. Vermaas, *et al.* (eds.), *Artefact kinds: Ontology and the human-made world*, Dordrecht: Springer.

Walton, Kendall (1990): *Mimesis and Make-Believe*. Cambridge, Mass.: Harvard U.P.

Yablo, Stephen (2001): "Go Figure: A Path through Fictionalism", Peter A. French & Howard K. Wettstein, *Midwest Studies in Philosophy* xxv, Oxford: Blackwell, 72-102.

Modesto Gómez-Alonso

University of Seville

Self-Determination and Rational Agency

Standard works in analytic epistemology are mainly concerned with justification, belief, representation, truth, and knowledge; namely, with phenomena pertaining to the domain of theoretical rationality. However, Telic Virtue Epistemology (TVE) has recently come to provide a performative account of knowledge, one which aims at explaining knowledge in terms of intentional relations, means/end rationality, attempts, competences, and apt success. On this view, being able to project and realize one's ends is a necessary condition for coming to know. It is true that TVE focuses on one's *cognitive* ends, and that their defenders make a categorical distinction between epistemic and practical ends. And yet, knowledge being not passive observation, TVE has been gradually led to acknowledge that epistemology should be embedded in a broader theory of *rational, purposive action*, one which would include both cognitive and practical rationality, as well as an account of how they are related. Additionally, and whatever may be the right view about how epistemic and practical aims are related, virtue epistemologists seem logically obliged to claim that beings deprived of practical ends would not be able to represent (truly or falsely): they would not be agents; they would lack the *intentionality of consciousness* which is the common root for representation and desire.

In this talk, I discuss self-awareness with an eye toward the function it plays within a performative account of knowledge. However, it is not only that by following the path opened by TVE, this talk aims at providing a performative, fully normative account of self- knowledge, so that the general structure of rational action which permeates all varieties of knowledge comes also to be applied to knowledge of oneself. It is also that self-activity (self- determination) is at the center of rational action, and so, that I-hood lies at the basis of the 'performative turn' in epistemology. Selfhood comes thus to be displayed as the *eliciting* act of intentionality, namely as the orientation of the mind towards itself which is required for the mind to be able to direct itself to something. Now, it is the nature of this eliciting act that has to be elucidated.

Capturing the epistemic structure of subjectivity cannot be the end of analysis. Far from it, it must be just its starting point. An epistemic approach to subjectivity leaves, at the very least, two problems unsolved.

On the one hand, if the core of subjectivity were merely a self-contained epistemic relation, self-knowledge would be epistemically idle, being thus irrelevant for making sense of the dynamics of cognition. On this view, the paradigm of

knowledge would (paradoxically) suffer from the greatest normative poverty, and self-consciousness would underlie experience without grounding experience. In other words: self-consciousness cannot be a mere explanatory reason for consciousness —a mere imposition on our passivity; it must rather sustain a basing relation so that acts of intentionality derive from it.

On the other hand, it is unclear how self-knowledge might be able to capture agency, and thus, how it might really refer, not to a given thing, but rather to a rational agent. What the epistemic view of the self, if unframed, attempts to capture (the heart of subjectivity, meaning self-activity) remains elusive. On this account, self-knowledge and action would always remain disconnected.

Call the first issue, the *grounding problem* for the *cogito*. Call the second issue, the *spectatorial, third-person-perspective problem* for self-knowledge. As it will be argued, those two problems are intrinsically related. What they suggest is that self-awareness, even if it is (as it is) conceptually tied to subjectivity, may not be logically prior for understanding the nature of selfhood.

The active nature of subjectivity is not, however, easy to be philosophically apprehended. This is why I have opted for a dialectical, indirect approach to the question; one which raises three (at first sight, hopeless) objections to the epistemological conception of subjectivity. Those objections lead us to the skeptical thought that self-knowledge is entirely unfounded, which, in turn, results in a logical conflict between basic principles of thought —a contradiction which is unlivable. The lesson to learn is that to avoid refutation, the structure of subjectivity cannot be primarily epistemic. Call this conclusion, which will be argued in detail, *the spandrel model of self-awareness*. Call the first objection, the *trivialization of self-knowledge*; while the second and the third objections are, respectively, the spectatorial problem for self-knowledge, and the grounding problem, which were previously mentioned.

The main problem lies, however, at the very heart of performative accounts of knowledge, hidden back in the notion of 'rational agency.' Actions, for being such, must be purposive, which means that they must be normatively structured and constrained by their aims. If, however, subjectivity is in its nature the power of the mind to initiate intentional, purposive, aim-guided acts (including representations), and thus, it is no other than the mind's capacity for being freely self-determined; then it seems obliged to conclude that rational acts at bottom are not rationally structured, and so, that normativity does not go all the way back to the foundations of rational action. We thus find radical contingency at the core of rational agency: rationality is not automatically given; *that* one represents an object (any object), and *how* one represents it, are free acts which, as such, require possible alternatives; even in the case of epistemic attempts, *libertarian freedom* seems to swamp the truth norm for belief.

Call the question here raised, the problem of *epistemic constructivism* for performative accounts of rational action. It is the problem to which this talk is mainly directed. Libertarian, unpremised freedom would be, if logically consistent, a mere brute act. Here I will take the opposite view: rational agency is not, and can't be an object of choice among others. Which means that my goal is not simply that of

mitigating the radical contingency of knowledge. I rather aim at showing that normative constraints and free self- determination are at bottom one in essence.

James Grayot
University of Porto

Why be a frame-sensitive reasoner? On the rational requirements of ultraintensional framing

In *Frame it Again* (FIA), J.L. Bermúdez defends the rationality of what he calls "frame-sensitive reasoning" (2020). The defense runs as follows: In contrast to mainstream perspectives about rationality in economics and decision theory, there are many contexts in which it is *permissible* for an individual to be sensitive to different framings of a decision and, thereby, to hold "quasi-cyclical" preferences. Unlike cyclical preferences, which are paradigmatic of irrationality, quasi-cyclical preferences are those in which the decision-maker would not change or alter their preference orderings upon learning that two options/outcomes share the same extension, i.e., refer to the same things. More specifically, quasi-cyclical preferences are permissible if/when they occur within an 'ultraintensional' context: this is when a decision-maker *knowingly* ascribes different valuations to the same option or outcome under different perceptual or semantic framings. Bermúdez further argues that there are also many contexts in which one *ought to be* a frame-sensitive reasoner. In scenarios of great social or political complexity, where there is no single frame of reference to interpret the options/outcomes of a decision, one ought to strive to consider as many (relevant) frames as possible. The more flexible one is in considering different frames, the more likely they will be to reveal salient emotions or values that bias their judgments (good or bad) toward a decision.

While we are sympathetic toward the view that frame-sensitive reasoning can lead to rational action, we remain skeptical about its theoretical underpinnings and justifications. In short, our paper asks: *Why be a frame-sensitive reasoner? –What establishes the rationality of frame-sensitivity?* Starting from the assumption that Bermúdez is primarily interested in the normative virtues of framing, FIA can be seen as having two overlapping goals: the first is to defend the permissibility of quasi-cyclical preferences; the second is to defend the rationality of frame-sensitive decision-making. The two are clearly related, but their justifications differ. Establishing the first goal depends on demonstrating that ultraintensionality, as a cognitive criterion, justifies holding quasi-cyclical preferences. Assuming that agents can and do engage in ultraintensional reasoning, we take it to be relatively straightforward that quasi-cyclical preferences are sometimes permissible. Establishing the second goal, however, depends on further showing that frame-sensitivity is not merely permissible but that individuals *ought to* be frame-sensitive

reasoners if they want to be rational in a broad range of real-world scenarios. The challenge for this second goal is that, by moving away from classical decision-theoretic contexts (i.e., 'small worlds') the justification for being a frame-sensitive reasoner becomes case-specific and opens itself up to questions and criticisms beyond those found in decision theory and philosophy of action.

Though Bermúdez is silent on what ultimately determines the rationality of one's actions, he identifies four components of what he calls 'due diligence' that are integral to the reasoned pursuit of one's actions (these are reflexive decentering, imaginative simulation, perspectival flexibility, and reason construction). But this suggests that the theory of frame-sensitive reasoning is not per se about rational action; rather, it is a theory about the benefits and appropriate applications of *rational deliberation*. Bermúdez muses that proper understanding and execution of the latter can help us to achieve the former.

Yet, many questions remain concerning what counts as an ultraintensional context. For instance, for a context to count as ultraintensional, must agents be (cognitively) *aware* that their double-framing causes quasi-cyclical preferences?

→ If **yes**, what cognitive capacities must we assume that epistemic agents have? And, what theory of cognition must one adopt in order to defend the idea that quasi-cyclical preferences can be epistemically virtuous?

→ If **no**, does this drive a wedge between the *mere permissibility* of quasi-cyclical preferences and the *practical rationality* of frame-sensitive reasoning in complex situations? Would it rule out simpler cases of quasi-cyclical preferences from being rational if the decision-maker lacks or fails to exercise the cognitive virtues needed to engage in ultraintensional deliberation?

Given the emphasis on rational deliberation, we approach these questions from the dual perspectives of virtue epistemology and philosophical psychology. Virtue epistemology holds that to know something is to know it in the right way, to have acquired it in virtue of good epistemic practices, such as *attentiveness*, *fair-mindedness*, *open-mindedness*, *intellectual tenacity*, etc. As such, we believe that an epistemic defense is, currently, the most plausible way to ground the rationality of frame-sensitive reasoning. And so, with respect to the titular question, *why be a frame-sensitive reasoner?* –the answer is that it not only inspires and instills cognitive virtues that enhance reasoning procedures, but also that, if executed in the right way, these virtues will serve as a filter for which actions an individual should pursue, thus avoiding the pursuit of non-rational ends.

But taking the epistemic defense seriously opens frame-sensitive reasoning up to important criticisms that have been leveraged against virtue epistemology. One formidable example is the *situationist critique*, which argues that individuals don't have stable character traits, and so, virtue is not realistic for guaranteeing pursuit or achievement of the truth. The situationist critique is important because, unlike reliabilist forms of virtue epistemology (which are only concerned with basic cognitive abilities, e.g., perceptual acuity, simple inductive and deductive ability), it takes aim

at higher-level reasoning capacities which require epistemic agents to have a mature and stable character.

In pursuing an epistemic defense of frame-sensitive reasoning, our goal is (i) to identify the extent to which the situationist critique holds, and (ii) to positively restate the requirements of ultraintensionality given some advances in cognitive science and psychology of human reasoning. If successful, this would help establish extra-normative principles for the rationality of frame-sensitive reasoning.

Margherita Grassi
University of Barcelona

Characterizing misogyny: a new approach

Misogyny has been traditionally understood as an emotion of hatred against women generally, or women *qua* women. Therefore, according to this definition, misogyny seems to be a property of *individuals* who hate women. *Contra* this view, Kate Manne (2018) characterized the concept of misogyny as a form of structural oppression aimed at maintaining, or restoring, the patriarchal order. Manne highlighted the *structural* facet defining 'misogyny' as a sanction distributing mechanism that punishes women who violate gender norms and rewards women who comply with them. I consider both conceptions insightful because they allow us to understand how misogyny could operate in our society. However, they both capture some aspects of the phenomenon and overlook others. In particular, I want to point out the risk of *supplanting misogyny with its expression*: too much attention is paid to the manifestations of misogyny throughout society, instead of focusing on its roots, and how it lives and survives socially. The concept of 'misogyny' captures a complex phenomenon: (i) it is important to emphasise what allows misogynistic actions to succeed structurally, but (ii) it is also crucial not to overlook how education and socialisation influence individuals' attitudes and behaviour permitting misogyny to exist and proliferate.

In academia, the problem of defining 'misogyny' has rarely been addressed directly. While previous authors had mostly dealt with oppression or related phenomena (among others, Dworkin 1974, Frye 1983, Young 1990, Cudd 2006), Manne critically addressed the ordinary definition of misogyny – hatred towards women – proposing an *ameliorated* one focused on what misogyny *does to* women. Her approach changed the debate on the issue because she decided, using the tools of conceptual engineering, to sharpen the term and make it easier to apply by de-psychologising it, de-individualising it, and making it epistemically accessible. In this way, she was successful in highlighting the *structural* facet of the phenomenon and in giving more political relevance to it. Moreover, she proposed to understand misogyny as a phenomenon that *does not end* in hostility: women can receive rewards and recognition if they submit to the patriarchal system in which they live, following the imposed gender norms. On the other hand, several criticisms have been raised about parts of her reasoning. Commentators have shown how certain deficiencies in the ordinary notion are not so decisive as to justify its outright rejection. Some elements of the common definition could be useful to understand what misogyny is, even without denying the need for conceptual redefinition (e.g., Simion 2021, Mikkola 2019). In particular, Samantha Pinson Wrisley (2023) critiqued Manne's proposal suggesting to recover the link between the term 'misogyny' and its affective dimension: misogynistic hatred is not a simple feeling of hate towards

women, it is a more structured and complex form of aversion. According to Pinson Wrisley, Manne has circumvented the challenge of defining what kind of hatred misogyny is by shifting the focus to the phenomenon's manifestations, that is by conflating misogyny with gendered violence. While I do not agree that (i) Manne's proposal is incompatible with recognising the affective facets of the issue (Lowen Walker 2022) and that (ii) Manne fully conflates the phenomenon with gender-based violence, I hold that Manne's work does run the risk of merging the phenomenon with its expressions. I argue, however, that characterising misogyny as hatred towards women also runs this risk.

Both characterizations can be useful in reframing the meaning of 'misogyny'. While the ordinary definition seems intended to capture how misogyny is individually experienced in an emotional key. Manne's conception primarily underlines the systematic nature of misogyny, illuminating especially how the patriarchal system reacts to its transgressions. At the same time, both definitions do not seem to capture the phenomenon in its entirety. The first only assesses the individual dimension, failing to explain how the phenomenon manages to be so socially influential. The second seems mainly concerned with structural actions and almost completely neglects the role of individual behaviour in perpetuating misogyny. I will suggest the need to broaden these views by considering what influence this phenomenon has on our way of reasoning, evaluating, and acting. It seems that certain forms of misogyny can only be explained by the way genders are socialised and the role that certain sets of social meanings and practices (Haslanger 2012, 2017, 2019) play in this process. Misogyny can manifest itself as a sanctioning mechanism, or a feeling of hatred, but there seems to be something beyond that. This approach seems promising in fixing many limitations of previous formulations of 'misogyny', helping us to understand the issue more fully.

References

Anderson, K.J. (2014). *Modern Misogyny: Anti-Feminism in a Post-Feminist Era*. Oxford: Oxford University Press.

Cudd, A. (2006). *Analysing Oppression*, Oxford, Oxford University Press.

Dworkin, A. (1974). *Woman Hating*. Plume Book, Penguin Books USA.

Frye, M. (1983). *The Politics of Reality: Essays in Feminist Theory*. The Crossing Press, Freedom, California.

Gilmore, D. (2018). *Misogyny: The Male Malady*. Philadelphia, PA: University of Pennsylvania Press.

Glick, P. and Fiske, S. (2001). An ambivalent alliance. Hostile and benevolent sexism as complementary justifications for gender inequality. *American Psychological Association*, 56(2), pp.109-18.

Haslanger, S. (2012). *Resisting Reality*. Oxford, Oxford University Press.

Haslanger, S. (2017). *Critical Theory and Practice*. Koninklijke Van Gorcum.

Haslanger, S. (2019). "Cognition as a Social Skill". *Australasian Philosophical Review*. n.3, 1: pp. 5-25. Loewen Walker, R. (2022). Call it misogyny. *Feminist Theory*.

Manne, K. (2018). *Down Girl*. Oxford: Oxford University Press.

Mikkola, M. (2019). Down Girl: The Logic of Misogyny, by Kate Manne. *Australasian Journal of Philosophy*, *97*(1), 198–201.

Pinson Wrisley, Samantha ([2021] 2023) "Feminist theory and the problem of misogyny". *Feminist Theory* 24 (2): pp. 188-207.

Simion, M. (2021). "Book reviews: Down Girl: The Logic of Misogyny by Kate Manne". *The Philosophical Quarterly*, vol. 71/3. pp. 681-684.

Young, I.M. (1990). *Justice and the Politics of Difference*, Princeton

Hugo Heagren

King's College London

The Conditions of Rigidity

Proper names are rigid designators—they designate the same thing in all possible worlds. Definite descriptions are non-rigid designators—they (often) designate different things in different possible worlds. At least two difficult questions emerge in the literature: 1) are there good notions of rigidity for expressions besides designators? 2) If so, which such expressions are rigid and which are not?

Answers to one question have generally presupposed answers to the other. Schwartz (2002) conflates rigidity with (rigid) designation, and thus concludes no other terms are rigid. Soames (2002) and Devitt (2005) both assume that if there *were* a good notion of rigidity for general or kind terms, only natural kind terms would be rigid, and use this to argue against such proposed extensions of rigidity. I present an account which presumes answers to neither question but can help answer both.

Before the arguments of the paper, I consider and reject the claim that a notion for kind terms of rigidity should, by definition, distinguish the natural from the non-natural kind terms. This would justify the assumptions made above, and all three authors rely on a principle like this. I argue this is not consonant with the original philosophical uses of rigidity: it was a *discovery* that proper names are rigid designators and (most) definite descriptions are not—rigidity was never supposed to be the way to distinguish the names from the descriptions. We find this in early discussion of rigidity by Marcus (1961) and Kripke (1980).

In the paper I develop two significant claims. First, I take the two widely agreed upon paradigm cases of proper names and definite descriptions, and develop an account of rigidity which confirms standard views about their status, while not assuming answers to the questions above. I begin by cleaning up Kripke's definitions to make clear the differing roles of the world of utterance and the world under discussion. I then distinguish de jure from de facto rigidity, by appeal to *facts* or *states of affairs*, which distinguish possible worlds from each other. This can also distinguish de jure from de facto rigidity: de jure rigid designators have their reference fixed *only* by states of affairs in the actual world, whereas de facto rigid designators do not, but are rigid all the same.

Second, I demonstrate how accounts of rigidity in general (mine or others) can be extended to non-designator terms. I argue that the class of designators is delimited by its function—*designating*—and that this makes it clear what is held fixed across possible worlds for members of that class to be rigid—in this case *designation*. Thus rigidity can only be profitably understood for other classes which are similarly delimited by function, such as predicate and logical terms. I discuss what would have to be held fixed to profitably understand rigidity for various other types of term. Conversely, one term with multiple functions might be rigid with respect to one and but not another. Indeed, definite descriptions are a good example: the designation of

a definite description like 'the president of the US' might change depending on which world is being discussed (so the description is not a rigid designator), but there *is* a sense in which its meaning is constant—in every world it picks out (whoever is) *the president of the US*. Such descriptions might be called *rigid describers*.

Finally, I take it that kind terms can function as predicators, and offer an analysis of rigidity for them in these terms. I show that the vast majority of kind terms are predicatively rigid, but there are some which are not—those whose satisfaction conditions depend on facts about the world under discussion, independent of the thing they are predicated *of*. Examples include *being rare* and *being common*. I also briefly discuss what de facto rigidity for predicators would look like, and consider some possible examples of unusual predicates from modal logic.

References

Devitt, Michael (2005), "Rigid Application", in *Philosophical Studies* 125, pp. 139–165.

Kripke, Saul (1980), *Naming and Necessity, Lectures Given to the Princeton University Philosophy Colloquium*, Harvard University Press.

Marcus, Ruth Barcan (1961), "Modalities and intensional languages", in *Synthese* 13, pp. 303–322.

Schwartz, Stephen P. (2002), "Kinds, general terms, and rigidity: A reply to LaPorte", in *Philosophical Studies* 109, pp. 265–277.

Soames, Scott (2002), *Beyond Rigidity: The Unfinished Semantic Agenda of Naming and Necessity*, Oxford University Press.

André Kfouri
NOVA University Lisbon

Procedural and Practical Knowledge

The concept of procedural knowledge, derived from the cognitive sciences, has been used in John Greco's The Transmission of Knowledge to accommodate Wittgensteinian hinge commitments as knowledge, and therefore as part of a broader framework of knowledge generation and transmission. There, procedural knowledge is defined broadly as "knowledge exercised in the performance of some task", that can be implemented in various ways according to how explicitly or implicitly it is held. Given the limited purpose of employing this concept, Greco remains intentionally neutral in regards to the many historical and modern philosophical questions regarding concepts of practical knowledge, though he cautiously suggests it might be identified with "knowledge how". This neutrality, however, leaves space for further considerations about the relationship between epistemology and the philosophy of action, and begs the question of how other accounts of practical knowledge might fit into his framework of transmission and generation.

The main object of this talk will be the comparison of Greco's account of procedural knowledge to Elizabeth Anscombe's account of practical knowledge, notable not only for its situation in her *Intention*, a landmark text in the philosophy of action, but also for taking center stage in that work, structuring her discussions and definitions around the intention with which something is done. For all the importance of this concept of knowledge in Anscombe's work, she is more concerned with its implications in action than in epistemology, and doesn't describe how it might interact with other sources to form an agent's full body of knowledge. My suggestion, then, is that a comparative look at Greco's and Anscombe's concepts of practical knowledge may be of mutual benefit to both accounts, providing the former with a more detailed notion of action and of acting as a source of knowledge, and the latter with a more thorough framework of how practical knowledge relates to other kinds.

I will present Anscombe's characterization of practical knowledge as the knowledge an agent has of his own action, subject to a number of epistemological particularities and privileges. In her account, it constitutes a mode of knowing distinct from observation even though both may have the same object, that is, that action as an event in the world, and may naturally overlap when the agent can see the results of his action. Nonetheless, practical knowledge is subject to a different kind of fallibility, in that, if it doesn't match the situation in the world, "the mistake is one of performance, not judgment", a maxim Anscombe adopts from the *Magna Moralia*, which she attributes to Theophrastus. When there is no such mistake, practical knowledge is "the cause of what it understands", as characterized by Aquinas, in

contrast to speculative, contemplative or theoretical knowledge, which is derived from what it knows. Finally, practical knowledge is also characterized by being structured in a means to ends order that Anscombe sees as characteristic of the description of any action as intentional, as she famously represents with chains of "why?" questions.

Then, we may consider a recent interpretation of Anscombe that seeks to clarify the relations between kinds of knowledge. John Schwenkler's *Anscombe's Intention: A Guide* attempts to track the historical roots of Anscombe's practical knowledge to a tripartite classification derived from Aquinas's commentary on *De Anima*. In Schwenkler's reconstruction of Aquinas, a piece of knowledge may be considered either practical or theoretical in three respects: the object, which is practical when it's a thing "producible by the knower"; the mode, which is practical when the producible thing is considered "as producible"; and the end, which is practical when directed towards the realization by action of what that knowledge represents. He argues that this classification is implicit in Anscombe's work, and that her concept of practical knowledge is simply knowledge that is practical according to all three of Aquinas's respects. I will argue that Schwenkler's Thomistic classification is not directly applicable as an interpretation of Anscombe. First, as Anscombe herself argues, the Aristotelian definition of a practical object of knowledge is far too broad, including everything that is contingent and that cannot be known by demonstration. Of course, it is very possible and common to reason about something that isn't necessary without having a view to action, and for theoretical knowledge to play a role in practical reasoning. And second, the distinction between a practical mode and ends of knowledge is also redundant, since the form of practical reasoning for Anscombe is a structured order of means to ends whose conclusion is an action, and practical knowledge is the knowledge the agent has of that action. Idle considerations about someone else's justifications for acting, or of Aristotle's 'classroom examples' of practical syllogisms, would not be considered by Anscombe to be a form of practical reasoning in any way, or conducive to practical knowledge.

One last problem may be posed about Anscombe's practical knowledge, concerning its continuation over time. Though it enjoys those peculiarities while the agent engages in the action, it would seem to become knowledge like any other when the full course of action is done: if someone misremembers something they did in the past, the mistake is surely already of judgment. I will conclude that Anscombe's practical knowledge is best thought of as a distinct *source* of knowledge, that may have a particular and privileged epistemic status during the performance of the action, but later mixes with knowledge from other sources, including testimony as well as observation, to form "a complicated network of received information", as Anscombe says in her paper on testimony "What is it to believe someone?". This network, I will argue, can be thought of in terms of Greco's framework, and in particular of his metaphor of a knowledge economy. Finally, in considering the possible integration of Anscombe's practical knowledge into Greco's framework, I will try to assess whether coming to have practical knowledge could be considered a case of generation comparable to other generative sources, or whether it constitutes another irreducible category of knowledge.

Samuel Laves
NOVA University Lisbon

A conversion of a special kind: changing worldviews and bridging deep disagreements

There have been numerous articles in recent years discussing a potential application of Wittgenstein's epistemological arguments in *On Certainty* (Wittgenstein 1969) to moral epistemology. Most of these have been focused on the topic of universal moral certainty and generally discuss one example. Comparatively less has been written on the topic of local moral certainty. This is despite the fact that, it could be argued, the phenomenon of local moral certainty leads to more difficult and interesting philosophical dilemmas. One such problem is that of deep disagreement, defined first by Fogelin as disagreements that "cannot be resolved through the use of argument, for they undercut the conditions essential to arguing" (Fogelin 1985, 8). According to his argument, rational arguments must take place against a background of shared certainties. When two parties do not share a certainty (or a group of certainties), a deep disagreement ensues.

However, the situation is perhaps not beyond hope. Wittgenstein alludes to this in *On Certainty* when he imagines a king who believes that the world began when he was born. He imagines Moore trying to convince the king that his belief was wrong, and that our certainty that the world has existed long before our birth is the correct one, saying: "And if Moore and this king were to meet and discuss, could Moore really prove his belief to be the right one? I do not say that Moore could not convert the king to his view, but it would be a conversion of a special kind; the king would be brought to look at the world in a different way" (OC 92). This passage illustrates the difficulty of bridging a deep disagreement while also pointing to a possible solution. Wittgenstein suggests that, since there is a lack of shared certainties, to overcome such a disagreement would involve more than a rational demonstration of the grounds for such a belief but would entail a change of worldviews. So, while a deep disagreement is not rationally resolvable, it could potentially be resolved through other, more radical means.

Picking up on this suggestion, this paper will explore how such a conversion might take place in a moral context. I will use the example of the wrongness of cannibalism and see how one might ultimately give up this certainty through a process of conversion. In order to do this, I will first begin by arguing that the wrongness of cannibalism is a moral certainty. I will support this argument by examining media that portrays cannibalism, showing that these take for granted the fact that the viewer has a deep moral aversion to it. I will also appeal to the real-world case of the Uruguayan Air Force Flight 571 and emphasize the fact that though the people involved in the

cannibalistic acts were rationally convinced that their choice to eat the dead was the right one, they still felt deeply disturbed by it in a moral sense. By examining this case and the reactions to it, I hope to establish that the wrongness of cannibalism was so deeply held that it can be labelled a certainty.

Then, I will examine historical cases of communities that regularly engaged in cannibalism to show that this certainty is a local one. In particular, following Mikel Burley (2016), I will examine the case of the Wari people in Western Brazil, who ate parts of deceased loved ones as a part of their grieving process. Finally, I will reflect on how someone might be fully brought to the view that cannibalism is acceptable through a conversion of a special kind. I will bring attention to the fact that many of the elements necessary for seeing the Wari's form of cannibalism as morally acceptable exist already in our form of life. For instance, we have our own rituals for respecting the dead, such as elaborate funerals, burials, and so on. Using this as a starting point, I will reflect on how this may serve as a starting point for such a conversion and consider to what extent it would be possible to fully embrace the opposite view.

References

Burley, M. (2016). Eating Human Beings: Varieties of Cannibalism and the Heterogeneity of Human Life. Philosophy (London), 91(358), 483–501.

Fogelin, R. (1985). The Logic of Deep Disagreements. Informal Logic, 7(1), 3-11.

Wittgenstein, Ludwig. (1969). *On Certainty*. (trans. GEM Anscombe). Oxford: Blackwell.

João Lemos
NOVA University Lisbon

Kant on deception and bad rhetoric

Kant describes bad rhetoric as a deceitful art. Literature on this topic is focused for the most on why bad rhetoric's deceptiveness is objectionable (Dostal 1980, Abbott 2007, Stroud 2014, Ercolini 2016, Leeten 2019). In this paper, I rather focus on what it means for Kant that bad rhetoric is deceitful in the first place.

I give an account of the deceptiveness of bad rhetoric. I argue that bad rhetoric deceives insofar as the semblance of bad rhetoric does not agree, or cannot harmonize, with the cognition of truth about bad rhetoric. I propose that not only does the disagreement between the semblance and the cognition of truth comprehend bad rhetoric's deceptiveness, it also expresses bad rhetoric's deceitfulness (intentional deceptiveness). I conclude with the moral and aesthetic repercussions of bad rhetoric being a deceitful art.

In the first part of my paper (sections 1-3), I give an account of what it means that bad rhetoric deceives. I start by reconstructing Kant's technical definition of deception as disagreement between the semblance and the cognition of truth (Anthropology from a Pragmatic Point of View and the transcripts of Kant's lectures on anthropology by others than Kant, as well as the address known as Entwurf). I go on to argue that bad rhetoric deceives in what it says about itself, that is, in what it promises, in comparison with what it provides (Critique of the Power of Judgment). In section 3, I show that the disagreement between what bad rhetoric promises and what it provides falls under Kant's technical definition of deception, and I conclude that bad rhetoric deceives insofar as the semblance of bad rhetoric does not agree, or cannot harmonize, with the cognition of truth about bad rhetoric.

 In the second part (sections 4-5), I propose that not only does bad rhetoric's disagreement between the semblance and the cognition of truth correspond to bad rhetoric's deceptiveness, it also expresses bad rhetoric's deceitfulness (intentional deceptiveness), and I also consider the moral repercussions of bad rhetoric being a deceitful art. I start by showing that Kant's definition of deception as 'Betrug' implies that deception is intentional. I go on to argue that a number of cases that literature on Kant takes to be deceptive – e.g. 'deception of the senses' (first Critique), deception by and deception of 'the deceiver in ourselves' as well as 'moral permissible illusion' (anthropological texts) – are not deceptive after all. I distinguish between bad rhetoric's deceitfulness and other kinds of intentional deceptiveness, such as inner lies (Metaphysics of Morals) and artificial deception of the senses (Anthropology). I conclude by arguing that, being deceitful, not only is bad rhetoric morally wrong, it does not give pleasure, therefore it cannot be neither beautiful nor even an agreeable art.

Keywords

Karol Lenart
University of Warsaw

Linguistic Ersatzism and Counterpart Theory

Linguistic ersatzism (Heller 1998, Melia 2001, see also Lewis 1986) explains modal facts in terms of ersatz possible worlds and ersatz possible individuals, understood as maximal and consistent collections of sentences of an idealised language L. In order to explain modalities *de re* about individuals, some linguistic ersatzists advocate counterpart theory. In this presentation, I identify and discuss several problems associated with such a view.

Let's start with the construction of ersatz possibilia. The idealised language L is assumed to be nonmodal, infinite, and to have a Lagadonian interpretation, which means that every individual and every property names itself. This is important because ersatz possibilia *qua* linguistic constructs represent possibilities by naming: Socrates exists in some world w iff w contains the name 'Socrates'. So, to ensure that everything is represented, everything must have a name. Now, since actual individuals exist, they can be named. Thus, ersatz worlds representing them can be identified with conjunctions of sentences with constants like: F(a) & G(b) & R(ab) &… . Because actual individuals are named, we can track them across ersatz worlds by tracking their names. Thus, modalities *de re* concerning actual individuals can be analysed in terms of trans-world identity. This, however, does not work in the case of alien individuals, i.e., individuals which do not exist, but could (like a talking donkey). Since they do not exist, they cannot be named. Therefore, they can be represented only generally, through quantificational sentences without constants. An ersatz world representing aliens could look like this:

(1) $\exists x \exists y …(Fx \& Gy \& Rxy …)$

Ersatz individuals representing aliens are no longer determinate parts of ersatz worlds. In order to recover an ersatz individual from (1) we need to unbind one variable in (1). We obtain:

(2) $\exists y …(Fx \& Gy \& Rxy \& …)$

(2) represents a maximal and consistent qualitative role that could be satisfied by something. An important feature of ersatz individuals like (2) is that they mirror the whole ersatz world representing them as existing. This, in turn, implies that (2) is worldbound: (2) cannot be represented by any other ersatz world than (1). This leads to the conclusion that all properties possessed by an alien are essential to it. Thus, a talking donkey is essentially a talking donkey.

In order to recover modal variability of aliens, the linguistic ersatzist can endorse counterpart theory (Heller 1998, Wang 2015). According to such a view, a

talking donkey could be a flying donkey iff an ersatz flying donkey is a counterpart of an ersatz talking donkey.

There are, however, several problems with this proposal (some of them have been indicated by Merricks 2003, Sider 2006 and Woodward 2011; some of them are new).

(1) There are many alternative candidates for the vocabulary of L: it could be numbers, sets, atoms, LEGO blocks, or whatever you prefer. This variability is acceptable as long as each candidate L-language is nonmodal, infinite, and has a Lagadonian interpretation. However, a problem is that which L-language we choose is an arbitrary matter. This is problematic because expressions of L provide a reductive basis for the modality. Thus, if the world-making language is chosen arbitrarily, then the reduction itself will also be arbitrary.

(2) Possibilia, apart from being the reductive basis of modality, are also supposed to be subjects of various metaphysical relations and concepts, e.g., closeness relations, counterpart relations, spatiotemporal relations, relations of dependence, or laws of nature. Thus, if there are various alternative ways in which the ersatzer can construct ersatz possibilia, then presumably our metaphysical concepts will not apply equally well to all of them. Moreover, if we arbitrarily choose L-language, the range of metaphysical applications of possibilia will also be arbitrarily determined.

(3) Linguistic ersatzism seems to be susceptible to a generalised Humphrey objection: how linguistic possibilia are is irrelevant to how modality works. Moreover, many of the metaphysical concepts central to modal metaphysics are inapplicable to possibilia *qua* linguistic constructs.

(4) Linguistic possibilia do not represent intrinsically, by themselves, but extrinsically, in virtue of our interpretations. Thus, a world without interpreters had been actualised, linguistic possibilia would not represent anything.

(5) There are several problems with ersatz counterpart relations, which are supposed to link ersatz individuals: (i) if they hold between linguistic constructs, it is unclear how to measure similarity between them; (ii) even if we manage to establish they do, similarity relations between linguistic constructs seem irrelevant to how represented individuals might have been; (iii) if, in turn, counterpart relations hold between represented individuals, and we consider the case of aliens, it is impossible to measure similarity between aliens, until worlds representing them as existing are actualised.

(6) Because ersatz individuals representing aliens mirror entire worlds, it follows that every such individual is similar to every other individual within the same world. So, every individual is a counterpart of every other individual within that world. This means that everything could be everything else within that world, which implies the controversial thesis of extreme haecceitism.

These analyses lead me to two conclusions. First, the linguistic ersatzer should abandon the counterpart theory and accept that ersatz individuals representing aliens are worldbound. I argue that this is not as problematic as it might seem. Second, many of the problems with linguistic ersatzism can be avoided

if we explain modality in terms of genuine abstract possibilia, such as sets of propositions or properties, rather than linguistic ones.

References

Heller, M. (1998). Property counterparts in ersatz worlds. The Journal of Philosophy, 95(6), 293-316.

Melia, J. (2001). Reducing possibilities to language. Analysis, 61(1), 19-29. Merricks, T. (2003). The end of counterpart theory. The Journal of Philosophy, 100(10), 521-549.

Lewis, D. K. (1986). On the plurality of worlds (Vol. 322). Oxford: Blackwell. Sider, T. (2006). Beyond the Humphrey objection. Unpublished manuscript.

Wang, J. (2015). Actualist counterpart theory. The Journal of Philosophy, 112(8), 417- 441.

Marcin Lewinski

NOVA University Lisbon

Metalinguistic appeals

The literature on various forms of metalinguistic disagreements and conceptual engineering typically assumes these forms manifest themselves, *qua* activities, in disputes where two parties normatively argue over the conceptual details of a given term (Cappelen, 2018; Ludlow, 2014; Plunkett & Sundell, 2013, 2021; Schiappa, 2003). The basic objective is to come up "with progressively more serviceable modulations via a normatively constrained process of argumentation" (Ludlow, 2014, p. 111). Participants in this process can, importantly, resort to arguments from the authority of agents external to the dispute itself in support of their metalinguistic positions (Ludlow, 2014). One underlying assumption behind this work – and yet conspicuously underexplored – is that the authority of speakers-qua-stakeholders plays an important role in such disputes. But if this is so, then the weight of one's engagement in various forms of metalinguistic disputes cannot be evenly distributed: the powerful epistemic and deontic authorities would seem to always have the upper hand. A growing body of recent literature starts addressing this issue (Kitsik, 2023; Podosky, 2022; Shields, 2021).

Against this background, in this paper I will take up two basic tasks.

First, I will theorise authority, in its most abstract sense, as inherently triadic. As a point of departure, I will take Bocheński's "logic of authority" (Bocheński, 1965; cf. Brożek, 2013; Kubalica, 2022). For Bocheński, "the formal structure of every authority is […] this: x is an authority for (the subject) y, in (the field) α — symbolically: $A(x, y, α)$" (Bocheński, 1965, p. 163). This dyadic relation between an authority and the subject is captured, among others, in Lewis' Master-Slave dynamic, whereby the Master exercises his authority by defining the sphere of permissibility for the Slave (Lewis, 1979a). Further analyses have brought this relation to bear on the details of various power dynamics in speech act exchanges, notably those where some form of subordination is in question (e.g., Langton, 1993, 2015; McGowan, 2019; Adams, 2020). By contrast, I will define authority as (minimally) triadic: $A(x, y, z, […], F, α)$, whereby agent x, rather than y, has authority over a set of subjects z[…], before the forum F, in (the field) α. Forum is understood as a source of normativity appealed to. So extended concept captures some basic authority relations, such as those of majority-minority, divide and-rule, mediation and arbitration, co-optation, etc. (e.g., Simmel, 1950; cf. Lewis', 1979b, Master-Slave-Kibitzer dynamic). It also captures the appeals to authority, a focus of this paper.

Second, I will isolate and analyse the class of speech act central to the non-ideal metalinguistic disputes: *metalinguis/c appeals*. Appeals are directive speech acts which, as understood here, require precisely (at least) a triadic communicative situation. They are speech acts in which one party, in a dispute with another party,

appeals to the third party – the Forum – for the authoritative recognition of its metalinguistic "modulation". In this sense, appeals have both similarities and differences with dyadic *metalinguis/c proposals* and *metalinguis/c provoca/ons* (Hansen, 2021): a speaker similarly acknowledges that "the speaker making the competing metalinguistic proposal would not accept" their proposal (Hansen, 2021, p. 8), but (whether in spite or because of that) they seek recognition and acceptance of it from the third, authoritative party, rather than merely a provocation of their opponents.

As a case in point, I will analyse Yasser Arafat's famous "rejection of terrorism in all its forms, including state terrorism" (Addressing the UN General Assembly in Geneva, 1988). Here, Arafat cannot be trying to reasonably convince his direct adversary – the government, citizens, and supporters of the State of Israel – that what Israel does against Palestinians is "included" in the extension of the concept TERRORISM, as its species ("*state* terrorism"), and should as such be equally condemned. Instead, he *appeals* to the UN and other international bodies, governments, and citizens that this is so. This, then, serves as a premise in an argument towards a balanced, two-sided approach to the Israeli-Palestinian conflict.

References

Adams, N.P. (2020). Authority, illocutionary accommodation, and social accommodation. *Australasian Journal of Philosophy, 98*(3), 560-573.

Austin, J.L. (1962). *How to do things with words*. Clarendon Press.

Bocheński, J. M. (1965). Analysis of authority. In *The logic of religion* (pp. 162–173). New York University Press.

Brożek, A. (2013). Bocheński on authority. *Studies in East European Thought, 65*(1-2), 115–133.

Cappelen, H. (2018). *Fixing language: An essay on conceptual engineering*. Oxford University Press.

Hansen, N. (2021). Metalinguistic proposals. *Inquiry, 64*(1-2), 1-19.

Kitsik, E. (2023). Epistemic paternalism via conceptual engineering. *Journal of the American Philosophical Association, 9*(4), 616-635.

Kubalica, T. (2022). The relation of master and disciple against the background of Józef M. Bocheński's logic of authority. *Studies in East European Thought, 74*(2), 225–240.

Langton, R. (1993). Speech acts and unspeakable acts. *Philosophy and Public Affairs, 22*(4), 293-330.

Langton, R. (2015). How to get a norm from a speech act. *The Amherst Lecture in Philosophy, 10*, 1–33.

Lewis, D. (1979a). Scorekeeping in a language game. *Journal of Philosophical Logic, 8*(1), 339–359.

Lewis, D. (1979b). A problem about permission. In E. Saarinen, R. Hilpinen, I. Niiniluoto, & M. Provence Hintikka (eds.), *Essays in Honour of Jaakko Hintikka* (pp. 163-175). D. Reidel Publishing Company.

Ludlow, P. (2014). *Living words: Meaning underdetermination and the dynamic lexicon*. Oxford University Press.

McGowan, M. K. (2019). *Just words: On speech and hidden harm*. Oxford University Press.

Plunked, D., & Sundell, T. (2013). Disagreement and the semantics of normative and evaluative terms. *Philosophers' Imprint, 13*(23), 1–37.

Plunked, D., & Sundell, T. (2021). Metalinguistic negotiation and speaker error. *Inquiry, 64*(1–2), 142–167.

Podosky, P-M. (2022). Agency, power, and injustice in metalinguistic disagreement. *The Philosophical Quarterly 72*(2), 441-464.

Schiappa, E. (2003). *Defining reality: Definitions and the politics of meaning*. Southern Illinois University Press.

Schields, M. (2021). Conceptual domination. *Synthese, 199*, 15043–15067.

Simmel, G. (1950/1910). *The Sociology of Georg Simmel*. Wolff, Kurt H. (trans. & ed.). The Free Press.

Matilde Liberti
University of Genoa

Neoptolemus and Huck Finn reconsidered. Alleged inverse akrasia and the case for moral incapacity

Cases of *akratic* behavior are generally seen as paradigmatic depictions of the knowledge-action gap (Darnell et al 2019): we know what we should do, we judge that we should do it, yet we often fail to act according to our knowledge. In recent decades attention has been given to a particular instance of *akratic* behavior, which is that of "inverse *akrasia*", where the agent possesses faulty moral knowledge but fails to act accordingly, thus ending up doing the right thing. In particular, two literary examples are considered as exemplifying this kind of *akratic* situation: Huckleberry Finn (Arpaly & Schroeder 1999, Arpaly 2015, Hursthouse 1999, Kleist 2009, Holton manuscript) and Neoptolemus as understood by Aristotle (NE; Arpaly & Schroeder 1999). In this paper I will argue that those of Neoptolemus and Huck Finn are not cases of inverse *akrasia* (Holton manuscript) but are much better explained as instances of what Williams (1993) called "moral incapacity". In particular, the reason why they fail to act according to their original judgments is due to a lack of motivation to act accordingly, which is grounded in their moral self-identities (Blasi 1984; Vigani 2016).

The paper will unfold as follows: I will, first, argue that neither Neoptolemus nor Huck Finn show *akratic* behavior; thus, they cannot be legitimately labeled as "inverse *akratics*". I will start from the following observation: if we consider Aristotle's structure of *akratic* action, we can see that inverse *akrasia* does seem to share the same psychological blueprint of *akrasia*: (i) it is action against one's *prohairesis*; (ii) it is voluntary; (iii) it is done out of some sort of ignorance; (iv) it is marked by conflict. Here, I shall argue that neither Neoptolemus nor Huck Finn are cases that satisfy (i); quite the contrary, they act in perfect accordance with their *prohairesis*.

However, if we exclude that those of Neoptlomeus and Huck Finn are cases of inverse *akrasia*, we are left with no explanation as per why they ultimately act the way they do. In the second part, I will argue that they act the way they do notwithstanding their faulty judgments because they are effectively motivated to do so. Such motivation originates in their moral self-identities (Blasi 1984; Vigani 2016) and is experienced through the threat of self-betrayal (Lapsley 2008). According to this picture, then, Neoptolemus cannot bring himself to carry on Odysseus' orders, because his commitment to honor is much more *central* (as opposed to *peripheral*)

to his moral self-identity than he himself would have imagined. Similarly, Huck Finn's commitment to friendship is so central to his moral self-identity that his "faulty" moral knowledge that slavery is fair and that Jim should return to his owner turns out not being motivating, or not as strongly as his *core* commitment to friendship.

Finally, I will end by suggesting that there is a more fitting label for cases such as those of Neoptolemus and Huck Finn, than that of "inverse *akrasia*", which is what Williams (1993) named "moral incapacity". In particular, when an agent is motivated to act in a way that is integral with her moral self-identity, acting otherwise is experienced as something one *cannot* ultimately do; that is, as a moral incapacity. This "cannot" is neither a metaphor for an "I shouldn't", nor an instance of what has been recently labeled a "moral impossibility" (Caprioglio Panizza 2020, 2021), since it does not arise from the normative force of deontic judgments, and it is neither physically nor psychologically impossible for the agent to act otherwise. In other words, these kinds of incapacities are not moral phobias: agents are *not* disgusted, nor impulsively pushed away from other options. What it is that does the job is moral self-identity; the fact that notwithstanding the attractiveness of other courses of action, the agent decides that she cannot ultimately betray her moral self.

References

Arpaly, N. (2000). On Acting Rationally against One's Best Judgment. *Ethics*, 110(3), 488–513. https://doi.org/10.1086/233321

Arpaly N., Schroeder T. (1999). Praise, blame and the whole self, *Philos Stud* 93: 161–188.

Blasi, A. (1984). Moral identity: Its role in moral functioning. In W. M. Kurtines & J. J. Gewirtz (Eds.), *Morality, moral behavior and moral development* (pp. 128–139). New York: Wiley.

Caprioglio Panizza, D. (2020) "If Veganism Is Not a Choice: The Moral Psychology of Possibilities in Animal Ethics", *Animals* 10 (145): 1-12. DOI: 10.3390/ani10010145

Caprioglio Panizza, S. (2021), "Forms of Moral Impossibility", *European Journal of Philosophy* 2021: 1-13. DOI: 10.1111/ejop.12651

Darnell, C., Gulliford, L., Kristja´nsson, K, Paris, P. (2019) Phronesis and the Knowledge-Action Gap in Moral Psychology and Moral Education: A New Synthesis?, *Human Development* 2019; 62:101–129, DOI: 10.1159/000496136.

Holton, R. Inverse *akrasia* and weakness of will, MIT, manuscript.

Hursthouse, R. (1999) *On virtue ethics*. Oxford University Press, NY.

Kleist, C. (2009). Huck Finn the inverse *akratic*: empathy and justice, *Ethical Theory Moral Practice* 12: 257–266.

Lapsley, D. K. (2008). Moral self-identity as the aim of education, in *Handbook of moral and character education*, Larry, P. Nucci & Darcia Narvaez (eds.), Routledge: New York.

Vigani, D. (2016). Moral Judgment and Motivation: Making sense of mixed intuitions, *Ethical Perspectives* 23 (2): 209-230.

Williams, B. (1993). Moral incapacity. *Proceedings of the Aristotelian Society*, 93, 59–70. https://doi.org/10.1093/aristotelian/93.1.59

Vassilis Livanios
University of Cyprus

Against the grounding conceptions of the Powerful Qualities View

Recently, some philosophers (Contessa 2019; Coates 2021; Azzano 2021) have appealed to the notion of metaphysical grounding to provide an unproblematic version of the main tenet of the most popular version of the Powerful Qualities View, that is, the Identity Theory. In this talk, I will focus only on Coates' version. According to his view, a property's qualitativity consists in its having an essence that does not involve genuinely distinct properties (that is, an intrinsic one), while a property's dispositionality consists in its having an essence that partly grounds its occupation of its dispositional role. Given those definitions, Coates suggests that the core principle of the Identity Theory can be stated in the following way: a property's qualitativity is identical to its dispositionality in the sense that both consist just in its having its particular nature, which is both intrinsic and a partial ground for its occupation of its dispositional role.

The major problem besetting this grounding view is that, depending on the formulation of its grounding claim, its main tenets fail to distinguish it from either the necessitarian versions of Pure Qualities Monism or a kind of Property Dualism. Let us first suppose that Coates opts for a predicational formulation of his grounding claims. The formulation in question is ontologically demanding because it presupposes that the relata of the grounding link are *entities* of some sort. So, if we take Coates' words literally, he is committed to the view that the relata of his grounding claim (that is, the essence of a powerful quality and the dispositions of its bearers) are both existing entities of some sort. Setting aside the issue of whether the essence of an entity can be plausibly regarded as an entity itself, I would like to focus here on the view that the *dispositions* of objects are entities of some sort. The reification of dispositions threatens the very *autonomy* of the grounding view. To illustrate that, note that on Pure Powers Monism the dispositions of objects are either identified with their properties (Mumford 2004) or constitute the essences of their properties (Bird 2007). The proponents of the grounding view, however, should reject those theses about dispositions. Since grounding is asymmetric, whereas identity is a symmetric relation, one cannot hold that the relation between properties and dispositions is both identity *and* grounding. In addition, since the typical view about grounding is that the grounded entities are less fundamental than their grounds, one cannot hold that properties are essentially constituted by (and hence depended on) dispositions. For otherwise the proponent of the grounding view would have to embrace the view that even those properties

which are fundamental are essentially depended on ontologically derivative entities. Hence, given that the advocate of the grounding view should reject the claim that the properties of things are identical with, or essentially dependent on, the dispositions of those things, if they want to reify the latter, they should regard them as belonging to a 'dispositional kind' of properties, a 'kind' which is *distinct* from the one to which powerful qualities belong. In this case then, the defenders of the grounding view should be committed to the view that objects have two 'kinds' of property (properties with intrinsic qualitative identity and properties which are dispositions) and the properties (or the essences of the properties) of the former 'kind' ground the properties of the latter kind. Since the properties of the former 'kind' have qualitative natures, they are in fact *qualitative* properties which necessarily determine the dispositions of their bearers. That in turn means that the view in question is actually a version of Property Dualism which differs from the 'traditional' one only in the fact that qualitative properties are ontologically more fundamental than the dispositional ones.

Given that the reification of dispositions raises difficulties for the grounding view, an advocate of this approach may embrace an ontologically non-committing operational formulation of their grounding claim of the form: "the object x has dispositions Di becauseg it possesses the property P", where the term "becauseg" stands for the grounding operator. In that case, the proponent of the Grounding View would not deny that there are *truths* about the dispositions objects possess: they would simply deny that the existence of those truths entails the existence of certain *entities* which we call dispositions. Yet, by simply rejecting the reification of dispositions of objects, the proponent of the grounding view cannot obviously 'save' its autonomy. For, even if the defenders of the Grounding View do not appeal to any ontological 'elements' of properties and simply assert that each property has an intrinsic qualitative identity and grounds, either by itself or in tandem with other properties, specific dispositions of its bearers, their view is in fact the core tenet of the necessitarian versions of Pure Qualities Monism. Hence, I cannot see any *substantial* difference between the grounding view and those versions of Pure Qualities Monism.

References

Azzano, L. (2021). Dispositionality, categoricity, and where to find them. *Synthese 199*, 2949–2976.

Bird, A. (2007). *Nature's metaphysics: Laws and properties*. Oxford: Oxford University Press.

Coates, A. (2021). Making sense of powerful qualities. *Synthese* 198, 8347–8363.

Contessa, G. (2019). Powerful qualities or pure powers? *Metaphysica 20*(1), 5–33.
Mumford, S. (2004). *Laws in Nature*. London and New York: Routledge.

Hugo Luzio
University of Lisbon

Animalism and Person Conativism

Animalism is the view that each of us (i.e., each being of our metaphysical kind) is essentially a human organism and contingently a person (Inwagen 1990, Snowdon 2014, Olson 1997). In this view, we persist over time by virtue of the continuity of our biological life. As such, no form of psychological continuity is necessary or sufficient for our persistence.

This view has been widely criticized by neo-Lockean philosophers, who believe each of us is essentially a 'person' in Locke's (1694) sense: a being with psychological persistence conditions (Parfit 1984, McMahan 2002).

According to animalists, 'person' is not a substance sortal, but a phase sortal (Wiggins 2001): each of us existed as a non-person in the past (e.g., as a newborn) and can come to exist again as a non-person in the future (e.g., as a patient in a persistent vegetative state). When it comes to our persistence, thereby, nothing hangs on our being a person.

Interestingly, however, most animalists are Lockeans about personhood. For them, what makes something a 'person' is the exhibition of complex psychological properties (e.g., consciousness and thinking) (Olson 1997: 102-105).

In being Lockeans about personhood, most animalists are person objectivists. For person objectivists, there is a natural, non-conventional fact as to whether something is a 'person'. According to animalists, whether an organism (i.e., one of us) is a 'person' at some moment is a matter of psychological fact.

Recently, however, some philosophers have argued that persons are not natural products, but rather conventional constructs. In this paper, I explore the relation between animalism and person conventionalism. In particular, I discuss whether animalism is compatible with the main form of person conventionalism in the literature: person conativism (Braddon-Mitchell and Miller 2004, 2020).

I begin by presenting animalism and person conativism. Then, I distinguish between two forms of person conativism, and argue that animalism is incompatible with the former, but not with the latter. Finally, I develop a phasal person conativist form of animalism and argue that it has several advantages over the Lockean conception of personhood that animalists traditionally endorse.

According to person conventionalism, in some sense or other, it's up to us which things are a 'person' at a moment, or the 'same person' over time. The most influential person conventionalist view is person conativism:

> (PC) which existing things count as the 'same person' over time is partly settled by which conative attitudes (hereafter, 'conations') obtain (Braddon-Mitchell and Miller 2020).

In this view, there are certain conative-independent relations (e.g., psychological and biological continuity). But it is the obtaining of certain conations that settles which of those relations, if any, is the 'same person' relation (hereafter, 'SP-relation'). What makes two things the 'same person' is thus partly a matter of conation. Here is the person conativist conception of the SP-relation:

> (SP-relation) A person x at a moment t and another person y at a distinct moment t* are the same person if and only: (1) a conative-independent relation holds between x at t and y at t*, and (2) the holding of that conative-independent relation is conatively relevant for x at t or y at t* (or both).

The compatibility between animalism and person conativism depends on the person conativist stance on the relation between the kind of thing we are (i.e., our basic metaphysical nature) and the property of being a person (i.e., personhood). Is each of us essentially a person, with partly conative persistence conditions? Or is each of us only contingently a person?

We can distinguish between two forms of person conativism. Substantial person conativism (SPC) is the view that each of us is essentially a person, with partly conative persistence conditions (Braddon-Mitchell and Miller 2004).

In this view, the SP-relation is a relation of numerical identity (between things that are essentially persons): if two persons are the 'same person', then they are just one of us (i.e., a single being of our metaphysical kind), rather than two. As such, in partly settling the SP relation, our conations partly settle our persistence conditions.

Animalism is blatantly incompatible with SPC. According to animalism, our persistence depends on the continuity of our biological life, rather than on our capacity to have conations (or any other psychological capacity). As such, our persistence cannot be a matter of conation.

There is, however, an alternative understanding of person conativism. Phasal person conativism (PPC) is the view that each of us is contingently a person and that being a person (i.e., exhibiting personhood) is a conventional property.

In this view, the SP-relation obtains between two of us when they are the 'same person' in a "sameness" sense weaker than numerical identity: when they are the same individual person. And what makes two of us the same individual person is partly a matter of conation.

Animalism is perfectly compatible with PPC. Since PPC is not a person essentialist view, in this view, our being a 'person' or the 'same person' is partly a matter of conation, but our persistence is indeed a matter of natural, non-conative fact. Moreover, there are good reasons to think that animalists should endorse PPC, rather than Lockeanism about personhood.

Whereas PPC is well equipped to accommodate 'same person' talk, there is no good contingentist Lockean view on what it means for a person to persist over time. And PPC offers substantial explanatory support to a wide range of plausible

positions animalists can adopt in challenging cases, which could hardly be sustained with a Lockean conception of personhood at hand.

For instance, PPC-animalists can maintain that: (i) one of us (i.e., a single human organism) can be a different person at distinct moments, without ceasing to exist (i.e., that we can change as a person, but nonetheless persist); (ii) two of us (i.e., two distinct human organisms) can be the same person at some moment (e.g., in brain transplant scenarios, including fission); and finally, (iii) one of us can be more than a single person at some moment (e.g., in conjoined twining cases).

References

Braddon-Mitchell, David; Miller, Kristie, 2004, "How to be a Conventionalist Person", *The Monist*, Vol. 87, No. 4, pp. 457-474.

Braddon-Mitchell, David; Miller, Kristie, 2020, "Conativism about Personal Identity", in Andrea Sauchelli (ed.), *Derek Parfit's Reasons and Persons: An Introduction and Critical Inquiry*, Abingdon: Routledge, pp. 159-269.

Locke, John, [1694] 1959, *An Essay Concerning Human Understanding*, Campbell (ed.), Dover Publications Inc.

McMahan, Jeff, 2002, *The Ethics of Killing: Problems at the Margins of Life*, Oxford University Press.

Olson, Eric, 1997, *The Human Animal: Personal Identity Without Psychology*, Oxford University Press.

Parfit, Derek, 1984, *Reasons and Persons*, Oxford University Press.

Snowdon, Paul, 2014, *Persons, Animals, Ourselves*, Oxford University Press.

van Inwagen, Peter, 1990, *Material Beings*, Cornell University Press.

Wiggins, David, 2001, *Sameness and Substance Renewed*, Harvard University Press.

Borbála Jász
Kodolányi János University

Limitations of Acquaintance in Architectural Meaning Construction

It has been a highly debated issue in the philosophy of art if relevant knowledge about works of art may be gained through testimony as well or only through acquaintance. In relation to photos taken of (visual) works of art, it might be argued that a photographic reproduction of a work of art displays the relevant features counterfactually. If that is so, then it can be questioned to what extent the photographic reproduction can be considered testimony and to what extent it can be considered an experience equivalent to perceptual acquaintance.

Kendall Walton argues that looking at a photograph of a person is equivalent to the original perception of the person – which he calls the transparency of photographic images. In his essay 'Transparent Pictures: On the Nature of Photographic Realism' (1984) he argues that „*the viewer of a photograph sees, literally, the scene that was photographed* [...] *my visual experience when seeing the photograph is counterfactually dependent on the object photographed: if X's visible properties were different, my visual experience when seeing the photograph would be correspondingly different.*" (Walton 1984. 252) Walton also notes that photographs preserve real similarity relations (such as shapes) as opposed to verbal descriptions that may also be counterfactually dependent on the visual properties of the scenes that they describe.

In my talk I will not argue for or against the transparency thesis itself, but I will rely on the idea that our visual experience when seeing photographs is counterfactually dependent on the visual properties of the scene photographed (and that they also preserve real similarity relations). I present architectural examples in which photographic testimony is arguably more fruitful (richer, more detailed, more relevant, etc.) from the point of view of aesthetic knowledge than the first hand experience, the acquaintance itself. I analyse three examples (belonging to two types) to support the argument that acquainting oneself with the built environment does not necessarily require personal perception and experiences; for all relevant purposes equally sufficient (or even richer, more detailed, more relevant) information can be obtained through photographs.

The first example presents questions of interior design, which can be illustrated with the Church of Sant' Ignazio di Loyola in Rome. The church was constructed in the Baroque style between 1626 and 1650, featuring the renowned illusionist ceiling fresco by Andrea Pozzo. The fresco, a highlight of the church, depicts an

imaginative architectural space and employs clever perspective to portray the heavenly sphere. It showcases Saint Ignatius being received by Jesus and the Virgin Mary, accompanied by allegorical figures representing the four continents. The church would provide a more detailed view if one could observe anything from this emblematic composition with the naked eye up close. In this case, however, there are simple physical limitations: without aids (mirror, binoculars), the ceiling fresco of the church is so far from the person entering the building that it cannot be sufficiently observed with the naked eye.

To address this limitation, a magnifying mirror known as "Europe's most beautiful mirror" was installed in the church. Despite this, even the mirror falls short of revealing the fresco in its entirety; the work is much betterstudied and appreciated through reproductions best taken by today's drone photography. (Nevertheless, the mirror serves as an excellent backdrop for selfies, allowing visitors to demonstrate their presence in Rome for the weekend, although the ceiling's details remain elusive in such photos.)

The second group of examples can include several types of buildings. In general, it can be observed that buildings can be better understood through photographs than with first-hand experience. This holds true to the extent that a building cannot be fully comprehended perceptually. Even during the planning phase, various views (such as ground plans, longitudinal sections) are employed because relying on a single perspective would not offer a complete understanding of the building, not even for the architect.

I will illustrate two cases exemplifying the limitations of acquaintance in this category. One instance is the Great Wall of China, which, due to its vast size, cannot be fully observed by the human eye. First-hand experience of the Great Wall is only possible for astronauts who have viewed it in its entirety from outer space. We can experience the Great Wall as a whole based on photographic satellite images only. When visiting the Great Wall, we may claim to have seen it, even though physical constraints limit us to viewing only a section of it with the naked eye.

Before delving into the second example of this category, providing some historical context is essential. The roots of architectural photography trace back to modernism between the two world wars. Photography emerged as a pivotal tool, enabling the display of intricate building details imperceptible to the naked eye. The lens allowed structures to be captured from angles inaccessible to human vision, introducing a new aesthetic dimension to architectural photography. Architects of this era often collaborated with dedicated photographers capable of presenting their buildings from innovative perspectives.

Postmodernism took this a step even further by designing buildings explicitly for photography. When standing in front of such a building, its aesthetic essence primarily designed for aerial photography presenting a new city skyline landmark may elude the naked eye. An illustrative case is the Berlin Jewish Museum, designed by Daniel Libeskind. Some of the most important visual properties of the postmodern wing, resembling an unfolded Star of David, remains imperceptible when standing in front of and walking through the building. The central concept of the structure becomes apparent only through aerial photographs, underscoring the limitation of the acquaintance.

In my paper I argue that – aesthetic – knowledge about architecture doesn't necessitate first-hand experience; in certain cases significantly more and more relevant knowledge can be derived from photographs of the building, acting as testimonial. Concerning the reception of architectural works, all of this implies that, owing to the physical limitations of our eyes, first-hand experience is not necessarily the most efficient means of acquiring the most comprehensive aesthetic knowledge and appreciation.

These examples vividly illustrate the limitations of acquaintance. However, these constraints are not exclusive to emblematic buildings; they apply to any building. This is because, with the naked eye alone, we are incapable of perceiving the entirety of a building. A visual medium is required to complement or even replace the first-hand experience. Taking it a step further, if we obtain a comprehensive view of the building through a remotely accessible visual device, it raises the question of whether a personal viewing, a first-hand experience of an architectural product is necessary at all to achieve the most thorough understanding and appreciation. In my paper, I also argue and explain how the attitude of contemporary architectural designers changes the appreciation process due to photographs.

Christos Kyriacou
University of Cyprus

Rationality, Eliteness and Reference Magnetism

As I have argued elsewhere, if we are to avoid universal skepticism and epistemic self defeat, some minimal reliability with the meaning and reference determination of the predicate 'epistemic rationality' should be taken to be *explanatorily indispensable.* Call this *'the minimal rationality requirement'.*

Roughly, this is because if any radically skeptical argument is to be epistemically rational *at all* (or indeed any other argument), then some minimal reliability with the application of 'epistemic rationality' *must* be the case. This modal result implies that it is plausible to think that there are some epistemic rationality norms that any rational argument, of necessity, should abide by. Minimal reliability with 'epistemic rationality' is one of them and elsewhere I have tried to conceptually excavate and bring to light some more indispensable epistemic rationality norms.

Detractors could argue that I am missing the skeptic's point because epistemic rationality *itself* is subject to a skeptical argument and, therefore, we are not making much progress. For example, we can use a skeptical argument from deep disagreement about epistemic rationality and alternative concepts of rationality (cf. Clarke-Doane (forth.), Eklund (2017)). Are we talking about Foley (1987) rationality, Nozick (1993) rationality*, Quine and Ullian (1970) rationality**, or what?. This is a serious challenge to be reckoned with, if we are not to beg the question against the skeptic about epistemic rationality.

In rejoinder, epistemic rationality seems to provide us with a semantic grip on what is harder, if not outright implausible, to be essentially contestable. For, if epistemic rationality is subject to radical skepticism and there is no argument that privileges some *elite* epistemic rationality property, then we are doomed to universal skepticism and its dire implications. Some, no doubt, could try to resist this skeptical implication, or would be happy to bite the bullet (e.g. Unger (1975, Streumer (2017), Rinard (forth.)), embrace a universally skeptical result and look to ameliorate the implications, but I am not very optimistic of the prospects of such radically revisionary metaepistemological projects. This is because of the severe problems such a version of radical skepticism would incur. I cannot delve into these problems here.

Suppose then that we have some minimal reliability with reference determination with regard to 'epistemic rationality'. In light of the explanatory indispensability of epistemic rationality, let us say that epistemic rationality is an *elite, referentially magnetic property* in the sense that the corresponding predicate must minimally reliably refer to the rationality property (as if the property has magnetic qualities and attracts the predicate's reference). As I understand eliteness, roughly, *elite*

properties are metaphysically fundamental properties that are explanatorily indispensable for objective rational argument. They are properties that should exist if we are to make sense of reality in a truth-tracking manner (and make sense of understanding itself, namely, the practice of gaining understanding).

I take it that such elite properties and relations include epistemic rationality, reference and truth. They are elite properties because any account of how we can be minimally rational should involve some minimally reliable reference (and thereby truth) to the predicate of 'epistemic rationality'. This suggests that we must also be minimally reliable with the application of the concepts of 'reference' and 'truth' themselves. Otherwise, if we had no minimal reliability with 'reference' and 'truth', we would not be in position to be minimally reliable with the application of 'epistemic rationality', which would imply some form of radical skepticism about rationality. The upshot would be that we are trapped in appearances about rationality (and everything else). If this is the case, then we are somehow stuck with some minimally reliable reference to 'epistemic rationality' because it is explanatorily indispensable.

Naturally, it could be objected that we can understand talk of epistemic rationality in terms of sophisticated antirealist positions, such as fictionalism/error theory, quasi-realist expressivism or relativism, without endorsing some form of radical skepticism about rationality. Such an antirealist position could argue that, although there is no robust epistemic rationality property (and no objective reference and truth about epistemic rationality), we can employ the concept of 'epistemic rationality' with pragmatic success. We can employ it with pragmatic success because we can speak *as if* there is such a property (error theory/fictionalism, cf. Streumer (2017)), or we can *earn the right* to speak of such a property in a deflated sense (quasi-realism, cf. Ridge (2018)), or that we can stipulate that such a property is relativist in some way ((cf. Stich (1990)), (Kusch 2009)) (MacFarlane 2016)).

Such sophisticated antirealist positions could be pragmatically successful in the sense that they could allow us to adequately explain how the conceptual practice of rationality (e.g. rationality assertions and attributions) works smoothly and allows us to be practically successful, both at an individual and at a social level. For example, suppose I assert that 'It is rational to believe that p, given evidence' (or something in the vicinity). An expressivist or fictionalist could argue that we speak as if there is a rationality property that believing that p exemplifies and this is pragmatically useful because it allows us to base our practical reasoning on otherwise useful beliefs. Such an account could also comport with a genealogical account (evolutionary, contractual, or other) of how and why we came to have the conceptual practice of epistemic rationality (cf. Craig 1990, Gibbard 1990, Kusch and McKenna 2018, 2020, Dogramaci 2012, Queloz 2021). Hence, there is no need for inflating our ontology with an elite epistemic rationality property.

But I think such antirealist arguments face a dilemma. Here is a very rough sketch. Either their argument for epistemic antirealism is epistemically rational or it is itself pragmatically rational; presumably it is rational in some sense, otherwise we would have no reason to accept it. On the first horn, it can't be categorically epistemically rational because there is no rationality property. So, there is no categorically epistemically rational reason to accept it. On the second horn of the dilemma, pragmatic reasons for belief (cf. Rinard forth.) seem to be the wrong kind of reasons

for believing something in the first place. I develop this argument from dilemma and illustrate that both horns appear unpalatable for the epistemic antirealist. I conclude that the possibility of elite, referentially magnetic properties is promising and should be further explored.

References

Cuneo Terence. (2007). *The Normative Web*. Oxford, Oxford University Press.

Clarke-Doane Justin. Forthcoming. 'Epistemic Non-Factualism and Methodology', Michael Klenk (ed.), *Higher-Order Evidence and Moral Epistemology*. Routledge.

Craig Edward. (1990). *Knowledge and the State of Nature*. Oxford, Oxford University Press.

Dogramaci Sinan. (2012). 'Reverse Engineering Epistemic Evaluations'. *Philosophy and Phenomenological Research* Vol. LXXXIV No. 3, 513-530.

Eklund Matti. (2017). *Choosing Normative Concepts.* Oxford, Oxford University Press.

Foley Richard. (1987). *The Theory of Epistemic Rationality.* Cambridge MA, Harvard University Press.

Gibbard Alan. (1990). *Wise Feelings, Apt Choices.* Cambridge MA, Harvard University Press.

Queloz Matthieu. (2021). *The Practical Origins of Ideas.* Oxford, Oxford University Press.

Quine W.V.O. and Ullian J.S. (1970). *The Web of Belief.* New York, Random House.

Kelly Thomas. (2003). 'Epistemic Rationality as Instrumental Rationality: A Critique'. Philosophy and Phenomenological Research 66(3):612-640.

Kusch Martin. (2009). 'Epistemic Replacement Relativism Defended'. In M.Suarez, Mauro Dorato and M. Redei (eds.), EPSA Epistemology and Methodology of Science. Dordrecht, Springer. pp.165-175.

Kusch Martin and McKenna Robert. (2018). 'The Genealogy of Relativism and Absolutism'. In eds., C.Kyriacou and R.McKenna, *Metaepistemology.* London, Palgrave Macmillan. pp.217-240.

MacFarlane John. (2016). *Assessment Sensitivity.* Oxford, Oxford University Press.

Nozick Robert. (1993). *The Nature of Rationality.* Cambridge MA, Princeton University Press.

Rinard Susanna. (forthcoming). 'Pragmatic Skepticism'. *Philosophy and Phenomenological Research.*

Ridge Michael. (2018). 'How to Be An Epistemic Expressivist'. In eds., C.McHugh, J.Way and D.Whiting, *Metaepistemology.* Oxford, Oxford University Press. pp.141-158.

Stich Stephen. (1990). *The Fragmentation of Reason*. MIT Press.

Streumer Bart. (2017). *Unbelievable Errors.* Oxford, Oxford University Press.

Unger Peter. (1975). *Ignorance.* Oxford, Oxford University Press.

Gabriel Malagutti
University of Lisbon

The structural flaw of a posteriori non-reductionism in the epistemology of testimony

Non-reductionism in the epistemology of testimony is the view that absence of defeaters suffices for testimonial justification. Lackey (2006) has put forward an objection against non-reductionism: the alien case. The case is one where an alien testifies that she had oranges for breakfast. She is both reliable and truthful, the human knows nothing about this kind of alien nor the planet of origin. Nonetheless, George forms the corresponding belief. Is George's belief justified? Lackey argues negatively. Given the lack of any sort of evidence, positive or negative, one ought to withhold belief. To push this interpretation, Lackey puts forward two kinds of sceptical concerns. One, regarding the alien's reliability, and the second, regarding the content of the assertion. Given the lack of evidence, one way or another, one is unable to dismiss these concerns, making testimonial acceptance not only unjustified, but irrational. This is problematic for non-reductionism for given the absence of negative evidence, one *should* be justified in believing the alien. If Lackey's assessment is correct, non reductionism has been falsified, for the absence of negative evidence is insufficient to justify testimonial acceptance.

There are two strategies available for proponents of non-reductionism. Either accept the absence of negative evidence, in which case one must provide some reason to dismiss the sceptical concerns; or deny the absence of negative evidence, enabling defeat, in which case one must provide a defeater. I'll pursue both venues, using the different accounts available:

Simion (2020)'s, Graham (2010)'s, and Goldberg (2010)'s. I will argue that their accounts are unable to do so. *A posteriori* versions of non-reductionism, such as the ones mentioned, are unable to successfully reply to the case. They either bring about reductionism, conceding the need of positive evidence, or are *ad hoc*.

The accounts are *a posteriori* due to the kind of reply given to the Source Problem. That is, given the no-defeater condition, justification seems to come too cheaply. Afterall, speakers may be mistaken or lying. The source problem is thus a challenge for non-reductionists to justify the no-defeater condition. Accounts such as the ones mentioned do so *a posteriori*. It's due to some feature of how testimony is employed that it enjoys such a standing. Simion appeals to social norms of truthfulness and reliability made in place by a social contract that motivates compliance. Goldberg appeals to *de re* reliance on the speaker, which is justified by expectations that come about due to social, political, and ethical considerations. Graham appeals to

assertion's e-function of producing true hearer beliefs. This function is satisfied reliably due to social norms that stabilize the use of assertion to make them sufficiently truthful and reliable. What connects these accounts is their appeal to background information, which is contingent and acquired empirically.

Given this framework, how can they reply to Lackey's case? Remember that testimony enjoys *prima facie* entitlement due to certain features about social norms, epistemic environments, and its normal conditions. All these features are contingent and come about from background information. As such, one is unable to justify acceptance in the alien case. The alien, due to being an alien, falls outside the domain of human testimonial background information. As Lackey states, humans lack any sort of evidence about the alien or its species. Lacking information about their environment, epistemic and social norms, and normal conditions, none of the accounts can preserve justified acceptance. It's important to point out that this doesn't mean that entitlement is defeated. Alien testimony lacks entitlement altogether. Entitlement comes from those contingent features that constitute human's background information. Is there a defeater that non-reductionists can appeal to?

I argue that the appeal to background information makes it impossible to formulate a defeater. Notice that the only feature non-reductionists can pick out in the case, is the fact that the speaker is an alien, or not-human. If one takes the former, then the reply is *ad hoc*, for its taking being an alien as negative evidence. The latter ultimately brings about Faulkner (1998)'s description of reductionism:

1. «testimony t is testimony of type T;

2. <u>type-T testimony has been established to be credible;</u>

3. therefore testimony t is credible.»

Replacing the variables to accommodate this picture makes it:

1. Testimony t is testimony of human-type;

2. <u>Human-type testimony has been established to be credible;</u>

3. Therefore, testimony t is credible.

The underlying problem is that human-testimony enjoys entitlement because human agents have positive evidence about it. It has established itself to be credible through background information. It's because of this that the alien's testimony poses a problem, for it hasn't been established to be reliable. The defeater can only come about due to absence of positive evidence, ultimately endorsing Faulkner's syllogism:

1. Testimony t2 is testimony of alien-type;

2. <u>Alien-type testimony has not been established to be credible;</u>

3. Therefore, one can't say that testimony t2 is credible.

Due to entitlement being linked to background information, non-reductionism turns out to be a framework about human-testimony. There are two ways to read this, both of which greatly undermine or falsify non-reductionism. The dilemma forces proponents to endorse one of two horns. Either take non-reductionism as a

normative theory of testimony *simpliciter*, in which case, the alien case is either unanswerable, or endorses Faulkner's reductionism. Or, to reduce its scope to be a framework about *human*-testimony. This is problematic for two reasons. Firstly, there is no scope restriction in how non-reductionism is presented. Backtracking when faced with the alien case seems a cheap way to save *a posteriori* non-reductionism. More worrisome is that making such a scope restriction is the same as endorsing reductionism as a normative account of testimony unrestricted. Notice that restricting the no-defeater sufficiency thesis is the same as applying Faulkner's reductionism to testimony *simpliciter*.

A posteriori non-reductionism thus faces a dilemma. Either take the account to be of testimony *simpliciter*, making it unable to reply to the alien case; or restrict it to an account of *human*-testimony. If one is to endorse the latter, reductionism emerges.

Emanuele Martinelli
University of Zurich

An Agent-Based Approach to Economic Planning

Mises (1963) and Hayek (1945) have crafted compelling arguments against central economic planning. For some time, these have been taken as definitive proof that a centrally plan economy managed by the government could be possible (Polanyi 2013). As a result, the market order has been declared the superior form of economic system. However, the exponential rise in the capacities of AI soon opened to the possibility that not human bureaucrats but supercomputers could have what it takes to plan the economy better than the aggregate action of market actors. The debate thus reignited, and the ball has been tossed back and forth quite inconclusively up to date. Arguably, this is because neither Mises nor Hayek have clearly and conclusively given a conceptual reason why central planning of the economy is impossible *in principle*. In this paper, I want to deliver such argument. I find that the most promising way to do it is to frame the problem of economic planning as an agent-environment interaction.

The core of the problem highlighted by Mises and Hayek is that the huge amount of information needed to plan the economy is never available to a single economic actor. Knowledge about people's preferences, needs, opportunities for innovation, availability to work, etc., is dispersed in society and cannot be retrieved by the central planner.

> The economic problem is thus not merely a problem of how to allocate "given" resources – if 'given' is taken to mean *given to a single mind* which deliberately solves the problem set by these 'data'. It is rather a problem of how to secure the best use of resources known to any of the members of society, for ends whose relative importance only these individuals know. Or, to put it briefly, it is a problem of the utilization of knowledge not given to anyone in its totality. (Hayek 1945: 519-520, emphasis mine)

The market order, on the other hand, can operationalize this knowledge because market competition does not require that any one economic actor knows about all the economic facts across society (Hayek 2016). This is realized through the 'price system', i.e. the distributed communication system that signals variations in the availability of certain resources, assets, or products as a result of economic actions and preferences (Hayek 1945; Mises 1996)

Much more has been said on both sides of the debate. But this much is enough to illustrate why, at some point, the socialist defendants of central planning have turned to AI and generally computational techniques to finally solve the calculation

problem. As Oskar Lange, the founding father of the so-called 'market socialism' solution, puts it:

> Were I to rewrite my essay today my task would be much simpler. […] Let us put the simultaneous equations on an electronic computer and we shall obtain the solution in less than a second. The market process with its cumbersome *tatonnements* appears old-fashioned. Indeed, it may be considered as a computing device of the preelectronic age. (Lange 2002: 1)

This depiction of the issues of economic planning nested into the Mises-Hayek thesis in essence concerns both the means and the ends of economic planning, or economic knowledge and the structure to convey and elaborate it. Now, that economic planning is *impossible* due to some of the above reason, as said before, can potentially mean two things. The impossibility can either be an empirical fact, meaning that we could in the future find ways to collect and manage all the data needed for economic planning e.g. with AI techniques, or that it is a conceptual impossibility and that the economic system is just not something that can be planned centrally. One could say, for instance, that the economy as a complex system cannot be planned because of x mathematical features, and yet that we might come up with an alternative conception of economic planning that accommodate them – perhaps through the wonders of quantum computing or general AI. Indeed, this is the route paved by the first responders to Mises' and Hayek's objections. Lange (1936) contended that 'they do not deny the *theoretical* possibility of a rational allocation of resources in a socialist economy, they only doubt the possibility of a satisfactory *practical* solution of the problem' (1936: 56). While his claim is debated by some commentators like Kirzner (1984), I do concede that both Mises and Hayek did not insist properly on the scope of their arguments. And thus much space for maneuver has been opened to subsequent proponents of central planning that tried to solve the practical impossibility through the wonders of AI development.

In §1, I will sum up the 'Mises-Hayek thesis' as a multi-faceted objection against economic planning, and I will highlight the problem of qualifying the claimed impossibility. In §2, I will introduce my own approach to the problem of economic planning, which is 'agent-based' in the sense that it moves from an ontology of the agents involved in the workings of the economy. I will here advance a general ontology of agents and then apply it to the problem at hand to show the different sets of agents that are considered a) in a market economy scenario where the price system is free to operate, and b) in a centrally planned economy equipped with the most sophisticated AI technology. In §3, I will present the most respected models that try to conceptualize a centrally planned economy: the Lange-Lerner model and the Cockshott-Cottrell model. In §4, I deliver my ontological argument against central economic planning, claiming that public institutions as planning bodies can substitute the market order, no matter the AI technology that aids and support them. The reason is that the elimination of the market entails the elimination of crucial kinds of agents that cannot be recreated through AI and careful social planning. More precisely, I will highlight the limitations of the socialist models presented in §3, based on the fact that public managers or AI systems that

automatize their decisions cannot replace the proactive action of entrepreneurs driving market allocation.

Maurizio Mascitti

Vita-Salute San Raffaele University

Fake news as violation of journalistic norms

Habgood-Cote (2019, 2022) recently argued that the philosophical community would benefit from omitting the term 'fake news' from its vocabulary. One of the reasons he presents is that 'fake news' signifies an unnecessary concept; existing terms adequately cover deceit and miscommunication. Furthermore, these terms tend to have clearer and more stable meanings in ordinary language, making them less prone to misuse. Habgood-Cote concludes that, given the slow and laborious nature of meaning change, we should abandon the concept of fake news.

While the eliminativist stance of Habgood-Cote may seem stringent, I find his diagnosis justified based on the current definition of fake news. Grundmann (2020) observes that the majority of accounts in the literature are hybrid views. These views rest on two assumptions: firstly, news, including fake news, is essentially propositional content communicable through various mediums like articles, pictures, and videos. Secondly, fake news must meet subjective and objective conditions. The subjective condition requires the producer to have deceptive intentions, while the objective condition involves the propositional content being literally false or capable of conveying something false through implicatures. The hybrid approach defines fake news as instances that meet both conditions, but this framing prompts the question: why retain the term 'fake news' when other concepts like disinformation already cover similar cases? Disinformation, typically referring to misleading content spread with the intent to deceive, already encompasses much of what hybrid definitions attribute to fake news, as do other terms like misinformation and propaganda.

I maintain that the first assumption of the hybrid model is inadequate and does not serve the purpose of defending fake news as a necessary concept. The notion of news employed by hybrid accounts is in fact so minimalist that it boils down to its propositional content plus a handful of other properties taken from the dictionary definition. Most of the time hybrid accounts differ only in the way in which they frame the subjective and the objective condition of the second assumption, leaving aside the question of what the real contribution of the term "news" in "fake news" is. Hybrid accounts treat fake news, and so news in general, only as a species of information. As Pritchard (2021) explicitly says, news «is *a kind of information* that one gets from a news source» (p. 52, emphasis mine), where information is something that possesses a semantic content and thus represents some part of the world as being a certain way. However, this assimilation to information does not render justice to the notion of news – and therefore also to fake news.

I argue that news and information are in fact two separate entities. Information is an epistemic entity, while news in my view is a sociological entity. The former has to do

(only) with meaning and truth values, and it has a structural connection with the concept of knowledge. That is why denying truth to information ends up fading it (Dretske 1983, Floridi 2010). News instead is an entity of a different kind; something that does not have a structural connection with knowledge. News is object that undergoes a certain process of collective regulation. The newscast report is in part the outcome of an accurate activity that is collectively carried out by agents of an acknowledged human institution; and this is journalism. Thus, news is a sociological entity inasmuch as it responds to some settled norms of journalism that characterize their process of production and publication. Consequently, in this paper I define fake news as any news item that significantly violates the norms of journalism.

Journalism stands as a dynamic social institution, with its members steadfastly dedicated to upholding a set of norms aimed at ensuring the seamless delivery of information. As posited by some scholars (Pepp et al., 2019), this collection of norms, encompassing commitments to reader interests and scrupulous information verification, serves as background expectations that readers harbor for the journalistic enterprise as a whole. These norms predominantly govern news production, constituting a crucial foundation upstream of the information delivery process. Yet, the distinctive character of journalistic enterprise extends beyond the norms of news production. I contend that a comprehensive array of conventions is necessary to regulate how information is ultimately presented to the reader, and these can be termed norms of news publication. Broadly speaking, we can view norms of news presentation as akin to Gricean maxims, albeit with some deviations. The pivotal point lies in recognizing that every significant violation of norms in news presentation constitutes an instance of potential fake news too – an aberrant manner of framing and presenting information to the reader.

Therefore, defining fake news as news that transgresses established norms of journalism, whether in production or publication, enables us to disentangle this concept from the realms of misinformation, disinformation, and their cognates. This distinction is crucial, as merely false news should not automatically be labeled as fake news, and vice versa. The critical determinant lies in how that false piece of news is presented to the audience. Only when it also violates the ideals of news journalism can we appropriately categorize it as fake news. In the absence of such normative breaches in presentation, it remains a case of misinformation (or disinformation). This nuanced approach to conceptualizing fake news, grounded in normative violations, preserves the conceptual necessity of the term. It underscores the importance of distinguishing between various forms of misleading information, aligning more closely with the intricacies of journalistic practices and their impact on the public's perception of reality.

References:

Dretske I. F., 1983, «Précis of Knowledge and the Flow of Information», *Behavioral and Brain Sciences*, 6, 1, pp. 55-90.

Floridi L., 2010, *Information: A very short introduction*, Oxford University Press, Oxford.

Grundmann T., 2020, «Fake news: the case for a purely consumer-oriented explication», *Inquiry*, 66, 10, pp. 1758-1772.

Habgood-Coote J., 2019, «Stop Talking About Fake News!», *Inquiry*, 62, pp. 1033–1065;

Habgood-Coote J., 2022, «Fake news, conceptual engineering, and linguistic resistance: reply to Pepp, Michaelson and Sterken, and Brown», *Inquiry*, 65, 4, pp. 488-516.

Pepp J., Michaelson E., Sterken R., 2019, «What's New about Fake News?», *Journal of Ethics and Social Philosophy*, 16, 2, pp. 67-94.

Pritchard D., 2021, «Good News, Bad News, Fake News», in Bernecker S., Flowerree A., Grundmann T. (eds.), *The epistemology of fake news*, Oxford, Oxford University Press, pp. 46-67.

Martí Bridgewater Mateu
University of Barcelona

Just how much obliged? Questions concerning deontic vagueness

Most contemporary literature on vagueness and the Sorites paradox concerns itself with logically interesting but normatively spurious concepts, such as baldness and tallness. It seems methodologically sound to study vagueness by means of examples whose vagueness is their only logically puzzling characteristic, and yet many vague concepts are also philosophically complex for other reasons, such as ethical notions like "good" and "just", but it is difficult to understand how these different conceptual difficulties interact. There is some work dealing with vagueness in the realm of ethics, but it tends to focus on issues of metaethics and value comparability, and it engages with the details of the relevant literature in logic and semantics to varying degrees (Asgeirsson, 2019). In this communication, I will offer some considerations on the interplay between different theories of vagueness and accounts of ethical obligation (broadly construed).

It is common to divide contemporary normative ethics into three main schools: consequentialism, deontology and virtue ethics. Of the three, the latter is the least concerned with the notion of obligation, although it is perhaps the one in which vagueness is most prominent as a phenomenon, and in fact it has been argued that Eubulides's original formulation of the Sorites paradox may have been prompted by the issue of the slippery distinction between virtues and vices in the scalar continuum of morally relevant character traits in Aristotle's ethics (Moline, 1969). The concept of obligation, however, is often sidelined or just altogether avoided in the contemporary literature on virtue ethics (Hursthouse & Pettigrove, 2023), and so it is more appropriate to focus analysis on the other two schools.

Obligation is arguably right at the conceptual core of deontological ethics, insofar as it is generally taken to be identical to or closely linked with the notion of duty. In the theoretical framework of deontology, a given agent has certain duties that it is obligatory for them to fulfil in whichever way they choose to, with all courses of action in compliance with these obligations remaining permissible. In this section I will discuss two main problems in deontological ethics in which vagueness features prominently, drawing on some preliminary work on the role played by degrees in deontological theories (Hurka, 2019). One concerns which issues matter enough to merit the establishment of duties, which I believe underlies the often-criticised arbitrariness apparent in Ross's canonical list of duties and other attempts in the same vein (1939). The other issue concerns which courses of action are taken to be compliant with a pre established list of duties: granted that one has a duty of

rescue in an accident, for instance, one can easily construct a version of the Sorites paradox concerning different possible responses and whether they satisfy this duty, with different views on the appropriate solution to the paradox involving different claims about the satisfaction of duties.

Consequentialist theories, on the other hand, derive their claims concerning obligation from claims concerning value, insofar as value is taken to be the sole and sufficient basis for normativity. Preexisting debates concerning the way in which value judgements relate to obligation are framed around the so-called "overdemandingness objection", which is generally settled in terms of an obligation to value-maximisation, with alternative satisficing and scalar views remaining mostly marginal in the field (Norcross, 2020). I will attempt to show that underlying this debate is fundamentally an issue of vagueness concerning how much relative value-fostering merits talk of obligation. I will further argue that this renders maximising deeply implausible *qua* implicit theory of vagueness, and prompts new difficulties for remaining alternative views, which intersect with questions about the plausibility of competing claims concerning the appropriate analysis of the Sorites paradox.

Finally, I will offer some more generalised discussion on the appropriate framework for thinking about the logical modelling of normativity about action in the realm of ethics. I will discuss some technical and philosophical difficulties that arise in attempting to resolve these issues in terms of a non-classical deontic logic, i.e., the assigning of degrees of truth, truth gaps, truth gluts and the like to deontic claims. I will also briefly discuss connections between these issues and alternative formulations of normativity about action in terms of questions concerning the notion of truth-bearing in the logic of imperatives (Vranas, 2013).

References

Asgeirsson, H. (2019). The Sorites Paradox in Practical Philosophy. In S. Oms & E. Zardini (Eds.), The Sorites Paradox (pp. 229–245). Cambridge University Press.

Hurka, T. (2019). More Seriously Wrong, More Importantly Right. Journal of the American Philosophical Association, 5(1), 41–58. doi:10.1017/apa.2018.41

Hursthouse, R., & Pettigrove, G. (2023). Virtue Ethics. Retrieved from The Stanford Encyclopedia of Philosophy website: https://plato.stanford.edu/archives/fall2023/entries/ethics-virtue/

Moline, J. (1969). Aristotle, Eubulides and the Sorites. *Mind*, *78*(311), 393–407.

Norcross, Alastair (2020). Morality by Degrees: Reasons Without Demands. Oxford University Press.

Ross, David (1939). Foundations of Ethics. Oxford University Press.

Vranas, P. B. M. (2013). Imperatives, Logic Of. In H. LaFollette (Ed.), The International Encyclopedia of Ethics (pp. 2575–2585). Blackwell.

Guido Melchior

University of Graz

The sensitivity of legal proof

The gatecrasher paradox results from conflicting intuitions concerning different types of fallible evidence in a court of law. We accept fallible individual evidence but reject fallible statistical evidence even when the conditional probability that the defendant is guilty given the evidence is the same. These intuitions create a puzzle or paradox because in legal theory, evidence quality is supposed to be closely tied to the conditional probability of an event having occurred given the evidence at hand.

This paper defends a solution to the gatecrasher paradox, building on a sensitivity account of checking and settling a question. The proposed sensitivity account of legal proof not only requires sensitivity simpliciter but sensitivity of each relevant step of the proof procedure and/or sensitivity concerning all relevant alternatives. This account avoids problems that have been identified for other sensitivity views of legal proof. Moreover, it is argued that sensitivity, rather than safety, is the crucial modal condition for legal proof.

Let me explain in more detail. Various solutions to the gatecrasher paradox have been suggested in the literature. A well-received proposal is sensitivity-based. Enoch et al. (2012) suggest that a principle very similar to the epistemic principle of sensitivity plays a crucial role in explaining the gatecrasher puzzle. Sensitivity is a modal feature of beliefs that Nozick (1981) introduced as a necessary condition for knowledge. S's belief that p is sensitive if, in the nearest possible worlds where p is false, S does not believe that p. Enoch et al. (2012) point out that statistical evidence is typically insensitive whereas individual evidence tends to be sensitive. Hence, sensitivity, respectively an action-guiding version of it, marks a distinction between these two kinds of evidence and allows for a solution to the gatecrasher puzzle. Sensitivity is a very controversial condition on knowledge and accordingly the account developed in Enoch et al. (2012) has been criticized from various perspectives and for various reasons. (See Blome-Tillmann (2015), Gardiner (2019), and Moss (2022).)

In this paper, I will defend epistemic sensitivity accounts of standards of legal proof. The account that I favor is based on a sensitivity account of checking and settling a question, a view developed in detail in Melchior (2019). According to this account, checking requires entertaining a proposition and intentionally using a method for determining whether the target proposition is true. Moreover, appropriate methods for checking and settling a question must be specified modally. Weak sensitivity concerning p is the crucial necessary condition for checking that p is true according to Melchior (2019). M is not an appropriate method for checking whether p is true, if M would indicate that p is true if p were false. Safety, in contrast, does not play a crucial role in theories of checking according to this account.

By using this account, I thereby position this account of legal proving in a larger context of sensitivity theories of norms for checking and settling a question. Importantly, the proposed sensitivity account of legal proof not only requires sensitivity simpliciter but sensitivity of each relevant step of the proof procedure and/or sensitivity concerning all relevant alternatives. This account avoids problems that have been identified for other sensitivity views of legal proof. Cases of sensitive statistical evidence raised against sensitivity accounts of legal proof still fail to be appropriate methods for proving in court because not each relevant step of the process is sensitive and/or the overall process is not sensitive concerning all relevant alternatives. Furthermore, the factivity problem of sensitivity, which implausibly rules out faultless false convictions based on misleading evidence, is avoided because the proposed account of sensitive *methods* does not suffer from the factivity problem. Moreover, it is argued that sensitivity, rather than safety, is the crucial modal condition for legal proof.

To summarize the central line of argumentation, there is a natural connection between sensitivity and legal proof. Sensitivity is a crucial mark of checking and settling a question and trials are paradigmatic instances of settling a question. Therefore, sensitivity is also a crucial mark of legal proof. Various questions concerning legal proof are still hotly debated, for example its relationship to knowledge or the possibility of faultless wrongful convictions. The provided sensitivity account can make a substantial contribution to all these accounts and topics. It might be an open question which role exactly sensitivity plays for trials and the law, for example, whether it is part of an incentive-oriented and action-guiding system of convictions (Enoch et al (2012) or whether it is a principle of ruling out relevant alternatives (Gardiner (2019b) and Moss (2022)). The provided account can fruitfully support any of these views on the relationship between knowledge and legal proof, a fact that provides further support that it captures a crucial feature of legal proof.

References

Blome-Tillmann, Michael (2015). Sensitivity, causality and statistical evidence in courts of law. *Thought* 4(2), 102–112.

Enoch, David, Fisher, Talia and Spectre, Levi (2012). Statistical evidence, sensitivity, and the legal value of knowledge. *Philosophy and Public Affairs* 40(3), 197–224.

Gardiner, Georgi (2019a). Legal Burdens of Proof and Statistical Evidence. In James Chase & David Coady (eds.), *The Routledge Handbook of Applied Epistemology*. Routledge, 179-195.

Gardiner, Georgi (2019a). The reasonable and the relevant: Legal standards of proof. *Philosophy and Public Affairs*, 47, 288-318.

Melchior, Guido (2019). *Knowing and Checking: An Epistemological Investigation*. Routledge.

Moss, Sarah (2022). Knowledge and legal proof. *Oxford Studies in Epistemology, Volume 7*, 176-213.

Nozick, Robert (1981). *Philosophical Explanations*. Harvard University Press.

Ismael Miguéns

University of Santiago de Compostela

Everything, step by step. Expansionism and mathematical richness

Can speakers achieve absolute quantification? Informally —and twisting our means of expression—, this question refers to the ability of speakers to quantify over absolutely everything there is. Absolutists answer with a resounding "yes". According to them, even if this is not always the case, speakers sometimes achieve this extreme form of generality. In turn, relativists deny this optimistic situation.

The debate between absolutists and relativists has long been animated by the discovery of several paradoxes affecting the naïve theory of sets. The weak assumptions needed for Russell's, Burali-Forti, and similar paradoxes to arise suggested that a deep revision of the state of the art concerning the foundations of mathematics was required. And, beyond their differences, all those proposals developed to fix the situation agreed on a particular point of the diagnosis: the classic, impredicative formulation of the Naïve Comprehension Schema (NCS)

$$\exists x \forall z (z \in x \leftrightarrow \phi(z)),$$

—where x does not occur free in $\phi(z)$ — should be rejected. After all, the correlation between sets and concepts it postulated was the cause of the contradiction – more specifically, the schema has contradictory instances.

However, once the standard formulation of NCS is rejected the following dilemma arises: either the comprehension schema is abandoned as an engine of set production, or we accept a liberal attitude towards collecting. Absolutists have been irremediably associated with the first horn of this dilemma. The existence of a comprehensive domain to which every set belongs and in which our quantifiers are interpreted leads to the complete abandonment of NCS. As soon as the universe of sets is constituted as the domain of first-order quantifiers, Russell's paradox tells us to let NCS go. In turn, relativists were bounded to the second horn and, with it, to a subtle restoration of the latter principle. Let's say that a Comprehension Schema is predicative just in case —at least some of— the objects it introduces

> (i) fall outside the domain of any quantifier in the condition $\phi(z)$,

> (ii) and outside the domain of the universal quantifier in the prefix.

Then, if we deny the existence of an absolute domain to which every set belongs — as the relativist does—, the predicative version of NCS can be safely adopted. As before, given a collection of sets A in which the first-order quantifiers in (i) and (ii) are interpreted, it asserts the existence of another set b whose members are exactly

those members of S satisfying a given property. However, now the set b does not necessarily belong to A. And, by letting the domain to expand, paradoxes are avoided. Against appearances, Russell's paradox is not just a matter of logic: it presupposes that speakers are fortunate enough to quantify over a comprehensive domain of sets —and thus, that sets form such a domain. By settling the debate on paradoxes, absolutists and relativists hoped to undermine their "enemy".

Relativists tend to see the adoption of a predicative version of NCS both as a solution to the paradoxes and as a reason for their position. Indeed, this expansionist process is generalized by the Dummettian (1981) notion of indefinite extensibility: informally, a concept is indefinite extensible if and only if for every domain containing only elements satisfying the former, there is another object (a) which satisfies the concept too, but (b) does not belong to the domain. The concepts of set, ordinal, and interpretation are regarded as paradigmatic examples of this phenomenon. And, the version of relativism that arises from the acceptance of a predicative comprehension schema —in general, from the acceptance of indefinite extensible concepts— is known as expansionism. According to expansionists, we cannot quantify over absolutely everything because we always can perform an expansion of the domain. Even though expansionism is in good condition and has been advocated by several philosophers and mathematicians, the truth is that it has not been the main course. Also, it has been associated with constructivist positions and mathematical revisionism. By default, absolutism has been accepted as the "standard" solution to the paradoxes: sets form a universe which is described by ZF set theory — where NCS is replaced by the Separation Schema. Thus, failing the test of mathematical naturalness, expansionism took a back seat.

However, expansionism has been recently revitalized by a series of books and papers (Linnebo, 2018; Stud, 2019) as a firm guard against paradoxes and a serious explanation of classical mathematics. In this talk, I will follow the direction pointed by these works and argue in favour of expansionism. First, I will argue that expansionism is supported by classical mathematical reasoning. Far from being a philosophical extravagance, expansionism and indefinite extensibility are part of our understanding of certain hierarchical domains of objects —such as sets, ordinals, or interpretations. The argument will be split in two steps:

> (1) I will show how Gödel's (1995) bootstrapping account of the iterative hierarchy of sets can be explained along the expansionist lines,

> (2) I will assess the options of explaining some modest large cardinal hypotheses—reflection and extensibility principles— in terms of domains expansions generated by the indefinite extensible nature of the concept of set.

As a result, not only is expansionism not alien to classical mathematics, but it also becomes a source of mathematical richness. Of course, the iterative conception and large cardinal hypotheses are consistent with absolutism. Rather, the point is that the latter cannot offer any justification for them on its own. Second, I will argue that expansionism —if it wants to be a support for classical mathematics— must be reconciled with absolute quantification. What we need, then, is a mild position between these two antagonistic conceptions. At this point, I will consider the intensional approach to quantification developed by Fine (2006) and Linnebo (2018) and cast some doubts on it based on two points: the recent criticism posed by Stud

(2022) and the paradoxes caused by semantic optimism (Williamson, 2003). I will finish by pointing out that an optimal solution to these challenges might be to differentiate between sets and other hyperintensional notions of collection —such as proper classes.

References

Dummett, M. (1981) Frege: Philosophy of Language. Harvard University Press.

Fine, K. (2006). Relatively unrestricted quantification. In Rayo, A. and Uzquiano, G. (eds.), Absolute Generality. Oxford University Press.

Gödel, K. (1995). Conference at Göttingen. In Feferman, S. (ed.), Collected Works, Volume III. Unpublished Essays and Lectures. Oxford University Press, pp. 127 – 155.

Linnebo, Ø. (2018). Thin Objects. An Abstractionist Account. Oxford University Press.

Studd, J. (2019). Everything, More or Less. A Defence of Generality Relativism. Oxford University Press.

Studd, J. (2022). Linnebo's abstractionism and the bad company problem. Theoria, 89(3): 366-392.

Williamson, T. (2003) Everything. Philosophical Perspectives, 17(1): 415-465.

Vittorio Morato

University of Padua

In defence of monism about metaphysical modality

Deflationism about metaphysical modality is the view that metaphysical modality should be deflated, reduced, explained away in terms of other non-metaphysical forms of modality or to some non-modal dimension altogether.

Even after the great renaissance of metaphysical modality due to the publication of *Naming and Necessity* in the 80s, there have been various deflationary projects in contemporary metaphysics. Among them, one might cite new forms of conventionalism (Sidelle (1989) or, more recently, Warren (2020)) and bi-dimensionalism. According to former metaphysical modality should be reduced to analyticity, according to bi-dimensionalism to the *a priori* or, more generally, to the conceptual dimension (see Tahko (2013)). In case, the *a priori* is further analyzed in terms of analyticity the two deflationary project tend to collapse. In both cases, the idea is that there is nothing distinctively metaphysical about modality, but there is something that metaphysical modality corresponds to in the end (see Soames (2005)).

Quite recently, however, another form of deflationism has emerged (Clarke-Doane (2019, 2021), Priest (2021)). According to "new deflationism", metaphysical modality can be deflated not by reducing it or to explain it away in terms of something else (modal or non-modal), but because there is *nothing special* about it. In particular metaphysical modality should not be considered special for its being the more inclusive, objective, non-epistemic, non-deontic kind of necessity. The project of finding out what metaphysical modality really is is misconceived.

There are various strategies this new form of deflationism might adopt. They are somewhat related to one another, even though sometimes they are not very clearly distinguished. The aim of the my talk will be to identify such strategies and criticize them.

Some preliminary definitions:

Broadness: $\Box_1$ is broader than another notion $\Box_2$ iff for every P, $\Box_1 P \rightarrow \Box_2 P$ and there is some Q such that $\Box_2 Q \land \neg\Box_1 Q$

(a necessity$_1$ is broader than necessity$_2$ in case everything that is necessary$_1$ is also necessary$_2$ but there is some proposition that is necessary$_2$, but it is not necessary$_1$.)

Absoluteness: $\square_1$ is an absolute notion of necessity iff for every $\square_i$, if $\square_1 P$ then $\square_i P$

(a necessity$_1$ is absolute iff something necessary$_1$ is necessary for any other kind of necessity)

Here is a brief summary of the three strategies:

Indefinite extensibility. As every kind of modality, metaphysical modality is defined with respect to some truths. The list is actually open, but at the minimum one might say that what is metaphysically necessary depends on some essentialist truths, mathematical truths, mereological truths, etc. Now assume that $\square_T$ is metaphysical necessity wrt to some set of metaphysical truths T and, for some P, $\square_T P$ due to some $t \in T$. The idea of new deflationists is that $\square_T$ is indefinitely extensible. Given $\square_T$, we can always define a broader notion $\square_{T-t}$. If this is so, metaphysical modality is not absolute and for every fixed notion of metaphysical modality we will be always able to define a broader notion.

Indeterminacy. A variant of the first strategy is the following. Modal metaphysics is about what could have been, but there seem to be many interpretations of "could have been". If the extensibility presented above works, for any alethic, objective sense of possibility for which, for example, you could *not* have different parents, there is another, alethic, objective sense for which you *could* have different parents. So even if it makes sense to ask, under the first interpretation, whether a proposition is true, the question is misconceived in general, because there is no "correct" way of understanding the notion of "could have been". This indeterminacy pushes towards some form of modal pluralism, analogous to other forms of pluralism in logic, set theory or the philosophy of mathematics (see Clarke-Doane and McCarthy (2023) for a similar argument).

Fineness of grain. Assume that we have a conception of metaphysical necessity and a conception of what propositions are according to which there might be two distinct propositions that are metaphysically necessary equivalent (every fine-grained conception where propositions are not simply sets of worlds will do). Assume that such propositions are P and $\top$ (a tautologous proposition). If this is so, we could in principle define a necessity operator O that applies to $\top$ and not to P. Given that, by definition, we have $O\top$, O is a normal necessity operator. But then O is broader than metaphysical necessity, because $\top$ is also metaphysically necessary, but P is not O-necessary. The same reasoning can be repeated for O as long as we define different or even more fine-grained conception of propositions (see Bacon (2018) for a similar argument).

My aim will be that evaluate and, in the end, criticize each of the following strategies and thus modal deflationism.

Against **indefinite extensibility** I will show, for example, that, given a notion of necessity $\Box^T$ defined wrt to a set of truths T, not every modification of this notion ($\Box^{T-t_i}$) automatically result in an extension of the previous notion. For example, metaphysical necessity minus the necessity of identity corresponds to a modal logic where variables range over intensional objects. The corresponding notion of necessity is thus more of a conceptual kind, rather than metaphysical and it is not very clear that conceptual necessity can be taken as a broader notion than metaphysical necessity.

Against **fineness of grain**, I will claim that it is not clear whether the process of "fine-graining" of a proposition is something that can be carried on indefinitely. Furthermore, from a methodological point of view, is quite strange to "discover" broader notions of necessity by means of a refinement of the notion of proposition. What we have here seems to be a case of reification of our tools of analysis.

I will in the end be more careful in the overall evaluation of **indeterminacy** (*pace* the arguments against **indefinite extensibility** on which **indeterminacy** somewhat depends) but I will nonetheless point to some important disanalogies between, for example, set-theoretic pluralism and pluralism about modal metaphysics.

References

Bacon, A. (2018). The broadest necessity. *Journal of Philosophical Logic*, *47*, 733-783.

Clarke-Doane, J. (2019). Modal objectivity. *Noûs*, *53*, 266–295.

Clarke-Doane, J. (2021). Metaphysical and absolute possibility. *Synthese*, *198*, 1861–1872.

Clarke-Doane, J., & McCarthy, W. (2023). Modal pluralism and higher-modal logic. *Philosophical Perspectives, Early view*.

Priest, G. (2021). Metaphysical necessity: a skeptical perspective. *Synthese*, *198*, 1873–1885.

Sidelle, A. (1989). *Necessity, essence and individuation: a defence of conventionalism*. Ithaca: Cornell University Press.

Soames, S. (2005). *Reference and description*. Princeton University Press.

Tahko, T. (2013). Soames's deflationism about modality. *Erkenntnis*, *78*, 1367–1379.

Warren, J. (2020). *Shadows of syntax*. Oxford University Press.

Hamza Naseer
USI Lugano

Amodal Completion leads to perceptual beliefs

Helton and Nanay (2023) argue that non-perceptual beliefs cannot be delineated from perceptual beliefs. In my model, however, there is a case being made that amodal completion requires familiarization to occur. The question that arises is: when we perceive something via amodal completion, does it share that content with what is seen? Note: there is no "seeing" in amodal completion; after all, what is being presented is the completion of the occluded bits of the object in question. In actuality, one has not seen the object one is completing. Can one share content with what they have not seen? The radical answer is yes. After all, IF there is a mind that has arrived at perceptual belief X, and X happens to coincide with the external object X, then it does not matter what the origins of perceptual belief X are - there can be sharing with the object without direct interaction with the object. However, I will go on to argue more in favor of the "shared object approach" (Helton and Nanay, 2023, pg. 95) and highlight that the indirect relationship between external object X and perceptual belief X is suggestive of their shared content.

Helton and Nanay criticize the shared object view by stating that it is "too strong with respect to beliefs emanating from non-visual modalities, in that it excludes too many of these beliefs from the class of perceptual belief. In particular, it makes the class of perceptual belief too impoverished to render the concept of perceptual justification feasible." (ibid) Their argument is as follows: when one has an auditory belief, for example, a belief that the noise of a passing train is too loud, one is forming a belief not about the object (the passing train) but about the noise (which is, according to them, not the object). This would then mean that most of our beliefs are about sounds or odors and not the objects that elicit them. I object to this, owing to the fact that I think the sounds or odors of objects are very much the properties of the object. If I can argue that they are the properties of the object, then I can argue that whatever beliefs we form about the odors or sounds of an object are very much beliefs about the object. This would then mean that

under the "shared object" view, we have objects that do share content with our perceptual beliefs about them. Kulvicki (2015) highlights the 'stable disposition view' according to which "sounds are objects' dispositions to vibrate in response to mechanical stimulation." (Kulvicki, 2015, pg. 206). This would mean that sounds are "qualities of objects." (ibid) In fact, one can think of sounds as parts of objects by the notion of the specific persistence of these sounds in tandem with the objects. If I have a guitar, and I strum my fingers along it in a particular way, it produces one note. If I do the same thing, it produces that same note again. If, all conditions remaining the same, I do this a million times, it will create the same note a million

times. The same holds true for other objects that elicit sounds i.e. all objects. As a result, I do not think that prescribing to the shared object view holds any strict problem for philosophers since whenever they refer to the sound of the object, they are still referring to the object. Furthermore, Helton and Nanay suggest that for the visual modality, this shared object view does allow for more perceptual beliefs. Even for visual states, I think we look at the properties of the object (the redness of the apple, the shape of the apple, and so on) from which we then form a holistic picture of the apple. My point being: all modalities more or less rely on the properties of objects to form holistic perceptual beliefs about the objects, so I don't see how the 'shared object' view in particular needs to be singled out as not allowing for a multiplicity of perceptual beliefs about objects across sense modalities.

Another question by Helton and Nanay is whether the content of the perceptual state is the same as the content of the perceptual belief. They suggest that for the 'shared content' approach to hold true, one would have to say they are either the same, or, at the very least, "they only need to be sufficiently similar." (Helton and Nanay, 2023, pg. 97) They argue that is not likely, owing to the differences that exist "between the way perception represents and the way beliefs represent." (ibid) Nonetheless, I don't think the differences are so extreme as to deny similarity altogether. One may envision perceptual representation and perceptual beliefs as entities speaking two different languages that are still able to communicate through some shared meaning-making ability. In fact, there must be a communication of this kind as otherwise, it would be difficult to have a coherent concept of a perceived object. Under a representationalist framework, for example, (see Nanay 2015) we know that multimodality is suggestive of a form of communication between all sense modalities that conspire together to complete our picture of the world. Similarly, I argue that perceptual representations and perceptual beliefs conspire together in order to comprehend the shared content they receive from the object. If the visual and the auditory can work together, so can, for example, a perceptual representation that is spatial and a perceptual belief that is propositional.

References

Helton, G., & Nanay, B. (2023). Against the very idea of a perceptual belief. *Analytic Philosophy*, 64, 93–105. https://doi.org/10.1111/phib.12277

Kulvicki, J. (2015). Sound stimulants: defending the stable disposition view. In Dustin Stokes, Stephen Biggs & Mohan Matthen (eds.), *Perception and Its Modalities*. New York, NY, USA: pp. 205-221.

Nanay, B. (2015). Perceptual Representation / Perceptual Content. In Mohan Matthen (ed.), *Oxford Handbook for the Philosophy of Perception*. Oxford University Press. pp. 153-167.

Alain Pe-Curto
USI Lugano/EPFL

Rethinking Value: An Aristotelian Challenge and Two Distinctions in Goodness

Korsgaard [1983] famously contrasted two distinctions in goodness: the distinction between extrinsic and intrinsic value versus the distinction between instrumental and final value. Moreover, Korsgaard, Kagan, Rabinowicz, and Rønnow-Rasmussen identified now classic cases of *final extrinsic* value. The cases suggest that final value and intrinsic value are distinct (cf. Tucker [2016]). They famously involve mink coats (Korsgaard [1983, 185]), a pen of Lincoln's (Kagan [1998, 285-6]), and Princess Diana's attire (Rabinowicz and Rønnow-Rasmussen [2000, 41 ff.]).

Final value obviously matters for moral philosophy: it is the value that something has as an end and the intuitive opposite of instrumental value. In contrast, intrinsic value feels like an arbitrary import from metaphysics into ethics (see, e.g., Kagan [1998, 290]). Thus, the now forty-year-old thesis that *final value* presents greater "normative interest" than intrinsic value has become greatly popular—as Tucker [2018, 131 fn. 2] notes (see Kagan [1998, 290-3], Olson [2015], Rabinowicz and Rønnow-Rasmussen [2003, 393; 2000, 48-9], and Rønnow-Rasmussen [2021, 3-23, 142-52, 2015, 2011]).

In contrast, I argue that *intrinsic value* is the normatively central category of moral philosophy. Before Korsgaard's distinctions and cases of final extrinsic value were introduced, my claim constituted axiological orthodoxy. G. E. Moore [1903] and his legacy offer a prime example of the relevant orthodoxy. Surprisingly, I show that Korsgaard's distinctions and the classic cases of final extrinsic value actually vindicate the past orthodoxy that intrinsic value stands at the core of moral philosophy.

1. An Aristotelian Challenge and Two Responses

In the *Nicomachean Ethics* [1.3.1094a18-26], Aristotle notes that the following regress would occur if all value were instrumental:

> **Instrumental Value Regress:** The instrumental value of *X* derives from the instrumental value of *Y*, which derives from that of *Z,* and so on to infinity.

Aristotle believes that our desires and, hence, our practical lives would be "empty and vain" if this regress occurred. We would be like archers without a mark. Thus, the *Nicomachean Ethics* begins with a challenge:

Aristotelian Challenge: We must identify the good(s) able to stop the Instrumental Value Regress.

For now, let us say that a value property—such as intrinsic or final value—is normatively central only if it plays the following central role in addressing the Aristotelian Challenge:

Normative Centrality: Value property *F* is normatively central *only if* all and only goods able to stop the Instrumental Value Regress instantiate value property *F*.

The first response to the Challenge that I consider is the Traditional Response:

Traditional Response: All and only good(s) instantiating intrinsic value are able to stop the Instrumental Value Regress.

The Traditional Response entails that intrinsic value is normatively central—as past orthodoxy would have it. As noted, though, the Telic Response rapidly gained traction after Korsgaard's two distinctions and cases of final extrinsic value were introduced.

Telic Response: All and only goods with *final* value are able to stop the Instrumental Value Regress.

The Telic Response entails that final value rather than intrinsic value is normatively central.

2. A New Aristotelian Challenge and the Metaphysics of Intrinsic Value

I present two arguments **(A-B)** to defend the past orthodoxy that intrinsic value stands at the core of moral philosophy.

(A) First, I show that the classic cases mentioned all instantiate a specific final extrinsic value, namely symbolic value, and that the following regress would occur if there were only symbolic value:

Symbolic Value Regress: The symbolic value of *X* derives from the symbolic value of *Y*, which derives from that of *Z*, and so on to infinity.

The Symbolic Value Regress resembles the Instrumental Value Regress. Plausibly, our practical lives would be "empty and vain" if all value were final symbolic value—as Aristotle noted that our practical lives would be empty and vain if all value were instrumental value.

This result reveals that the Aristotelian Challenge and Normative Centrality as defined above unduly focus on the Instrumental Value Regress. They neglect other regresses, like the Symbolic Value Regress, that could also make our practical lives empty and vain. Therefore, I introduce a New Aristotelian Challenge and a New Normative Centrality test that account for various value regresses. Moreover, I address objections that take final extrinsic values other than symbolic value—such as personal value, rarity value, and the value of knowledge—to be immune from such regresses (see Rabinowicz and Rønnow-Rasmussen [2003, 393-5 and fn. 10], Rønnow-Rasmussen [2021, 2011]; cf. Lemos [2022] and Moore [1903, §113- 20, 237-48]).

(B) Second, I argue that intrinsic value is not an arbitrary import from metaphysics. Extrinsic value (and its types, like symbolic value) depends on at least one external relation between its bearer and something distinct from its bearer. This dependence allows precisely for the value regresses examined— whether the relevant extrinsic value is final or not.

In contrast, intrinsic value depends solely on the intrinsic properties (and internal relations) of its bearer. Therefore, no Intrinsic Value Regress involving external relations can get off the ground: the *metaphysical* nature of intrinsic value makes it immune to the *value* regresses

considered. Only intrinsic value can meet the New Aristotelian Challenge and satisfy the New Normative Centrality test. In other words, only intrinsic value can ensure that our practical lives are not empty and vain due to the regresses examined.

In sum, Korsgaard's two distinctions and the classic cases of final extrinsic value call for revising the Aristotelian Challenge *itself* rather than merely revising the Traditional Response to the Challenge. Intrinsic value may well be normatively central and the mark at which we should aim as moral archers.

References

Aristotle. *N.E. The Nicomachean Ethics* (Trans. W. D. Ross rev. by J. O. Urmson; rev. by J. Barnes and A. Kenny). In *Aristotle's Ethics: Writings from the Complete Works (2014),* eds. J. Barnes and A. Kenny, Princeton and Oxford*: Princeton University Press,* pp. 272-372.

Beardsley, Monroe C. 1965. Intrinsic Value. *Philosophy and Phenomenological Research* 26/1, 1-17.

Feldman, Fred 1998. Hyperventilating About Intrinsic Value. *The Journal of Ethics* 2/4, 339-54.

Hirose, Iwao, and Olson, Jonas (eds.) 2015. *The Oxford Handbook of Value Theory*. Oxford University Press.

Hurka, Thomas 2003. Moore in the Middle. *Ethics* 113/3: 599-628.

Kagan, Shelly 1998. Rethinking Intrinsic Value. *Journal of Ethics* 2/4: 277-97.

Korsgaard, Christine M. 1983. Two Distinctions in Goodness. *Philosophical Review* 92/2: 169- 95.

Lemos, Noah M. 2022. Defeat, Pluralism, and Indispensable Goods. *Philosophical Studies* 179/ 10: 3039-53.

Moore, George E. 1903 (1993). *Principia Ethica*. Cambridge University Press.

Olson, J. 2015. Doubts about Intrinsic Value. In *The Oxford Handbook of Value Theory*, eds. I. Hirose and J. Olson, Oxford University Press, pp. 44-59.

Rabinowicz, Wlodek and Rønnow-Rasmussen, Toni 2003. Tropic of Value. *Philosophy and Phenomenological Research* 66/2: 389-403.

Rabinowicz, Wlodek and Rønnow-Rasmussen, Toni 2000. A Distinction in Value: Intrinsic and for Its Own Sake. *Proceedings of the Aristotelian Society* 100/1: 33-51.

Rønnow-Rasmussen, Toni 2021. *The Value Gap.* Oxford University Press.

Rønnow-Rasmussen, Toni 2015. Intrinsic and Extrinsic Value. In *The Oxford Handbook of Value Theory*, eds. I. Hirose and J. Olson, Oxford University Press, pp. 29-43.

Rønnow-Rasmussen, Toni 2011. *Personal Value.* Oxford University Press. Tucker, Miles 2018. From an Axiological Standpoint. *Ratio* 32/2: 131-138.

Tucker, Miles 2016. The Pen, the Dress, and the Coat: A Confusion in Goodness. *Philosophical Studies* 173/7: 1911-22.

Zimmerman, Michael J. 2015. Value and Normativity. In *The Oxford Handbook of Value Theory*, eds. I. Hirose and J. Olson, Oxford University Press, pp. 13-28.

Zimmerman, Michael J. 2001. *The Nature of Intrinsic Value.* Lanham: Rowman and Littlefield.

José Antonio Pérez-Escobar
University of Geneva

The logical structure of analogies between artifacts and biology, and epistemic circularity: implications for scientific practice

In this talk I analyze the logical structure of analogies between the functions of artifacts and biological phenomena to show that they involve epistemic circularity. Moreover, such an epistemic circularity is not always malignant, and accounts for successful instances of analogies of this kind. I also conduct a case study from scientific practice that illustrates both the positive and malignant effects, and reflects on how to optimize the net effect of these analogies.

Analogies between artifacts and biological phenomena have been common in science since the start of the 20th century. However, despite being at the base of important theoretical and practical achievements, like cybernetics and biorobotics, many argue that these analogies are, at least sometimes, pernicious for biological domains (see, for instance, Lewens, 2004; Nicholson, 2013; Nicholson, 2014). An example is conceiving natural selection as a designer in a sense (Nicholson, 2013).

A promising approach to better understand these analogies and their heterogeneous effects in scientific practice is to investigate their logical structure. I am concerned with functional analogies: those that establish parallelisms between the functions of artifacts and biological phenomena. I argue that these analogies involve a type of teleology close to the family of etiological and Selected Effects approaches (e.g., Wright, 1973; Neander, 1991; Garson, 2017). Thus, to understand the logical structure of these analogies, the logical structure of this teleology must be laid out first. This involves reviving and refining/updating previous work on the logical structure of teleological explanations (Wimsatt, 1972) and subsequently applying it to the context of functional analogies between artifacts and biological phenomena.

When a function has explanatory power over the trait that fulfills such a function, some expectations on the characteristics of the trait take place (e.g., it should be shaped in a given manner and not another, or else it would have been unlikely to be selected). In terms of the underlying logical structure, the fulfilled function behaves as both language and metalanguage, and yields some degree of epistemic circularity.

In functional analogies between artifacts and biological phenomena, often the function of an artifact (original domain) is deemed analogous to that of a target biological domain (and less often, the other way around). It is the unknown characteristics of traits (e.g., complex biological traits like brain systems) that cannot be established as analogous and therefore must be investigated. However, as in the original domain, the analogous function establishes specific expectations on the studied traits of the target domain. This entails that, sometimes tacitly, trait characteristics across the two domains are deemed similar even when this is in principle outside the scope of the positive domain of the analogy (what the two domains are known to share; in this case, a similar function). Although this can manifest, for instance, as conceiving biological phenomena as machines too literally, it can also aid scientific discovery productively.

I conduct a case study on the brain's "compass", a neurocognitive system that is believed to play the function of a compass (providing a sense of angular directionality) for spatial orientation. The brain's compass is one of the several neurocognitive systems framed by functional analogies with navigation artifacts (others being cognitive map systems and cognitive speedometers). I survey the effects of the compass analogy on the study and understanding of the brain's compass as a trait (its neurophysiological properties) since the 90s (when the analogy was considered extremely useful) until recent developments (when several limitations of the research program due to the analogy were manifested and prompted revisions). The analysis is conducted in the terms of the underlying logical structure of the compass analogy. Overall, I argue that understanding the underlying logical structure of functional analogies between artifacts and biological phenomena, and the epistemic circularity that they involve, may optimize the net effect of these analogies.

References

Garson, J. (2017). A generalized selected effects theory of function. *Philosophy of Science*, *84*(3), 523-543.

Lewens, T. (2004). *Organisms and artifacts: Design in nature and elsewhere*. Cambridge, MA: MIT Press.

Neander, K. (1991). Functions as selected effects: The conceptual analyst's defense. *Philosophy of science*, *58*(2), 168-184.

Nicholson, D. J. (2013). Organisms≠ machines. *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences*, *44*(4), 669-678.

Nicholson, D. J. (2014). The machine conception of the organism in development and evolution: A critical analysis. *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences, 48*, 162-174.

Wimsatt, W. C. (1972). Teleology and the Logical Structure of Function Statements. *Studies in History and Philosophy of Science 3*, 1-80.

Wright, L. (1973). Functions. *Philosophical Review, 82*(2), 139–168.

Niccolò Aimone Pisano
University of Stirling

Against an interventionist solution to the exclusion problem

The causal exclusion problem has been one of the strongest and most debated metaphysical arguments against nonreductive physicalism in the philosophy of mind over the course of the past 25 years. The core idea of the problem is that, if one is committed to a rather minimal form of nonreductive physicalism, and if some plausible (although by no means uncontroversial) principles are granted, then one is forced to admit that higher-level properties, and mental properties in particular, do not possess causal powers.

Recently, there have been several notable attempts at offering solutions to this problem based on the adoption of Woodward's (2003) interventionist account of causation (e.g. List and Menzies (2010); Zhong (2014); Woodward (2015, 2022)), instead of directly rejecting one or more of the principles giving rise to the problem. However, it has also been argued (for example by Baumgartner (2018)) that the exclusion problem cannot be dispelled as a result of the adoption of interventionism: this would only be the case if the supervenience bases of the relevant higher-level variables were independently fixable, which they are not for metaphysical reasons.

Since the interventionist arguments against the exclusion problem differ significantly from one another, the focus of my talk will be on just one of those interventionist solutions, namely the one proposed by Woodward (2015, 2022). This particular proposal revolves around an amended version of Woodward's original account of interventionism. The main difference between this "interventionism*" and the original version, consists in the fact that the latter is only concerned with purely causal graphs (i.e. with graphs only involving causal relations among variables), while the former explicitly takes into account causal graphs also involving metaphysical determination relations different from causation (in particular, supervenience relations). Specifically, interventionism* highlights the need *not* to hold the supervenience bases fixed while performing interventions on the relevant higher-level supervening variables, thus avoiding the aforementioned "independent fixability objection".

The main claim that I will argue for in my talk is that the reasons why one may be inclined to adopt interventionism* cannot be accepted in the context of the debate over the exclusion problem. To be clear, this is not to say that one may not legitimately adopt interventionism* for pragmatic purposes (hence in conformity with the general interventionist spirit) in ordinary causal reasoning. Rather, my claim is

that, when it comes to the assessment of the exclusion problem, the considerations that play a role in deciding whether some variable is to be held fixed or not tilt the scale in favour of the need to hold the variables representing supervenience bases fixed, when an intervention is performed on the variables representing the relevant supervening (mental) properties. This is because failing to do so would make it impossible to answer the causal questions that motivate the discussion around the exclusion problem itself, specifically: "Do higher-level (mental) entities possess causal powers distinct from those of their supervenience bases?". But, as Woodward (2022, p.22) maintains, the causal questions one is interested in answering are part of the contextual factors guiding the decision to consider some variable as a confounder or not, and hence the choice of whether to hold it fixed or not. Therefore, failing to hold the supervenience bases fixed, as recommended by the amended version of interventionism, prevents one from ascertaining whether the causal powers that one may attribute to mental properties are also thereby possessed by the relevant supervenience bases, or whether they are indeed distinct.

My talk will be divided into two parts. In the first part, I will begin by offering a general presentation of the exclusion problem as a set of jointly inconsistent claims (*Nonreductionism*; *Physicalism*; *Closure*; *Exclusion*; *Causal efficacy*), as it has become more or less customary over time. After that, I will outline the core idea behind the interventionist account of causation: *X* is a direct cause of *Y* with respect to variable set **V** if there are possible interventions on *X* that will change the value of *Y* when all other variables in **V** are held fixed at some value by interventions (Woodward (2022, p.9). In doing so, I will explain how Woodward's recent amended version may be used to address the exclusion problem. Then, in the second part of my talk I will argue that this revised version of interventionism should not be accepted because it does not take into account the specific theoretical interests associated with the debate over the exclusion problem, as it instead should, based on Woodward's own recommendations.

References

Baumgartner, M. (2018). The inherent empirical underdetermination of mental causation. *Australasian Journal of Philosophy*, 96(2), 335-350.

Menzies, P., & List, C. (2010). The Causal Autonomy of the Special Sciences. In C. Macdonald, and G. Macdonald (eds), *Emergence in Mind*. Oxford: Mind Association Occasional Series.

Woodward, J. (2003). *Making Things Happen: A Theory of Causal Explanation*. New York: Oxford University Press.

Woodward, J. (2015). Interventionism and causal exclusion. *Philosophy and Phenomenological Research*, 91(2), 303-347.

Woodward, J. (2022). Modeling interventions in multi-level causal systems: supervenience, exclusion and underdetermination. *European Journal for Philosophy of Science*, 12(4), 59.

Zhong, L. (2014). Sophisticated exclusion and sophisticated causation. *The Journal of Philosophy*, 111(7), 341-360.

Pedro Santos
University of Algarve

If only

The thesis that *if A then B* and *A only if B* (indicative) conditionals are semantically equivalent is adopted by the average logic textbook. Nevertheless, on closer scrutiny, it is fairly problematic (see McCawley's much cited remarks in McCawley 1981). Pairs such as (1)/(2) and (3)/(4) illustrate the point:

(1) If Magda took one these pills, she fell asleep in 5 minutes.

(2) Magda took one these pills only if she fell asleep in 5 minutes.

(3) If my pulse rises above 100, I do physical exercise.

(4) My pulse rises above 100 only if I do physical exercise.

Clearly, (1) and (4) could plausibly be described as true, whereas their counterparts (respectively, (2) and (3)) couldn't. This undermines the notion that *if A, B* and *A only if B* are no more than syntactic variants of each other and, more specifically, gives credence to the suggestion that the two kinds of conditional cannot be construed as truth conditionally equivalent. Interestingly, this behaviour cuts across different conceptions on the correct way to make compositional sense of the relation of "only" and "if" in "only if" conditionals (see, for discussion on the latter issue, von Fintel 1997).

A noticeable fact about "if" conditionals is that they are ambiguous between the kind of reading displayed in (1)/(3) (antecedent-situations referring to the cause and consequent situations referring to the effect) and the kind of reading displayed in (2)/(4) (where the causal relation is reversed), as shown by (5) and (6):

(5) If the firm went bankrupt, John emigrated.

(6) If Ann stayed home, so did Tim.

In fact, by means of (5) one could be stating either that the bankruptcy made John emigrate or that John's emigrating caused the bankruptcy; whereas by means of (6) one could be stating either that Tim stayed home because Ann did or that Tom's staying causally explains Ann's staying. These generate cases of partial symmetry with their "only if" counterparts, as opposed to the "pure" symmetry displayed by (1)-(4). In fact, (5)/(6) and their "only if" versions (respectively, "The firm went bankrupt only if John emigrated" and "Ann stayed home only if Tim also did"). can be interpreted as equivalent, but only in one of the readings of (5)/(6), namely the one in which the consequent refers to the cause and the antecedent to the effect.

McCawley's intuition that the contrasts have truth-conditional import seems to be borne out by the evidence. Take the following example (based on remarks in Lycan 2001). At a party, A and B are discussing John's predictable behaviour:

A: John leaves only if I leave.

B: No, that's false - if the booze runs out, John leaves.

B seems to be correctly implying that the *truth-conditions* of A's statement are breached because there is a possible cause for John to leave other than A's leaving, which arguably means that the causal reading of A´s conditional has indeed truth-conditional import (a very counter-intuitive follow-up to A's conditional would be to describe it as true but pragmatically questionable, along the following lines: *That's true, but misleading - if the booze runs out, John leaves*)

On the other hand, if "if"/"only if" contrasts impinge on the truth-conditions of the conditionals involved, then accounting for the same kind of contrasts in "if" conditionals like (5) and (6) gives credence to the same kind of claim about the latter.

It is worth noting, in this connection, that the extensive formal semantics literature attempting to make compositional sense of the role played by "only" in "only if" conditionals has built on the "exclusionary" meaning of "only" in other constructions (e.g. bare plurals) to account for the *necessary condition* reading of the if-clause in those conditionals (which reverses the *sufficient condition* reading assigned to if-clauses in simple "if" conditionals). This kind of project, if successful, would provide a compositional account of the way "only if" conditionals express a connection between a sufficient and a necessary condition that is symmetric to the one expressed by their "if" counterparts. But it could hardly be expected to provide a compositional account of the kind of the causal nexus between antecedent and consequent that isthe basis of the noted semantic discrepancies between both types of conditionals. In other words, even if it is a component of the truth conditional meaning of "only if" conditionals, such a nexus is unlikely to be a function of the semantics of "only" and "if". Any account successfully dealing with that nexus would therefore have to be non-compositional. *A fortiori*, the same will be true of any account of the way indicative "if" conditionals can be made to express contrasting causal nexus.

This, in turn, lands us in contextualist territory. Scepticism about contextualist approaches to indicative conditionals is well-known. To give two staple examples, Bennett (in Bennett 2003) claims that contextualism is unable to account for communication involving indicatives. Williamson 2020 extensively criticizes any semantic approaches to indicatives that build on the truth-conditional import of such contextual idiosyncrasies as the ones illustrated above.

Bennett and (more directly) Williamson offer their criticisms in connection with the infamous Gibbardian stand-offs (GOS)-situations where both *if A then C* and *If A then not-C* seem well warranted as asserted by different speakers from different epistemic backgrounds. Williamson claims that speakers can endorse indicative conditionals, assign truth-values to them and perform inferences involving them

even if they are not privy to the kind of contextual information that vindicates the assertion of a particular indicative conditional in a GOS- thus a third, neutral, party would be justified, for example, in deriving the falsity of the common antecedent to both conditionals under the assumption that they are both true. In his 2003 book, in turn, Bennett offers a global *reductio* argument, to the effect that speakers typically don't have the tools to properly decode the alleged context-sensitive truth-conditional component of indicative conditionals-thus, if contextualism were true, basic interchanges involving them would be impossible.

In this talk I claim that both kinds of scepticism are unwarranted. Local scepticism overlooks the fact that linguistic behaviour involving indicatives as displayed in a GOS does not entail actual mastery of their truth-conditions. Global scepticism overlooks a general fact about human linguistic interactions, namely that a considerable degree of indeterminacy and vagueness is consistent with effective communication: even when not privy to the truth-conditional content of each other's statements, speakers monitor each other, make amendments and mostly succeed grasping the content their interlocutor intended to share. My claim is that interchanges involving GOS-like indicative conditionals bear this out: a speaker's assertion of an indicative in a GSO effectively conveys information about her epistemic idiosyncrasies, by making them public. Pooling each other's partial information goes towards making communication effective in such cases.

References

Bennett, J. (2003) *A Philosophical Guide to Conditionals*. Oxford: Clarendon Press. Lycan, (2001) *Real Conditionals.* Oxford: Oxford University Press.

von Fintel, K. (1997) "Bare Plurals, Bare Conditionals and *Only*". *Journal of Semantics* 14: 1-56.

McCawley (1981) *Everything that Linguists have Always Wanted to Know about Logic* *but Were Afraid to Ask*. Chicago: Chicago University Press.

Williamson, T. (2020) *Suppose and Tell*. Oxford: Oxford University Press.

Elisabetta Sassarini

University of Oxford

Purely Factual Definitional Disputes

Some disputes arising in various domains of inquiry, including taxonomy, astronomy, metaphysics, as well as ordinary everyday disputes, have been labeled as merely verbal. These include disputes as to whether whales are fish, Sedna is a planet, free will is compatible with determinism, the Pope is a bachelor, and so forth. The idea is, roughly, that if the disputants appear to be equally well-informed about some relevant details of the cases at issue but still find themselves stuck in apparently intractable disagreement, they must somehow be talking past each other in a way that involves some implicit linguistic misunderstanding. For example, if the disputants agree that whales are warm blooded aquatic vertebrates with lungs, or that the Pope is an unmarried man ineligible for marriage, but still dispute whether whales are fish, or whether the Pope is a bachelor, they must be using "fish" or "bachelor" differently.

Most accounts of verbal disputes fall into either of two main categories. On the one hand, on what I will call "semantic consistency" accounts (e.g. Hirsch 2005, 2009, Sider 2006, 2009, Jenkins 2014), divergence in the disputants' uses of the relevant terms entails semantic divergence, so that the speakers' apparently conflicting utterances and beliefs are in fact consistent, and the speakers fail to disagree about the subject matter of their disputes. On the other hand, on what I will call "metalinguistic" accounts (e.g. Chalmers 2011, Plunkett 2015, Plunkett & Sundell 2021), linguistic divergence and agreement on the relevant facts are instead taken to warrant the ascription of conflicting (normative or descriptive) metalinguistic beliefs to the speakers.

In the first part of this paper, I will argue that there can be disputes where the parties agree on the relevant facts and differ in their uses of the relevant terms, but which satisfy neither of the accounts of verbalness above: linguistic divergence and agreement on the relevant facts don't necessarily entail semantic consistency, nor do they warrant the ascription of inconsistent metalinguistic beliefs to the speakers. Disputes involving agreement on some relevant facts and linguistic divergence might instead arise from disagreement about "definitional" but *purely factual* matters, typically concerning real definitions or (higher-order) identities. Two speakers who disagree about such definitional matters may not have any metalinguistic disagreement, and could even lack metalinguistic beliefs about the relevant terms altogether. For example, the parties might dispute whether whales are fish, despite agreeing that whales are warm-blooded aquatic vertebrates with lungs, in virtue of disagreeing about the *non-metalinguistic* definitional question of whether to be a fish is to be an aquatic vertebrate (as opposed to a cold-blooded

aquatic vertebrate with gills). I will argue that this may not involve any metalinguistic disagreement.

I will thus conclude that many, though not all, disputes that are typically considered as verbal can be instead considered as purely factual definitional disputes (PFDDs) arising from *genuine* disagreement about definitional but *non-metalinguistic* facts. Purely factual definitional disputes reduce to verbal disputes only under rather controversial assumptions: they meet the semantic consistency requirement only within a semantic framework leaning towards holism; they meet the metalinguistic disagreement requirement only given some form of deflationism whereby definitional claims are interpreted as metalinguistic.

Although PFDDs are not, strictly speaking, verbal, they may be taken to involve agreement on all the relevant facts, and thus to be in some sense not fully substantive either. Agreement on the facts is, indeed, often considered as the "tell-tale sign" of verbalness (Sidelle 2007, Chalmers 2011, Jenkins 2014). In the second part of the paper, I will argue that this objection holds only given a fairly coarse grained individuation of facts. Under this assumption, however, many perfectly legitimate disputes can be taken to involve agreement on the facts under a certain description. Therefore, agreement on the facts should not be considered as the "tell-tale" sign of verbalness nor does it entail that a dispute is defective or non-substantive in some other sense.

Interpreting the disputes in question as PFDDs, particularly if factual definitional disagreements are taken to concern "heavyweight" matters such as which properties are more joint-carving, explains why such disputes can sometimes persist and appear unsolvable. This account seems preferable to metalinguistic accounts, which portray the disputes in question as implicitly concerning language — a characterization which the disputants themselves would often reject.

The paper also aims to make explicit some core assumptions about semantics, propositional attitude ascription, and the individuation of facts, which implicitly underlie existing accounts of verbal disputes. These crucial issues have, quite surprisingly, received very little attention in the existing literature on verbal disputes.

References

Chalmers, David (2011) "Verbal Disputes", *Philosophical Review* 120 (4):515-566.

Hirsch, Eli (2005) "Physical-Object Ontology, Verbal Disputes, and Common Sense", *Philosophy and Phenomenological Research* 70 (1):67–97.

Hirsch, Eli (2009) "Ontology and Alternative Languages", in David Chalmers, David Manley & Ryan Wasserman (eds.), *Metametaphysics: New Essays on the Foundations of Ontology*, Oxford University Press.

Jenkins, C. S. I. (2014) "Merely Verbal Disputes", *Erkenntnis* 79 (S1):11-30.

Plunkett, David (2015) "Which Concepts Should We Use?: Metalinguistic Negotiations and The Methodology of Philosophy", *Inquiry: An Interdisciplinary Journal of Philosophy* 58 (7-8):828- 874.

Plunkett, David & Sundell, Tim (2021) "Metalinguistic Negotiation and Speaker Error", *Inquiry: An Interdisciplinary Journal of Philosophy* 64 (1-2):142-167.

Sidelle, Alan (2007) "The Method of Verbal Dispute", *Philosophical Topics* 35 (1-2):83-113. Sider, Theodore (2006) "Quantifiers and Temporal Ontology", *Mind* 115 (457):75-97.

Sider, Theodore (2009) "Ontological Realism", in David Chalmers, David Manley & Ryan Wasserman (eds.), *Metametaphysics: New Essays on the Foundations of Ontology*. Oxford University Press.

Wolfgang Sattler

University of Lisbon

Ontological Priority without Separation in Aristotle

Aristotle characterises sensible substances in several passages as being separate or separable (χωριστὸν), i.e. as ontologically independent in some sense. What being separate exactly amounts to here is controversial (cf. e.g. Corkum (2008), Katz (2017)). A central dilemma underlying this controversy concerns the substantial forms of sensible substances. For there seem to be textual and philosophical reasons to make three distinct, but jointly incompatible assumptions: (1) forms are primary substances; (2) whatever is a primary substance and thus ontologically prior is separate (from everything else); and (3) forms are not separate. Hence, forms, being primary substances, should be separate, but it seems they are not. Different ways to resolve this dilemma have been suggested in the literature. Some argue that forms are separate after all (cf. e.g. Spellman (1995), Peramatzis (2011)). Others hold that forms are not separate (i.e. they are not separate from *the respective compound substances*) and argue that forms are consequently not even substances, but only *substances-of* sensible substances (cf. e.g. Wedin (2000), Angioni (2012)). And some hold that there are different senses of separation, and forms are separate in one sense, but not in others (cf. e.g. Gill (1984), Dufour (1999), Katz (2017)). However, all these suggestions fail to accommodate crucial textual evidence and thus are not convincing. I suggest that the problem lies with (2), and more precisely with the following underlying view:

(S) A is ontologically prior to B iff A is separate from B and B is not separate from A.

Ontological priority is usually identified here with priority 'in nature' in *Metaphysics* 5.11. This view (S) has been explicitly defended, for example, by Fine (1984), Corkum (2008), and Peramatzis (2011), and is highly influential. Yet it seems to be wrong in its *unqualified* form, as I argue in my paper. For forms are clearly ontologically prior to matter (cf. e.g. *Metaphysics* 7.3). Yet there is strong evidence that forms are *not separate from matter*, contrary to what is standardly assumed (with exceptions, like Berti (2012)). At the same time, the evidence allegedly supporting (S), for example passages in *Metaphysics* 7.1, *Eudemian Ethics* 1.8, and *Categories* 5, either does not support it; or supports only a restricted form of (S), which excludes the relation between form and matter. Hence, the dilemma about the separation of forms can be resolved by rejecting (2) in its unqualified form.

In this talk I focus on the evidence that form is not separate from matter. For example, in *De Anima* 2.1 the soul, i.e. the form of a living being, is said not to be separate from the body, i.e. the matter of the living being. The only exception here is the so-called active intellect, a soul part of human souls, which alone is separate (cf. *De Anima* 3.5). And in *Physics* 1.2 accidentals are said not to be separate from

substance as their substrate. Hence, a substantial form should likewise not be separate from its substrate, namely matter. Moreover, Aristotle frequently contrasts being separate with being 'in' something, arguably in the sense of being 'in' in which a (substantial or accidental) form is 'in' a substrate (cf. e.g. *Metaphysics* 5.23, 6.1, 13.2, *Physics* 4.3). Hence, a substantial form, being 'in' matter as its substrate, is consequently not separate from matter. Furthermore, Aristotle's so-called prime mover is a non-sensible, eternally existing, unchanging cosmological principle of motion. The fact that it exists eternally and never changes seems to be due in particular to the fact that it is separate from any sensible material substrate, which would contain a capacity for change (cf. e.g. *Metaphysics* 6.1, 12.6). By contrast, substantial forms, such as the souls of living beings, do not exist eternally, but cease to exist when their sensible substances perish. Hence, they seem not to be separate from a material substrate with a capacity for change, i.e. matter. Moreover, according to Aristotle, Plato conceives Platonic Forms (such as the Horse itself) as being separate from the corresponding sensible particulars (e.g. individual horses) as their material substrates. A main motivation for this conception is to ensure the eternal existence of Platonic Forms as objects of knowledge (cf. e.g. *Metaphysics* 13.4). Since substantial forms do not exist eternally, they should, again, in contrast to Platonic Forms, not be separate from a sensible material substrate, namely matter.

References

Angioni, L. (2012). *As Noções Aristotélicas de Substância e Êssencia*. Editora da Unicamp, Campinas

Berti, E. *Estrutura e significado da Metafísica de Aristóteles*. Coleção Filosófica, traduzido do italiano para o português por J. Bortolini. São Paulo: Paulus, 2012.

Dufour, R. (1999). La Séparation chez Aristote. *Les Études philosophiques* No.1, 47-65. Corkum, P. 'Aristotle on ontological dependence'. *Phronesis,* v. 53, 65-92, 2008. Fine, G. 'Separation'. *Oxford Studies in Ancient Philosophy,* v. II, 31-87, 1984.

Gill, M.L. (1989). *Aristotle on Substance. The Paradox of Unity*. Princeton University Press, Princeton

Katz, E. 'Ontological separation in Aristotle's *Metaphysics*'. *Phronesis,* v. 62, 26-68, 2017.

Spellman, L. (1995). *Substance and Separation in Aristotle*. Cambridge University Press, Cambridge

Peramatzis, M. *Priority in Aristotle's Metaphysics*. Oxford: Oxford University Press, 2011

Wedin, M. V. (2000). *Aristotle's Theory of Substance. The Categories and Metaphysics Zeta*. OUP, Oxford

João Pinheiro da Silva

Central European University

Hylomorphism and the Causal Closure of the Physical

The causal closure of the physical (CCP, from now on) has been a topic of hot debate for now various decades in analytic philosophy. There have been various formulations of the principle, some more modest, some more ambitious, but all equally used as premises in a bigger argument for various forms of either materialism, reductionism or eliminativism.

CCP was born and has been mostly deployed in debates surrounding mental causation, more often than not, to deny its possibility. The argument is rather simple: if all of causality is closed under physics, then there is nothing left for the "mental" to do. This simple yet effective argument has been recently expanded beyond the borders of philosophy of mind. Howard Robinson, who is deeply acquainted with CCP's applications in its original context, has recently criticized the contemporary revival of hylomorphism in analytic philosophy for being inconsistent with CCP and, by consequence, modern science.

Robinson has argued that modern hylomorphists faces a dilemma: while they want to be realists about structure (or form) and its causal power, they also want their position to be compatible with CCP. But this is impossible. If one accepts CCP, "it is sufficient for the concept of structure to be applicable that elements be appropriately related in the world, and these relations can be characterized without employing the notion of structure. This could be done by specifying the spatio-temporal location of the elements and their causal influence on each other" (Robinson 2014:11-12). That is, if one accepts CCP's assumption that all of causality is physically closed and, as such, purely bottom-up, there is nothing left for structure to do.

This line of reasoning can be strengthened by taking into account Jaegwon Kim's famous causal overdetermination argument, according to which higher-level phenomena are fully determined by lower-level physical phenomena, and any attempt to posit top-down causation leads to causal overdetermination. By combining the two arguments – which rely heavily on CCP – we will back the hylomorphist into a tough spot: in failing to encompass CCP, not only does he, according Robinson, renounce contemporary science; he also, according to Kim, embraces an inconsistent view of causality.

In order to show the insufficiencies of the CCP argument against hylomorphism, we will first consider CCP as such. We start by settling the exact meaning of CCP and then arguing that its deferral of the meaning of "physical" to physics is doubly problematic: for if it means *present* physics, then CCP is false; and if it means *future* physics, then it is undetermined. That CCP is wrong in terms of present physics will become clearer by carefully looking at various developments in physics that defy

CCP's view of causality as a purely bottom-up endeavor. We will consider superconductivity, the Paoli Principle and quantum physics as examples of how the macroscopic level has a causal role to play in physics, thus defying CCP's implicit microphysicalism. We thus conclude that CCP is not actually *given* by physics, as its proponents often claim, but rather *imposed* on it.

Having considered the problems with CCP, we turn our attention to the philosophical framework of hylomorphism. We first argue that contemporary structural accounts of hylomorphism do indeed fall prey to Robinson's and Kim's arguments by taking a very minimal account of form as mere structure. We will shortly analyze the versions of hylomorphism presented by Koslicki and Jaworski, arguing that both of them fail to provide a meaningful distinction between mere heaps on the one hand and real substances on the other. In taking form to be a mere mereological aspect of a substance or as the configuration of discrete parts, these accounts commit what Robert Koons has called the "statue fallacy", taking the analogy in "Physics II.3 and Metaphysics V.2 as providing a literal example of material and formal causes" (Koons 2014, 153). There, Aristotle says that the form of the statue is its shape. But it is clear in other passages that a statue can only be said to have a substantial form in a loose manner. For it would be strange to say that the configuration of the statue has a capacity for "statueness". The statue has an accidental unity, not a substantial unity. It does not display the dynamicity of a real form, the capacity for self-actualization, the inherent teleology, or the ability to guide and organize the development and behavior of its material constituents. That is because Aristotelian substances are not mere shapes or configurations, but dynamic principles that imbue substances with their unique properties and capacities.

Against structural hylomorphism, we present a version of transformational hylomorphism according to form is better understood as an irreducibly holistic principle of specificity of a thing (that by which it is what it is) that dynamically re-identifies the material parts of an entity towards the same end as a singular causal system, thus preserving the unity of the substance and distinguishing it from the mere accidental unity of a heap.

This account of form, by being more in line with the original Aristotelian account of formal unity, provides us with a cohesive distinction between substances on the one hand and mere heaps or aggregates of matter on the other. This will prove to be particularly advantageous when considering various examples from contemporary biology that defy CCP's microreductionism in showing the importance and causal relevance of biological higher-level phenomena in shaping the its lower-level physical and chemical surrogates.

We end by contrasting two pictures of nature, the first afforded by CCP, the second by hylomorphism. Contrary to what its proponents often argue, CCP's picture of nature, with its monistic view of causation, is not in line with our best science. Contra Kim and Robinson, contemporary science, from physics to biology, provides us a picture of nature where there is a constant and dynamic interplay of higher and lower-level phenomena, with both shaping each other in novel and unpredictable ways. Aristotelian hylomorphism, with its pluralistic approach to causation, provides a stronger metaphysical framework to ground such picture.

References

Aharonov, Y., Cohen, E., & Tollaksen, J. (2018). Completely top–down hierarchical structure in quantum mechanics. *Proceedings of the National Academy of Sciences of the United States of America*, *115*(46), 11730–11735. https://doi.org/10.1073/pnas.1807554115

Anjum, R. L., & Mumford, S. (2011). *Getting Causes from Powers*. Oxford University Press.

Anjum, R. L., & Mumford, S. (2017). Emergence and Demergence. In M. P. Paoletti & F. Orilia (Eds.), *Philosophical and Scientific Perspectives on Downward Causation* (pp. 92–109). New York: Routledge.

Austin, C. J. (2018). A Biologically Informed Hylomorphism. In W. M. R. Simpson, R. C. Koons, & N. J. Teh (Eds.), *Neo-Aristotelian Perspectives on Contemporary Science* (pp. 185–210). New York: Routledge.

Austin, C. J. (2020). Contemporary Hylomorphisms: On the Matter of Form. *Ancient Philosophy Today*, *2*(2), 113–144. https://doi.org/10.3366/anph.2020.0032

Austin, C. J. (2021). Form, Cause, and Explanation in Biology. In L. Jansen & P. Sandstad (Eds.), *Neo-Aristotelian Perspectives on Formal Causation* (pp. 238–268). Routledge.

Austin, C. J., & De La Rosa, L. A. (2018). Dispositional Properties in Evo-Devo. In L. N. De La Rosa & G. B. Müller (Eds.), *Evolutionary Developmental Biology: A Reference Guide* (pp. 1–13). https://doi.org/10.1007/978-3-319-33038-9_64-1

Barbour, I. G. (1990). *Religion in an Age of Science*. Harper San Francisco.

Crane, T. (2001). *Elements of Mind: An Introduction to the Philosophy of Mind*. Oxford University Press.

Crane, T., & Mellor, D. H. (1990). There is No Question of Physicalism. *Mind*, *XCIX*(394), 185–206. https://doi.org/10.1093/mind/xcix.394.185

Dodds, M. J. (1997). Top Down, Bottom Up or Inside Out? Retrieving Aristotelian Causality in Contemporary Science. *Paper Presented at the Annual Thomistic Institute Sponsored by the Jacques Maritain Center, University of Notre Dame*. https://www3.nd.edu/~afreddos/papers/dodds.htm

Drossel, B., & Ellis, G. F. R. (2018). Contextual Wavefunction collapse: an integrated theory of quantum measurement. *New Journal of Physics*, *20*(11), 1–35. https://doi.org/10.1088/1367-2630/aaecec

Dupré, J. (2013). I—John Dupré: Living Causes. *Aristotelian Society Supplementary Volume*, *87*(1), 19–37. https://doi.org/10.1111/j.1467-8349.2013.00218.x

Eliot, T. S. (1943). *Four Quartets*. Houghton Mifflin Harcourt.

Ellis, G. F. R. (2008). On the Nature of Emergent Reality. In P. Clayton & P. Davies (Eds.), *The Re-Emergence of Emergence: The Emergentist Hypothesis from Science to Religion* (pp. 79–107). https://doi.org/10.1093/acprof:oso/9780199544318.003.0004

Ellis, G. F. R. (2018). Top-down causation and quantum physics. *Proceedings of the National Academy of Sciences of the United States of America*, *115*(46), 11661–11663. https://doi.org/10.1073/pnas.1816412115

Feser, E. (2014). *Scholastic Metaphysics: A Contemporary Introduction*. Editions Scholasticae.

Fine, K. (1999). Things and Their Parts. *Midwest Studies in Philosophy*, *23*(1), 61–74. https://doi.org/10.1111/1475-4975.00004

Gabbani, C. (2013). The Causal Closure of What? An Epistemological Critique of the Principle of Causal Closure. *Philosophical Inquiries*, *1*(1), 145–174. https://doi.org/10.4454/philinq.v1i1.14

Hacker, P. M. S. (2011). A Plague On Both Your "Isms." *American Philosophical Quarterly*, *48*(2), 97–111.

Hempel, C. G. (1980). Comments on Goodman's Ways of Worldmaking. *Synthese*, *45*(2), 193–199. https://doi.org/10.1007/bf00413558

Jaworski, W. (2016a). *Structure and the Metaphysics of Mind: How Hylomorphism Solves the Mind-Body Problem*. Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780198749561.001.0001

Jaworski, W. (2016b). Why Materialism Is False, and Why It Has Nothing To Do with the Mind. *Philosophy*, *91*(2), 183–213. https://doi.org/10.1017/s0031819116000036

Johnston, M. (2006). Hylomorphism. *The Journal of Philosophy, 103*(12), 652–698. https://doi.org/10.5840/jphil2006103125

Kim, J. (1993). *Supervenience and Mind: Selected Philosophical Essays*. Cambridge University Press.

Kim, J. (1999). doi.org/10.1023/A:1004563122154. *Philosophical Studies*, *95*(1/2), 3–36. https://doi.org/10.1023/a:1004563122154

Koons, R. C. (2014). Staunch vs. Faint-hearted Hylomorphism: Toward an Aristotelian Account of Composition. *Res Philosophica*, *91*(2), 151–177. https://doi.org/10.11612/resphil.2014.91.2.1

Koslicki, K. (2006). Aristotle's Mereology And The Status Of Form. *The Journal of Philosophy, 103*(12), 715–736. https://doi.org/10.5840/jphil2006103127

Laughlin, R. B. (1999). Nobel Lecture: Fractional quantization. *Reviews of Modern Physics*, *71*(4), 863–874. https://doi.org/10.1103/revmodphys.71.863

Lewis, D. (1986). *On the Plurality of Worlds*. Oxford: Blackwell Publishing.

Marmodoro, A. (2013). Aristotle's hylomorphism without reconditioning. *Philosophical Inquiry, 37*(1), 5–22. https://doi.org/10.5840/philinquiry2013371/28

Mayr, E. (2017). Powers and Downward Causation. In M. P. Paoletti & F. Orilia (Eds.), *Philosophical and Scientific Perspectives on Downward Causation* (pp. 76–92). New York: Routledge.

Montero, B. (2003). Varieties of causal closure. In S. Walter & H.-D. Heckmann (Eds.), *Physicalism and Mental Causation.* Imprint Academic.

Morrison, M. (2000). *Unifying Scientific Theories: Physical Concepts and Mathematical Structures*. Cambridge University Press.

Mulder, J. (2021). The Limits of Reductionism: Thought, Life, and Reality. In O. Passon & C. Benzmüller (Eds.), *Wider den Reduktionismus - Ausgewählte Beiträge Zum Kurt Gödel Preis 2019* (pp. 25–40). Springer Nature Switzerland. https://doi.org/10.1007/978-3-662-63187-4_4

Noble, D. (2008). *The Music of Life: Biology Beyond Genes*. Oxford University Press.

Noble, D. (2012). A theory of biological relativity: no privileged level of causation. *Interface Focus*, *2*(1), 55–64. https://doi.org/10.1098/rsfs.2011.0067

O'Connor, T. (2020). Emergent Properties. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy (Winter 2021 Edition)*. https://plato.stanford.edu/entries/properties-emergent/

Oderberg, D. S. (2007). *Real Essentialism*. Routledge Studies in Contemporary Philosophy.

Oderberg, D. S. (2017). The Great Unifier: Form and the Unity of the Organism. In W. M. R. Simpson, R. C. Koons, & N. J. Teh (Eds.), *Neo-Aristotelian Perspectives on Contemporary Science* (pp. 211–234). New York: Routledge.

Paoletti, M. P., & Orilia, F. (2017). Downward Causation: An Opinionated Introduction. In M. P. Paoletti & F. Orilia (Eds.), *Philosophical and Scientific Perspectives on Downward Causation* (pp. 1–21). New York: Routledge.

Papineau, D. (1998). Mind the Gap. *Noûs*, *32*(S12), 373–388. https://doi.org/10.1111/0029-4624.32.s12.16

Papineau, D. (2001). The Rise of Physicalism. In *Cambridge University Press eBooks* (pp. 3–36). https://doi.org/10.1017/cbo9780511570797.002

Robinson, H. (2014). Modern Hylomorphism and the Reality and Causal Power of Structure: A Skeptical Investigation. *Res Philosophica*, *91*(2), 203–214. https://doi.org/10.11612/resphil.2014.91.2.5 pagination follows the version available at https://www.researchgate.net/publication/274449029

Robinson, H. (2021). Aristotelian Dualism, Good; Aristotelian Hylomorphism, Bad. In P. Gregoric & J. L. Fink (Eds.), *Encounters with Aristotelian Philosophy of Mind* (pp. 283–306). Routledge.

Simpson, W. M. R. (2021). From Quantum Physics to Classical Metaphysics. In W. M. R. Simpson, R. C. Koons, & J. Orr (Eds.), *Neo-Aristotelian Metaphysics and the Theology of Nature* (pp. 21–65). Routledge.

Simpson, W. M. R. (2022). The Myth of the Physical Given. *Revue Philosophique De Louvain*, *119*(1), 151–164. pagination follows the version available at https://www.academia.edu/74195586

Simpson, W. M. R. (2023). *Cambridge University Press*. Cambridge University Press. https://doi.org/10.1017/9781009026475

Sterelny, K., & Griffiths, P. E. (1999). *Sex and Death: An Introduction to Philosophy of Biology*. University of Chicago Press.

Stoliar, D. (2001). Two Conceptions of the Physical. *Philosophy and Phenomenological Research*, *62*(2), 253–281. https://doi.org/10.2307/2653699

Tabaczek, M. (2019). *Emergence: Towards a New Metaphysics and Philosophy of Science*.

Tabaczek, M. (2023). Classical and Contemporary Versions of Hylomorphism: Philosophical and Scientific Perspectives. *Proceedings of the Eleventh International Thomistic Congress*, 1–44. Draft Provided by the Author

Walsh, D. (2006). Evolutionary Essentialism. *The British Journal for the Philosophy of Science*, *57*(2), 425–448. https://doi.org/10.1093/bjps/axl001

Walsh, D. M., & Wiebe, K. (2020). The being of living beings. In A. S. Meincke & J. Dupré (Eds.), *Biological Identity: Perspectives from Metaphysics and the Philosophy of Biology* (pp. 107–127). Routledge.

Weisberg, M., Needham, P., & Hendry, R. (2019). Philosophy of Chemistry. *The Stanford Encyclopedia of Philosophy (Spring 2019 Edition)*. https://plato.stanford.edu/entries/chemistry/

Wilson, J. (2015). Metaphysical emergence: Weak and Strong. In T. Bigaj & C. Wüthrich (Eds.), *Metaphysics in Contemporary Physics* (pp. 251–306). Brill/Rodopi.

Rui Silva

University of the Azores

Value-ladenness and the integrity of science

Traditional philosophy of science was strongly influenced by the so-called fact value dichotomy, according to which scientists, in order to produce objective knowledge, should avoid value judgments in the proper domain of science or (to use a distinction proposed by Reichenbach) in the "context of justification", as opposed to the "context of discovery", where values may inspire the direction of research or the formulation of hypotheses. In the second half of the 20th century, however, many authors argued that values play an important role in core activities of science like theory choice. We can divide their arguments in three main categories: methodological, linguistic and sociological arguments. The first category involves, for instance, underdetermination arguments, based on the premiss that the gap between evidence and theory cannot be bridged by logico-methodological principles (Longino 1990), or the argument of inductive risk, according to which the acceptance of rejection of hypotheses depends on value judgments concerning the consequences of error (Rudner 1953 and Douglas 2000). Linguistic arguments claim that the scientific language contains, especially in the human sciences, "thin terms", terms that are both descriptive and evaluative (Root 1993). Sociological arguments emphasize the presence of social values in the categories used in the social and behavioural sciences.

The debate on the role of values in science has been enriched in the last decades by a series of distinctions that contribute to a more rigorous discussion. The first important distinction is between epistemic and non-epistemic values. Accuracy, consistency, scope, simplicity and fruitfulness are examples of epistemic values (Kuhn 1977). If the epistemic/non-epistemic distinction holds and if we only allow the presence of epistemic values in the core activities of science, we avoid the danger of political or ideological contamination. However, the distinction can be challenged. Longino (1996), for instance, argued that values like novelty and ontological heterogeneity are in the borderline of the epistemic and non-epistemic distinction. They have a feminist motivation: novelty opposes traditional biases and ontological heterogeneity is a counterweight against the neglect of differences. Still, we can legitimately accept a theoretical distinction even if there are borderline cases in practice.

Another important distinction is the distinction between the direct and indirect role of values in science (Douglas 2009). According to it, only epistemic standards like empirical adequacy and logical consistency can play a direct role in theory choice; other values may only influence the lowering or raising of evidential standards and the characterization of data. In a similar vein, Steel (2017) argued for

a prioritizing of epistemic values over non-epistemic ones. However, it might be objected that the epistemic priority thesis is false, because the progress of scientific research involves choices at several steps, and the presence of non-epistemic values at the start of investigations can even irradiate its influence over later stages of the research (Brown 2020).

I propose a distinction between a disclosing and constructive role of values in science. The disclosing role can be analysed in two different ways. On the one hand, value judgments can have a debiasing role in science by discovering new evidence. Feminist philosophy of science offers several examples of how feminist values expose traditional biases in scientific research. There is, for instance, feminist research that challenges the traditional women-gatherer/man-hunter paradigm, according to which women played a subordinate role in comparison to men conceived of as natural "breadwinners". We must notice that research guided by this sort of values does not compromise the integrity of science; on the contrary, it promotes it by denouncing its contamination by erroneous values.

Value judgments can also have a disclosing role without gathering new evidence, simply by highlighting relevant aspects of reality. An example is the important historiographic debate between the so-called "ideological-intentionalist" explanation of the Holocaust, according to which the participation of ordinary Germans in the Holocaust can be explained by the "eliminationist antisemitism" prevalent in German society (Goldhagen 1997) and the dominant, "structural-functionalist" explanation of the Holocaust, which emphasizes adverse economic and historical circumstances, as well as psychological-social mechanisms, like obedience to authority or peer pressure. Each explanation has its own difficulties in integrating all the available evidence, but let us imagine that a historian, starting from a positive evaluation of the German culture, rejects the ideological-intentionalist explanation (which emphasizes the role of cultural factors). Let us also assume that this historian adequately justifies his or her positive valuation, invoking, for instance, the existence of a rich literary, artistic and philosophical tradition, as well as a neo-humanist movement that shaped the German educational system since the 18th century. If one accepts that such a value judgment, besides being insightful, also provides a reason to prefer the structural/functionalist explanation, then in addition to its disclosing role, it also plays a constructive role by directly justifying a theory. Even if it is possible to present some cases where the constructive role of values in science seems acceptable, this role is problematic because it may contaminate scientific research with ideological assumptions. Feminist and pragmatist accounts of science or the movement of critical theory claim that the "right values" make science better. If we take values in its disclosing role, this claim is correct. If we take them in its constructive role, the claim is incorrect. "Right values" might be controversial and they can also lead to false conclusions.

To conclude, although it is necessary to recognize the inevitable presence of values in science, we should combine this recognition with a Weberian reluctance towards the instrumentalization of science by political or ideological agendas. The value-ladenness of science can, in fact, endanger the integrity of science, a problem that can be mitigated with some remedies: transparency (the willingness to declare the values behind one's work), intersubjective discussion (to detect the unconscious

influence of values) and empirical adequacy, in the sense that the projection of erroneous values in science typically fails to integrate the available evidence into a coherent whole.

References

Brown, Matthew. 2020. *Science and Moral Imagination*. University of Pittsburgh Press. Douglas, Heather. 2000. "Inductive risk and values in science". *Philosophy of Science*, 67, 559-579.

Douglas, Heather. 2009. *Science, Policy, and the Value Free Ideal*. University of Pittsburgh Press.

Goldhagen, Daniel. 1997. Hitler's Willing Executioners: Ordinary Germans and The Holocaust. New York, Vintage Books.

Kuhn, Thomas. 1977. "Objectivity, values, and theory choice". In: *The Essential Tension*, Chicago, IL, Chicago University Press, 320-339.

Longino, Helen. 1990. *Science as Social Knowledge*, Princeton, NJ: Princeton University Press.

Longino, Helen. 1996. "Cognitive and Non-Cognitive Values in Science: Rethinking the Dichotomy". In L. Nelson & J. Nelson (eds.), *Feminism, Science, and the Philosophy of Science*. Dordrecht: Kluwer, 39-58.

Root, Michael. 1993. *Philosophy of Social Science*. Oxford: Blackwell. Rudner, R. 1953. "The scientist qua scientist makes value judgments". *Philosophy of Science* 20, 1-6,

Steel, Daniel. 2017. "Qualified epistemic priority". In K. Elliot and D. Steel (eds.), Current Controversies in Values and Science. London: Routledge.

Pinelopi Stylianopoulou

University of Glasgow

Logical consequence in a sub-classical pluralist framework

Critics of Logical Pluralism have argued that the position is untenable when classical logic belongs to the plurality of logical accounts that the pluralist endorses. Since the alternative logics usually considered are sub-classical, Logical Pluralism often collapses into Logical Monism with classical logic being the one true logic. This is a rough sketch of the so-called 'Collapse Argument' which aims to show that, since such cases exist, pluralism about logic is false. In this paper, I consider what happens when the logics endorsed by the pluralist are exclusively sub-classical and comparable in logical strength. Specifically, I examine a pluralist account containing the logics of relevance and intuitionism and observe how trumping among these occurs in the presence of certain argument schemata. The paper offers a novel case study assessing this narrower pluralist position compared to the wider-ranging and less-viable account that is typically defended by Beall and Restall (2006). If successful, the case study presented, will highlight a potentially viable shape that pluralism about logic can take.

I begin with an exposition of the Collapse Argument (henceforth, CA). The CA as I refine it, has four premises: one saying that every logic is normative for reasoning, one relating normativity to conformity with argument schemata, a third relating the obligation to conform with trumping among logics, and a fourth relating trumping among logics to pluralism. I proceed by challenging the fourth premise of the CA. The premise states that, if logical pluralism is true, then there are logics L_1 and L_2 and argument schemata that are L_1-valid but not L_2-valid, and such that L_1 does not trump L_2 with respect to arguments that are instances of that schema.

In response, I adopt a pluralist framework comprised of logical systems of incommensurable strength, thereby allowing for trumping to mutually occur with respect to distinct argument schemata. Since classical logic is such that it trumps all its standard rivals, the version of logical pluralism I consider endorses the sub-classical accounts of relevance (L_r) and intuitionism (L_i). I begin by considering two arguments: one in the form of double negation elimination (DNE), and one in the form of disjunctive syllogism (DS) which are L_i-valid but L_r-invalid and L_r-valid but L_r invalid respectively. Assuming that the pluralist takes the premises in each of the arguments to be true, they ought to see to it that they accept the inference in DNE via L_r while rejecting L_i, and further, that they accept the inference in DS via L_i while rejecting L_r. Both cases of argument evaluation stand as paradigmatic instances of trumping among each of the admitted logical accounts and indicate that trumping can (and indeed does) mutually occur in sub-classical frameworks of logical

pluralism. If so, it can be argued, it is not the case that trumping is incompatible with pluralism about logic.

However, another kind of collapse into monism might arise via trumping, whereby the right logic is a disjunction of the plurality of logics. I call this Disjunctive Pluralism. I observe that, when trumping occurs, a single logic is admitted, and when trumping does not occur, a plurality of logics is redundant since a single logic suffices to do the job. A pragmatic consequence of endorsing the pluralist view then, might be that, in fact, there is no genuine pluralism in any possible case of argument evaluation. Thus, suggesting that pluralism about logic is false. In response, I consider an expanded argument in which I amalgamate the steps of the inferences reached in DNE and DS via L_r and L_i respectively and employ conjunction introduction as a final step. This argument gives rise to a further kind of pluralist consequence relation, one that follows from the transitive closure of the logics that act as the disjuncts in Disjunctive Pluralism. I call the resulting pluralism, Closure Pluralism, whereby the right logical consequence relation is one that stands as the generalised transitive closure that extends the disjunctive relation. But transitivity ensures a single logical consequence relation which is yet another kind of monism about logic. This implication poses a dilemma for the logical pluralist: either deny that the inference is a valid one, or admit a single logical consequence relation, namely the one put forward by the Closure. I argue that what we require for the purposes of achieving a viable shape of pluralism, is a way to account for each of the conjuncts we inferred via L_r and L_i respectively, while rejecting their conjunction. To do this, I revisit Disjunctive Pluralism and argue that it can stand as a viable shape of logical pluralism by drawing an analogy between compartmentalisation of inferences and Lewis's idea of compartmentalisation of beliefs (1996). We can think of the pluralist as an agent who reasons in more than one way, all of which are captured in different compartments. The sub classical logical pluralist in our case, has a compartment that reasons relevantly and another that reasons intuitionistically. Each of the compartments gives rise to (potentially) different inferences, and while the pluralist takes both these inferences to be correct, they don't take their conjunction to be correct also. This is because, in the absence of a compartment that is an instance of the closure of the plurality of the compartments, the pluralist cannot deduce the conjunction. The structure of the paper is as follows: I begin in sect. 1. by presenting a refined version of the CA. In sect. 2., I challenge the fourth premise of the argument through the adoption of a sub classical pluralist account and show how trumping mutually occurs. Section 3. considers two shapes that pluralism takes following trumping, and whether these constitute yet another type of monism about logic. I conclude in Section 4. that only a pluralist framework of the disjunctive sort can maintain the plurality of logical consequence relations while avoiding collapse.

References

Beall, J. C., and Restall, G. 2006. *Logical Pluralism*. Oxford: Clarendon Press.

Lewis, D. 1996. Elusive Knowledge. *Australasian Journal of Philosophy*. 74: 4, 549-567.

Foivos Syrigos

University of Athens

Anschauung in Kant's conception of geometry

In the present paper, I will argue that Kant's conception of mathematics isn't just a part of the history of philosophy but it's role is timely and largely overlooked. In particular, I will claim that the appeal to Kantian intuition is, in some cases, still necessary. In the first section of my paper I will analyze the classic reading of the role of Kantian intuition in geometry, based on the distinction that Brittan (2006) proposes. Moreover, I will briefly mention past and recent objections that have been presented against Kant's arguments. In the second section I will propose a reading of the A 716–717 / B 744–745 of the first critique, where I believe that the use of Kantian intuition remains timely and philosophically important. This role of intuition, which I call "revelating", remains indispensable despite the advances that have been made in mathematics and logic since Kant's times. Furthermore, I will argue that almost the whole of Kant's interpreters that have translated the first critique in the English language have disregarded this function of intuition, which also holds for the official translations in French and also in Greek. The goal of this paper is to show why these interpretations doesn't fit with the context in which Kant develops his philosophy of mathematics, as much as to claim that the revelating role of intuition isn't offended by the objections that have been set in the classic reading of Kant's geometry.

The classic reading of the role of the Kantian Anschauung can be distinguished in two groups (Brittan 2006), although both groups share a common characteristic: Mathematical propositions, according to Kant, can't be cognized through concepts alone. The first group, with its main representatives being Frege and Hilbert, thought that we need intuition in order to *verify* the truthfulness of the unproven premises, these being (famously) Euclid's postulates. In contrast, the second group emphasized the role that intuition plays *during* the deductive process. In this view, intuition's role is to guarantee the truth or the correctness of mathematical arguments as is, in particular, the assurance of the existence of points in geometrical propositions.

In today's context, with the logical tools of the existential and universal quantifiers, existence seems to be a mere matter of logic. Plus, Kant's treatment of the figure as a space-time object, although plausible in his times, seems almost incredible to modern mathematicians. Therefore, what role can we assign to Kantian intuition in order to be considered relevant in the modern context?

My answer to this question comes from the part of the *CPR* called "The discipline of pure reason in its dogmatic use" where Kant compares the mathematical with the philosophical method. Even though both the mathematician and the philosopher are

experienced in working with abstract objects, Kant strikingly notes that only the first uses the whole of his intellectual capacities. If you give a philosopher the concept of a triangle, Kant says, no matter how long he meditates on this and relevant concepts, he *won't* produce anything new. Something seems to be missing from his intellectual toolbox. This missing piece is shown by analogy of the geometer's treatment of the problem.

When the geometer takes up this question, he constructs a triangle, reflects on the figure and the premises/ propositions proven before, then extends some sides of the triangle and, after a chain of constructions and inferences, proves the desired result. What is this thing that leads the geometer to the solution of the problem? *Intuition*. Kant uses the word "einleuchtenden" to refer to the solution guided by intuition which, literally, can be translated as what sheds light to something. Therefore, while the geometer is working on the problem, intuition *guides* him through it and, when the result has been reached, we have something like a *revelation*, the discovery of new knowledge. Guyer and Wood (G& W, 1998) have translated "einleuchtenden" as "fully illuminating", which agrees with the present interpretation. Unfortunately, that's the only English translation which does so.

Both Haywood (1838, the first English translation) and Meiklejohn (1855) have selected the word "clear" to refer to "einleuchtenden". Müller (1896), in contrast, preferred the word "convincing" which, although not as appropriate as the one G & W use, I believe is better than the others that have been proposed through the years. Since Müller and until G & W, the whole of translators (Smith 1929, Pluhar 1996, Weigelt 2007 (based on Müller's (!))) preferred the word "evident", whereas Dimitrakopoulos (2006) in the present Greek translation uses the words "σαφής ή προφανής" which is similar to "evident" and Tremesaygues and Pacaud (1975), in the contemporary French translation, use the word "claire" which is akin to the one Haywood and Meiklejohn selected.

In short, a major goal of Kant throughout the CPR is to show the existence of a priori synthetic propositions. If Kant succeeds in proving that in mathematical propositions, he would have refuted one of the main arguments of rationalists, who support that they are knowledgeable solely through the understanding. Kant in A 716–717/744–745 describes , in an extraordinary manner, both the mental state of the practicing geometer but also the *limits* of the understanding. In simple terms, without the Anschauung *leading* us to the solution, we can't have it at all. *That's also true today*. Regarding the other translations proposed, they are relevant with the Hintikka-Parsons debate on the role of intuition but, unfortunately, out of the context presented here.

Guido Tana

NOVA University Lisbon / IUSS Pavia

Recognizing Others in Arguments: Epistemic Injustice and the Adversarial Stance

Does critical engagement in argumentation entail epistemic injustice against oppressed subjects? This presentation addresses this question, raised in recent contributions and debates within social epistemology and argumentation theory. The adversarial requirement of philosophical argumentation consists of subjecting one's dialectical opponent to a high degree of scrutiny. This is predicated on the supposition that the aim of argumentation is its aim towards truth. However, recent research has raised the specter that this adversarial element of argumentation might be yielded as a weapon against those subjects who suffer from an underprivileged or underrepresented social position. The skeptical element inherent to the adversarial disposition towards the arguer entails that when such underprivileged agents are subjected to skeptical scrutiny, they are victims of epistemic, specifically hermeneutic, injustice.

This presentation aims to argue against this verdict and defend the idea that it is precisely by sparing them such an application of skeptical scrutiny that they become victims of hermeneutical injustice. This stance will be defended as follows: after presenting the case for argumentative skepticism entailing hermeneutical injustice, it is shown how the cases presented in its defense establish testimonial injustice instead. Consequently, it is argued that lack of skeptical argumentation deprives the underprivileged subject of the capacity to present her reasons and defend them as being epistemically cogent. The lack of adversarial engagement amounts to a form of silencing, enacting hermeneutical injustice. This assessment is then refined by engaging with other positions in the debate that have recognized the importance of adversarial argumentation to avoid epistemic injustice. An account of skeptical argumentation based on the practice of mutual recognition then is provided

The adversarial understanding of argumentation endorses the following constraints: Formal constraint) arguing for a certain position p at least entails the belief that not-p is incorrect. Pragmatic constraint) those who defend not-p are one's opponents (Grovier 1999).

The formal constraint is assumed to be necessary according to the adversarial view of arguments (Aikin 2011; Casey 2018), and the pragmatic aspect is its simple consequence (Vollbrecht 2022, p.271).

Together they entail that the parties of the dispute must reason under what has been called

> *Default Skeptical Stance*) the evaluation of an argument can only be obtained by subjecting it to 'the strongest or most extreme position' (Moulton 1983, p.153, Hundleby 2010, p.284).

Non-adversarial stances on argumentation view this default skeptical stance as leading to epistemic dysfunction. Phyllis Rooney (2010; 2012) argues that this dysfunction targets epistemic agents belonging to underprivileged groups in society, instantiating hermeneutical injustice. Placing the burden of proof on the underprivileged subjects enables an unfair kind of scrutiny whose endpoint will be their silencing and exclusion from the epistemic community.

Against this reading, it will be shown how Rooney's understanding of such cases is mistaken. First, in the cases she describes of unfair and unwarranted dismissal of arguments, silencing is not due to hermeneutical injustice. Hermeneutical injustice concerns the availability of conceptual and

hermeneutical resources available to members of a community to have an appropriate understanding of their experiences (Fricker 2007, p.148). A harm to the members of a community is done in their capacity as knowers being undercut by this conceptual lacuna that does not allow them to make sense of what they experience as members of a marginalized community (ibid, p.151). However, the problematic issue in Rooney's case is that the reasons and arguments brought forth by the members of the marginalized community are unfairly dismissed and ignored. This means that the issue is not one of a lacuna of conceptual self-understanding but one where the subject endures an unfair deficit of credibility due to systematic asymmetries in place, the hallmark of *testimonial injustice* (ibid, p.17; 1998).

The second main issue with Rooney's understanding concerns the taking away from the marginalized subject the possibility of defending their reasons and concepts. Having one's arguments subjected to skeptical scrutiny is an essential tool for their self-understanding and development of a subject's position in an epistemic community. Exempting marginalized subjects from the very possibility of such an exchange leads to a lack of recognition towards them as candidates for full participation in communal epistemic practices. By doing so, the subject suffers an unfair lacuna concerning their capacity as knowers and arguers. Rejecting a priori the default skeptical stance engenders hermeneutic Injustice, against Rooney's contention.

The above argument is then further developed via an engagement with the current literature. Allan Hazlett's (2020) stance envisions the mere participation in the adversarial dialectic as being a remedy to injustices, but as Lucy Vollbrecht (2022) aptly recognized, simple participation in dialogical practices cannot suffice to avoid the peril of epistemic injustice affecting the exchange. By taking inspiration from the Pyrrhonian method of inquiry, Vollbrecht proposes a cooperative conception of skeptical argumentation where what is forbidden is to rule out the other's arguments on non-reason-based considerations. However, it is argued here that this picture is incomplete. The requirement of critical uptake seems to clash with the fact that even when such requirements are taken in at face value, implicit biases might still create enough epistemic dysfunction to generate cases of testimonial injustices (Saul, 2013, Ballantyne, 2015). Secondly, Vollbrecht's picture ignores that the application of a default skeptical stance in situations of institutional asymmetry is not

merely the purview of the subject in the dominant position. Refusal of engagement and a stance of distrust are conceptual tools available to marginalized communities as well (Baghramian&Panizza 2022; Brennan 2023).

To solve these issues, the skeptical aspect of argumentative exchanges ought to be understood under the lens of mutual epistemic recognition. Recognizing others as autonomous and authoritative sources of knowledge means that no lack of credibility befalls them, avoiding the epistemic dysfunction that leads to forms of testimonial injustices (cf. Congdon 2018, Giladi 2018). In turn, recognizing the skeptical adversary as acting upon their commitments of fair skeptical scrutiny ensures the necessary kind of trust the marginalized subject demands to present their reasons and have them recognized as valid in the argumentative community, avoiding hermeneutical injustice. While critical uptake might fall prey to biases, mutual recognition is fundamentally a normative activity where subjects assume responsibility in their argumentative behavior for what they are committed to before the wider community. This means that the community, in turn, assesses them concerning their ability to fulfill what they are entitled to do as argumentative partners. Secondly, confronted with the possibility of pervasive, systematic biases affecting the evaluation of one's arguments, it allows the adoption of an attitude of skeptical distrust on behalf of the arguer, allowing them to enact epistemic resistance (Medina 2018).

References

Aikin, S. (2011), A Defense of War and Sport Metaphors in Argument, *Philosophy and Rhetoric*, 44:3, 250-272

Baghramian, M & Panizza, S. (2022), Skepticism and the Value of Distrust, *Inquiry*, forthcoming, 1-26

Ballantyne, N. (2015), Debunking Biased Thinkers, *Journal of the American Philosophical Association*, 1:1, 141-162

Brennan, J. (2023), Skepticism, the Virtue of Preemptive Distrust, *Journal of the American Philosophical Association*, forthcoming, 1-18

Casey, J. (2018) Regarding the Adversary Paradigm, in *ISSA 9 Conference Proceedings*, edited by B. Garssen, D. Godden, G. R. Mitchell, J. Wagemans, 155-163, Amsterdam: Sic Stat

Congdon, M. (2018), "Knower" as an Ethical Concept: From Epistemic Agency to Mutual Recognition, *Feminist Philosophy Quarterly*, 4:4, 1-26

Fricker, M. (1998), Rational Authority and Social Power: Towards a Truly Social Epistemology, *Proceedings of the Aristotelian Society*, 98:2, 159-177

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

Giladi, P. (2018), Epistemic Injustice: A Role for Recognition?, *Philosophy and Social Criticism*, 44:2, 141-158

Grovier, T. (1999), *The Philosophy of Argument*, Newport News, VA: Vale Press

Hazlett, A. (2020), Critical Injustice, *American Philosophical Quarterly*, 57:2, 129-144

Hundleby, C. (2010), The Authority of the Fallacies Approach to Argument Evaluation, *Informal Logic*, 30:3, 279-308

Medina, J. (2018), Misrecognition and Epistemic Injustice, *Feminist Philosophy Quarterly*, 4:4, 1-16

Moulton, J. (1983) A Paradigm of Philosophy: the Adversary Method, in *Discovering Reality*, edited by S. Harding and M. B. Hintikka, Boston: D. Riedel

Rooney, P. (2010), Philosophy, Adversarial Argumentation, and Embattled Reason, *Informal Logic*, 30:3, 203-234

Rooney, P. (2012), When Philosophical Argumentation Impedes Social and Political Progress, *Journal of Social Philosophy*, 43:3, 317-333

Saul, J. (2013), Scepticism and Implicit Bias, *Disputatio*, 5:37, 243-263

Vollbrecht, L. (2022), Why We Need Skepticism in Argument: Skeptical Engagement as a Requirement for Epistemic Justice, *Argumentation*, 36:2, 269-285

Jesús Vega-Encabo
Autonomous University of Madrid

Intellectual autonomy, understanding, and intellectual ethics

In the first chapter of *Epistemic Explanations*, E. Sosa (2021) explores the nature of a particular kind of epistemic achievement: understanding-why and first-hand knowledge. It is primarily a kind of understanding that involves some explanatory knowledge of phenomena. But it is not this aspect of advancing understanding that is highlighted in the discussion. Sosa claims, first and foremost, that there are certain issues that demand that epistemic agents do not settle them by deference to others but significantly through the exercise of their own competences and resources. This is so at least in all those matters in which "rational appreciation" is at stake; these are matters that could not be adequately settled by consulting others, however expert they may be considered, and by deference to their judgement and authority. One's own insight and understanding of the issues must be the guide in forming belief, and no amount of deference can properly close the deliberation or settle the question. In other words, it would be epistemically inappropriate to defer in these matters, even if deferring to others could provide us with true, reliable beliefs, knowledge and also some form of understanding. This will be "truncated", Sosa suggests. Understanding could perhaps be gained under conditions of epistemic dependence, but there would still be something particularly desirable about attaining knowledge (and understanding, in particular) at first hand. But what makes deference in these cases epistemically inappropriate? Is it a matter of some epistemic achievements only being attainable in the first person?

It is easy to think that the acquisition of this first-hand knowledge involves the manifestation of a virtue that we have traditionally identified as intellectual autonomy. Sosa, in several passages of his work, stresses the special value we attach to intellectual autonomy, without which we would be unable to place in its proper place what other sources, including testimony, might give us. But what is the place of intellectual autonomy in the explanation of the normativity of epistemic achievements? Is it only under conditions of autonomy that certain achievements can be constituted as attributable to the agent? In this contribution I explore different lines of interpreting how it might be explained within the framework of Sosa's epistemology of virtues or competences what makes desirable, in epistemic terms, the aspiration to constitute achievements under conditions of full agency, achievements that might perhaps involve the cultivation of dispositions and attitudes characteristic of the intellectually autonomous self.

I will defend that Sosa's distinction between gnoseology and intellectual ethics allows to see how intellectual autonomy can figure in epistemology insofar as it explains how epistemic agents could aspire to full agency. Here my set of claims: (i) Intellectual ethics is concerned with how we cultivate those dispositions, attitudes, and traits that can contribute to place us in a position to know (or in general to achieve something epistemically valuable); (ii) intellectual ethics deals with how we are engaged as intellectual beings in our cognitive performances and so shape our intellectual world and conduct our intellectual life; (iii) the ideal behind our aspiration to become intellectually autonomous is to exhibit authority over our one's cognitive performance; (iv) at least a good part of the intellectual vices we develop are rooted in heteronomy, in a lack of authority - sometimes deriving from a lack of recognition by others (Tanessini 2021) - over our doxastic performances and inquiries, and leading to defective epistemic identities that affect our epistemic agency and jeopardise the acquisition of more epistemic achievements. This is why we should care about being intellectually autonomous; it is not because it is the only way to constitute distinctive *epistemic* achievements; it is rather because otherwise we risk to lose our capacity as agents and therefore as epistemic agents.

References

Battaly, H (2022), "Autonomy and Intellectual Interdependence". In Matheson, Jonathan & Lougheed, Kirk (eds.) (2021). Epistemic Autonomy. New York, NY: Routledge: 153-172.

Carter, A. (2022), *Autonomous Knowledge. Radical Enhancement, Autonomy, and the Future of Knowing.* Oxford: Oxford University Press.

Matheson, J. (2022), "The Virtue of Epistemic Autonomy". In Matheson, Jonathan & Lougheed, Kirk (eds.) (2021). Epistemic Autonomy. New York, NY: Routledge: 173-194

Sosa, E. (2021), *Epistemic Explanations*, Oxford: Oxford University Press.

Vega-Encabo, J. (2022), "Understanding and the Value of Intellectual Autonomy". In Matheson, Jonathan & Lougheed, Kirk (eds.) (2021). Epistemic Autonomy. New York, NY: Routledge: 195-214.

José Xarez
University of Porto

Ceticismo em relação ao livre-arbítrio, responsabilidade moral sem merecimento e punição não-retributiva os problemas da proporcionalidade e da inocência

Nesta comunicação pretendo argumentar que o ceticismo em relação ao livre-arbítrio é compatível com uma justificação adequada e satisfatória da existência do sistema penal. A discussão em torno da questão do livre-arbítrio tem uma longa história filosófica. No âmago do debate contemporâneo encontramos o problema da compatibilidade do livre arbítrio com o determinismo causal. A possibilidade para o desacordo não fica, todavia, por aqui. Isto porque o conceito de livre-arbítrio não é um conceito isolável ao contexto metafísico, estando intimamente relacionado com uma série de práticas morais, e de teorias que as visam sustentar. A responsabilidade moral e as práticas sociais que nesta se baseiam ocupam um lugar de destaque na discussão e é amplamente aceite que um conceito de livre-arbítrio será tanto mais satisfatório quanto melhor cumprir o propósito da justificação das mesmas. A posição que informará este trabalho, partilhada, em grosso modo, com autores como Galen Strawson, Saul Smilansky, Derk Pereboom, e Gregg Caruso, tem no seu núcleo duas afirmações:

1) O livre-arbítrio é uma capacidade necessária para a inteligibilidade da responsabilidade moral convencional, especificamente na medida em que esta sustenta a legitimidade da atribuição de juízos de merecimento (*basic desert*).

2) Esta noção de livre-arbítrio é metafisicamente incoerente, independentemente da verdade (ou falsidade) do determinismo causal.

Ajuizar que um agente merece algo significa considerar que é intrinsecamente bom recompensá-lo, se realizou, livremente, um ato moralmente louvável, ou puni-lo, se realizou, livremente, um ato moralmente reprovável. De acordo com a posição

de que parto este tipo de juízos implica a noção de controlo último (*ultimate control*) e a tentativa de fundamentação desta capacidade conduz a uma regressão infinita. A questão central do meu trabalho prender-se-á com a viabilidade moral e prática da renúncia à conceção convencional da responsabilidade moral, na medida em que inclui juízos de merecimento. Ainda que outras formas de responsabilidade moral sejam independentes do merecimento (como a capacidade de atribuir atos a agentes, aos seus carateres e às suas disposições), a eliminação deste suscita à revisão de um conjunto significativo de atitudes e práticas morais. Alguns autores consideram que esta revisão seria catastrófica, por ser incapaz de fornecer justificações para práticas sociais indispensáveis, das quais o sistema penal é o exemplo por excelência (Smilansky, 1990;2011) (Lemos, 2016). A questão da justificação moral da punição, no contexto do sistema penal, será a principal questão visada pelo meu trabalho. Segundo as justificações retributivistas da punição o merecimento de um agente deve constar como parte da justificação da inflição da mesma, e, por conseguinte, este grupo de teorias é incompatível com a negação da responsabilidade moral convencional. As justificações consequencialistas da punição são, por sua vez, perfeitamente conciliáveis com esta visão acerca do livre-arbítrio e da responsabilidade moral. São, porém, frequentemente consideradas incapazes de salvaguardar intuições éticas essenciais neste contexto. No cerne deste tipo de objeções podemos encontrar duas intuições fundamentais:

a) A intuição segundo a qual deve existir uma proporcionalidade entre a punição e o ato, que se calibre, pelo menos em parte, no merecimento do agente.

b) A intuição segundo a qual, *ceteris paribus*, a punição de um inocente é moralmente pior do que a punição de um culpado, devido a diferenças no merecimento destes agentes.

Recentemente têm sido desenvolvidos modelos que procuram resistir a estas objeções sem apelar a noções retributivistas, tais como o modelo da quarenta de Gregg Caruso e Derk Pereboom, o modelo da correção de Michael Corrado e o modelo neo Kantiano de Benjamin Vilhauer. Neste trabalho pretendo argumentar que estes modelos não conseguem integrar a intuição b). A objeção que se baseia nesta premissa, tipicamente endereçada a posições não-retributivistas, salienta a relação meramente indireta entre o valor da punição de um agente julgado culpado por uma ação moralmente reprovável, e o facto desse agente ser julgado culpado. Argumentarei, porém, que isto não é particularmente problemático, por dois motivos:

1) A inclusão de considerações consequencialistas, ou contratualistas, é suficiente para que, na maioria dos casos, esta distinção se mantenha relevante, uma vez que a punição de inocentes se revela ineficaz a alcançar os vários propósitos não-retributivos da existência do sistema penal (nomeadamente a dissuasão e a reabilitação).

2) A existência de casos extremamente específicos, nos quais esta distinção é sobreposta por outras considerações, não é limitada a modelos não-retributivistas, e por isso não se afigura uma desvantagem específica destes últimos.

Por considerar também que os modelos que referi fornecem princípios de proporcionalidade não-retributivistas satisfatórios, concluo que a visão que defendo acerca do livre-arbítrio e da responsabilidade moral é compatível com a possibilidade de uma justificação adequada para a existência de práticas penais em sociedade.

Bibliografia

1. Caruso, Gregg. 2021. *Rejecting Retributivism: Free Will, Punishment, and Criminal Justice*. New York: Cambridge University Press.

2. Caruso, Gregg; Elizabeth, Shaw. 2019. *Free Will Skepticism in Law and Society: an overview*. New York: Cambridge University Press.

3. Caruso, Gregg; Morris, Stephen. 2017. "Compatibilism and Retributivist Desert Moral Responsibility: On What is of Central Philosophical and Practical Importance". *Erkenntnis*: 837-855.

4. Corrado, Michael. 2015. "Punishment, quarantine, and preventive detention". *Criminal Justice Ethics*: 3-13.

5. Corrado, Michael. 2019. "Criminal Quarantine and the Burden of Proof". *Philosophia*: 1095-1110.

6. Corrado, Michael. 2021. "The limits of the state's power to control crime". *Journal of Legal Philosophy*: 126-131.

7. Focqueart, Farrah. 2018. "Free Will Skepticism and Criminal Punishment: A Preliminary Ethical Analysis". Em *Neuroexistentialism Meaning, Morals, and Purpose in the Age of Neuroscience.* New York: Oxford University Press.

8. Greene, Joshua; Cohen, Jonathan. 2004. "For the law, neuroscience changes nothing and everything." *The Royal Society:* 1775-1785.

9. Honderich, Ted. 2006. *Punishment: The Supposed Justifications*. London: Pluto Press.

10. Lemos, John. 2016. "Moral concerns about responsibility denial and the quarantine of violent criminals". *Law and Philosophy:* 461-483.

11. Levy, Neil. 2012. "Skepticism and Sanction: The Benefits of Rejecting Moral Responsibility". *Law and Philosophy:* 477-493.

12. Levy, Neil. 2015. "Less Blame, Less Crime? The Practical Implications of Moral Responsibility Skepticism". *Journal of Practical Ethics:* 1-17.

13. Pereboom, Derk. 2001. *Living Without Free Will*. Cambridge: Cambridge University Press.

14. Pereboom, Derk; Caruso, Gregg. 2018. "Hard-Incompatibilist Existentialism: Neuroscience, Punishment, and Meaning in Life". Em *Neuroexistentialism Meaning, Morals, and Purpose in the Age of Neuroscience.* New York: Oxford University Press.

15. Schoeman, Ferdinand. 1979. "On Incapacitating the Dangerous". *American Philosophical Quarterly*: 27-35.

16. Smilansky, Saul. 1990. "Utilitarianism and the 'Punishment' of the Innocent: The General Problem". *Analysis*:256-261.

17. Smilansky, Saul. 2011. "Hard Determinism and Punishment: A Practical Reductio". *Law and Philosophy*: 353-367.

18. Vilhauer, Benjamin. 2011. "Persons, punishment, and free will skepticism". *Philos Stud*: 143-163.

19. Vilhauer, Benjamin. 2012. "Taking Free Will Skepticism Seriously". *Philosophy Quarterly*: 933-852.

20. Vilhauer, Benjamin. 2013. "The People Problem". Em *Exploring the Illusion of Free Will and Moral Responsibility*. New York: Lexington Books. 21. Vilhauer, Benjamin. 2015. "Free Will and the Asymmetrical Justifiability of Holding Morally Responsible". *The Philosophical Quarterly*: 772-789. 22. Vilhauer, Benjamin. 2023. "Free Will Skepticism and Criminals as Ends in Themselves". Em *The Palgrave Handbook on the Philosophy of Punishment*. Washington: Palgrave Macmillan.

Erica Yu

Erasmus University Rotterdam

From Signed Orders to Committee Rankings

Consider the following example:

> Eleanor is asked to rank two alternative committees against each other. The first is composed of her favorite candidate, Anna, and a candidate she greatly dislikes, Daphne. The second is composed of a candidate, Betta, whom she likes less than Anna but would still be happy to have in the committee, and Chiara, whom she is indifferent about whether to include or exclude in the committee.

How should Eleanor rank the committees {Anna, Daphne} vs. {Betta, Chiara} against each other based on her preferences over the candidates?

What information do we have about Eleanor's preferences over the four candidates? First, we can say that she prefers Anna the most, followed by Betta, Chiara, and then Daphne. Second, we can also say that she approves of Anna and Betta to be included in the committee, disapproves of Daphne to be included in the committee, and is indifferent to Chiara being included in or excluded from the committee.

Notice that this information is *not* enough to be able to make a comparison between the two committees. We would need to know how much Eleanor disapproves of Daphne relative to how much she approves of Anna or Betta. For instance, if Eleanor disapproves of Daphne more than she approves of Anna and Betta, then {Betta, Chiara} should be ranked higher than {Anna, Daphne}. But in all other cases, we would need a greater specification of the "value" attached to the approvals and disapprovals of candidates based on their relative ranking.

The questions that I address in this paper are the following: How do we capture these preferences over candidates? How can these preferences be extended to preferences over committees (i.e., sets of candidates)? The signed orders framework proposed by Fishburn (1992) is able to provide a starting point for answering these questions.

In the signed orders framework, inclusions and exclusions of candidates to the committee can be ranked against each other. For example, one can say that "excluding Daphne from the committee is preferred to including Anna in the committee". The ranking of a candidate's inclusion in and exclusion from the committee also enables a partitioning into positive and negative sets of candidate inclusions and exclusions. For example, since Anna and Betta are preferred to be included in the committee rather than excluded from it, then including them in the committee is in the positive set of inclusions. Including Daphne, on the other hand,

is in the negative set because she is preferred to be excluded from the committee rather than included in it.

How do we then extend these signed orders into rankings over committees (i.e., sets of candidates)? The extension strategy I propose proceeds as follows: First, I assume that in any committee, including candidates whose inclusions are in the positive set and excluding candidates whose inclusions are in the negative set are desirable. Given the example above, this assumption entails that including Anna and Betta in and excluding Daphne from the committee is desirable. Second, based on this assumption, we can then assign positive value to including candidates whose inclusions are in the positive set and to excluding candidates whose inclusions are in the negative set. In the example, positive value is then attached to including Anna and Betta in and excluding Daphne from the committee. Third, I argue that the magnitude of these values should be dependent on the rankings of candidate inclusions and exclusions. For instance, if excluding Daphne is preferred to including Anna, which is preferred to including Betta, the values attached to these inclusions and exclusions should reflect this ordering. Finally, I argue that committees be ranked against each other based on the summed values of their desirable components.

In line with this extension strategy, I then characterize the Signed Borda ranking, which is a modification of the Borda ranking over sets of objects by Darmann and Klamler (2019). In essence, a Borda score is assigned to each positive candidate inclusion and each negative candidate exclusion based on their relative rankings against each other, and these scores are summed to get the value of a committee. I show that the Signed Borda ranking is characterized by four axioms, namely the extension rule, strict independence, trading, and irrelevance from indifference.

In summary, this paper proposes a framework which allows for more informational richness to be accounted for in representing individuals' preferences over the inclusion and exclusion of candidates to a committee. I then show how these preferences can be extended to rankings over committees, and characterize a signed version of the Borda ranking.

References

Fishburn, Peter C. "Signed orders and power set extensions." *Journal of Economic Theory* 56, no. 1 (1992): 1-19.

Darmann, Andreas, and Christian Klamler. "Using the Borda rule for ranking sets of objects." *Social Choice and Welfare* 53, no. 3 (2019): 399-414.

Tarek Yusari
University of Liverpool

with

Daniel J. Hill
University of Liverpool

Stephen K. McLeod
University of Liverpool

Attila Tanyi
The Arctic University of Norway

What is (Distinctively) Wrong about Entrapment?

Entrapment occurs whenever one party, the *agent*, intentionally brings it about that another, the *target*, commits a criminal offence, intending to have the target prosecuted for committing the offence. *State entrapment*, which is of uppermost concern here, happens when the agent is, or is a deputy of, a law enforcement agent. When the agent lacks this status, we have *private entrapment*.

Scenarios of state entrapment and complicity scenarios are treated differently by the courts.

At least some instances of entrapment are commonly considered morally problematic, warranting a judicial response, to the target's benefit, called a 'remedy' (e.g., mitigation at sentencing, restriction on the admissibility of evidence, or acquittal).

When an individual—the *principal*—is induced (short of inducement by means of threats of grievous bodily harm or death) by another party—the *accessory*—to commit a criminal offence, this is not considered to warrant any remedy that benefits the principal.

There has been no convincing account in the literature of the justification, if any, of the differential treatment between the entrapment and complicity scenarios and of the conditions that may license it. It is plausible, however, that whether the differential treatment is warranted depends on whether there is an objectionable or wrong-making feature that is *distinctive* to entrapment.

The main contention of the paper is that, if there is anything distinctively wrongful or otherwise objectionable about entrapment, it must depend upon entrapment's distinguishing feature relative to complicity scenarios. This feature is the agent's

law-enforcement aim: in entrapment scenarios, the agent intends to have the target prosecuted for having committed the criminal offence.

The entrapment literature struggles to offer a persuasive account of the distinctive wrongfulness of entrapment. We argue that this is because it fails adequately to engage with the law-enforcement aim as entrapment's distinctive feature.

We begin by critically engaging with several accounts of the moral wrongfulness of entrapment. On some accounts, the wrongfulness of entrapment inheres in a feature that is (often) common to entrapment and complicity scenarios. For instance, on one view, the agent *subverts* the target's moral capacities, increasing the target's likelihood of culpably choosing to offend. On another view, the agent becomes *morally allied* with the target's unlawful or wrongful conduct. On various versions of a third view, the agent *tempts* or *manipulates* the target, thereby undermining the target's moral autonomy to an extent that the target's culpability is diminished or nullified.

Their substantive merits notwithstanding, these accounts fail to explain why remedies that benefit the defendant are fitting for entrapment scenarios but not for complicity scenarios.

Other views misidentify what is distinctive about entrapment. For instance, some argue that the differential treatment between state entrapment and complicity scenarios is apt because of the agent's status, as an agent of the state, in the former. We argue against this view.

First, the features associated with the agency of a public official in an entrapment scenario are either inessential to entrapment or may be shared by a private entrapper. Second, although the moral position of the state vis-à-vis an individual is different to that of private persons with one another (for instance, it is plausible that the state has more stringent duties than private persons to steer others away from wrongdoing, criminal or otherwise), these differences do not warrant treating entrapment scenarios differently depending solely on the agent's status. Third, it is true that by inducing the target's offence, the agent is in most relevant respects an accomplice or accessory to it. Thus, it may be thought that, when the agent is a state agent, the state's standing to prosecute and convict the target is undermined, whereas this is not true when the agent is a private person. This claim, however, speaks only to a defect in the state's standing rather than to a wrong-making feature that inheres to the entrapping conduct itself. Indeed, the objection to entrapment remains when agent and prosecuting/adjudicating body are entirely separate entities.

We then challenge some accounts that, while they do focus upon the entrapping agent's law enforcement aim, we consider mistaken about the moral import of this feature. One of these accounts has it that the target is wrongfully manipulated because the agent purposefully hides their law enforcement aim from the target. We argue, however, that the target has no right not to be deceived about the risk of capture, prosecution, and punishment ensuing from the prospective commission of a criminal offence. Moreover, the target is not wronged by the agent's deception as to their law enforcement intention.

The paper suggests, instead, that the distinctively objectionable feature of the entrapment scenario inheres in the *anticipatory* character of agent's intention to prosecute and punish the target for an offence yet to be committed. This means that entrapment, contrary to many accounts, is always—and distinctively—*pro tanto* wrongful, no matter the target.

We then ask whether any factors *defeat* this objection to entrapment, thereby making some acts of entrapment permissible, and thus rendering it sometimes appropriate to treat the target just as the principal in a complicity scenario. To answer this question, we again focus on the law-enforcement aim. After rejecting the common view than entrapment is permissible if the target would have still committed an offence of the same kind as the entrapped offence absent the inducement of the agent, the paper develops one such defeater: under some conditions, those who have become liable to punishment may also become liable to be entrapped as a means of bringing them to justice. Still, a morally problematic feature is retained in the form of a *discontinuity* between the grounds for the agent's conduct (i.e., the target's past, unpunished, offences) and the grounds (i.e., the entrapped offence) for the claim advanced in the criminal trial towards the target's conviction and punishment. This discontinuity, we argue, violates a core tenet of procedural justice in the criminal law.

Lastly, based on our account of the distinctive wrongfulness of entrapment, we provide a new defence of the widespread claim that the appropriate remedy for impermissible entrapment is a permanent stay of proceedings.

Elia Zardini

Complutense University of Madrid

Vagueness across the Type Hierarchy

When is an entity vague (or precise)? Perhaps a natural answer would be to say that a property is vague iff it (possibly) presents borderline cases (and precise otherwise), but such an answer is problematic on at least two counts. Firstly, it is not clear how to generalise the answer to other types of entities such as *e.g.* objects. Secondly, the answer overgenerates as it also makes vague *e.g.* the paradigmatically precise property *x*-is-a-straight-line (for there might be concrete lines that are borderline straight). This paper critically examines one influential proposal for characterising vagueness across the type hierarchy and then offers an alternative. While most of the discussion will centre on the task of simply providing a nontrivial necessary and sufEicient condition for an entity to be vague (using 'characterisation' as a shorthand for such a condition), some remarks will also be made concerning the more ambitious task of providing an analysis of what it is for an entity to be vague (using 'deEinition' as a shorthand for such an analysis).

According to an influential proposal going back at least as far as Rolf [1980] (and recently defended *e.g.* by Bacon [2018]), we should take the notion of vagueness as primitive for some types (say, objects and propositions) and characterise vagueness for other types by saying that an entity is vague iff it takes at least one precise input and yields a vague output. For example, assuming that 1 is precise and that the proposition <1 is small> is vague, this Rolf-style characterisation correctly implies that the property *x*-is-small is vague.

I'll argue that the Rolf-style characterisation embodies an objectionably "purist" conception of vagueness. For example, consider a property ("schbaldness") taking any precise object *x* to yield, say, <*x* is a number> (plausibly assuming that the property *x*-is-a-number is precise) and any vague object *x* to yield <*x* is bald>. Schbaldness would seem vague, for, say, it takes a man, Harry, whose vagueness (we may so suppose) only resides in the vagueness of where its right toe ends and who has 50,000 hairs, to yield the borderline and therefore vague <Harry is bald>. If taking Harry to yield <Harry is bald> is sufficient for baldness to be vague (and it is!), how could it not be sufficient for schbaldness to be vague? Where else could the vagueness of <Harry is schbald> come from, if not from the vagueness in schbaldness (the only other entity at play is Harry, but <Harry is schbald> is vague because <Harry is bald> is, and Harry's vagueness resides in a feature that is totally irrelevant for the vagueness of the latter proposition)? However, schbaldness is precise on the Rolf-style characterisation, for it takes any precise object *x* to yield the precise <*x* is a number>.

This train of thought leads to the issue that, on the Rolf-style characterisation, it is not even clear that baldness is vague, since objects capable of having hair on their scalp and for which therefore the question of baldness could arise are typically—and, one may well suspect, invariably—vague (and those of them that are vague are anyway those that paradigmatically support the idea that baldness is vague). Typical precise objects (such as numbers, graphs, points in space *etc.*) are not objects capable of having hair on their scalp and for which therefore the question of baldness could arise, and, even granting the possibility of precise objects that are capable of having hair on their scalp and for which therefore the question of baldness could arise, such extravagant objects are certainly not necessary for supporting the idea that baldness is vague. Nor, for analogous reasons, is it clear that a paradigmatically vague object like Kilimanjaro is vague, since properties nontrivially applying to a mountain are typically—and, one may well suspect, invariably—vague (and those of them that are vague are anyway those that paradigmatically support the idea that Kilimanjaro is vague). For example, properties of the kind *x*-is-at-least-*i*m-high paradigmatically support the idea that Kilimanjaro is vague, but, *pace e.g.* Bacon [2018], these are arguably vague, as manifested by the following kind of series: start with a *i*m-high mountain with a thin protuberance rising up to *i*+1m, and then gradually enlarge the protuberance, eventually ending up with a *i*+1m-high mountain.

Turning now to my favoured alternative, let a soritical series for an entity be a series along a dimension relevant for the entity's presence (*i.e.*, depending on the entity's type, its existence or occurrence or application *etc.*), where at the start the entity is clearly present while at the end it is clearly not present, and where each successive case in the series represents a tiny worsening of the conditions for the entity's presence. Further, let an entity lack a sharp boundary on a soritical series for it iff, for no pair of adjacent cases in the series, the entity is present in one and not present in the other. Then, the same characterisation of vagueness that many have thought to apply for properties can be defended to apply to all other types as well: just as a property is vague iff it (seemingly) possibly lacks a sharp boundary on some soritical series for it, so is any entity of any other type. (I'd propose the version with 'seemingly'—understood epistemically rather than psychologically in terms of *prima facie* justification—as a characterisation, whereas, within the nontransitive system developed in earlier works, I'd propose the version without 'seemingly' as a deEinition.) It's true that, in the case of *e.g.* objects, for different cases, the (seeming) possible lack of a sharp boundary is realised on different dimensions (spatial, temporal, mereological *etc.*) and, for each particular case, good judgement is needed to set up a compelling soritical series for it manifesting such lack, but so it is also in the case of properties (because of their pervasive multidimensionality).

In conclusion, on this view, there is one single nonprimitive notion of vagueness—(seeming) possible lack of a sharp boundary—that gets realised in different irreducible ways among and within different types, as opposed to the Rolf- style characterisation, on which there are primitive separate notions of vagueness for certain types to which vagueness of all other types is reduced (plus, as indicated, the proposed characterisation can be turned into a much more satisfying definition than the Rolf-style one).

References

Bacon, A. [2018], *Vagueness and Thought.* Oxford University Press, Oxford.

Rolf, B. [1980], 'A Theory of Vagueness'*, Journal of Philosophical Logic* **9**: 315–325.

Dan Zeman

Slovak Academy of Sciences

A Polysemy Account of Evaluative Terms

There are many natural language expressions that we use to evaluate parts of the world – people, things, situations, events, courses of action, etc. Take, for example, *slurs*: their function is to convey certain negative evaluations towards people on the basis of belonging to a certain social group (Hom (2008), Anderson & Lepore (2013), Camp (2013), Jeshion (2013, 2020), etc.). Predicates of taste ("tasty", "disgusting", "fun"), aesthetic terms ("beautiful", "ugly"), moral terms ("good", "bad", "ought to") form the class of *evaluatives* and have been at the center of much discussion in recent semantics (Kölbel (2004), Lasersohn (2005, 2016), Stojanovic (2007), Cappelen & Hawthorne (2009), MacFarlane (2014), Silk (2016), Zakkou (2019), etc.). A subset of these are *thick terms* – "courageous", "lewd", "balanced", "symmetrical", etc. (Hare (1963), Gibbard (1992), Väyrynen (2015), Cepollaro (2020), etc.). *Dual-character expressions* – "mother", "father", "teacher" and many other vocational terms – have also been considered evaluative (Knobe, Prasada & Newman (2013), Leslie (2015), Del Pinal & Reuter (2015, 2017), Reuter (2019), etc.). Finally, *expressives* – from interjections like "oops", "ouch" and "damn" to more loaded terms like "jerk" or "asshole" (Kaplan (ms), Potts (2005), Richard (2008), Predelli (2013), Gutzmann (2015), Berškytė & Stevens (2019), Marques (2021), etc.) – are perhaps the best example of evaluative expressions due to their more direct way of conveying evaluations.

What most of these expressions have in common is that they encode both descriptive information *and* evaluations. Providing a satisfactory account of this double aspect has not been an easy task within traditional truth-conditional semantics. In addition, many of the expressions above can be used in conformity with the type of evaluation they encode, but also with a switched one ("sick", "wicked", etc. are vivid examples of expressions used to convey positive evaluations while encoding a negative one). Interestingly, they can also be used *neutrally*: that is, as mere descriptions or as stating matters of fact. Several neutral uses of slurs have been attested: Anderson's (2018) "referential use" and Zeman's (2022) "identificatory use" are just two examples. In relation to thick terms, it has been stressed that they can be used in a purely descriptive manner in certain contexts (e.g., "balanced" can be used simply to describe a work of art, without casting an evaluation), and similarly for dual-character expressions ("father" is frequently used in the biological sense only). In relation to the class of evaluatives generally, it has recently been pointed out that they give rise to indeterminate evaluations in certain contexts, or to no evaluation at all (Stojanovic & Kaiser (2022) call the set exhibiting this trait "middling adjectives"). Finally, expressives can be used neutrally due to repeated use that erases their negative connotations.

Neutrality seems thus to be exemplified by all the expressions mentioned. This raises an additional challenge for theories of evaluative expressions.

In this paper, I aim to address this challenge by proposing a unitary treatment of all the expressions above that accounts for their double aspect and for their various uses (including neutral ones), based on the idea that the evaluative expressions mentioned are polysemous. I implement this idea by using a framework used in lexical semantics – what is known as "rich-lexicon" theories. One of the main applications of such theories is to nouns, in relation to which rich-lexicon theories postulate various *dimensions* of meaning, among which the one that an expression has in a certain context (its "sense") is selected. How to determine the meaning dimensions that constitute a noun's lexical entry is a crucial issue; one common way to do so has been to appeal to the old Aristotelian idea of qualia: units of meaning that encapsulate certain types of information and which are interrelated in certain ways (Pustejovsky (1995)). Thus, following the work of various lexical theorists (and in particular the framework proposed by Del Pinal (2018)), I take the lexical entry for nouns to comprise at least the following: perceptual information about the objects referred to with the noun (call that dimension PERCEPTUAL), information about what those objects are made of or their parts (CONSTITUTIVE), how they came to being or the purpose of their creation (AGENTIVE), their typical function (TELIC), etc. As for the mechanisms responsible for the selection of senses, there are many options on offer (Pustejovsky (1995), Frisson (2009), Asher (2011), Schumacher (2013), Del Pinal (2018), etc.), but I will remain neutral on this issue and use the placeholders "foregrounding" and "backgrounding" instead.

When it comes to evaluative expressions, the application of the framework requires the introduction of at least one meaning dimension that captures their evaluative character. Thus, in relation to slurs, Zeman (2022) has proposed the introduction of EVALUATIVE, a meaning dimension comprising attitudinal elements and evaluations, but also more objective, socially relevant ones such as stereotypes, perspectives, ideologies, etc. The lexical entry for slurs would then comprise some of the usual meaning dimensions plus the new one (and perhaps other novel ones). This can easily be generalized in application to all the other evaluative expressions, with the exact components of EVALUATIVE varying with the expression at stake (for example, thick terms and expressives don't seem to be dependent on ideologies). What is important is that, by the same mechanisms used to select a particular sense of a noun, different senses of an evaluative expression are selected in different contexts. To illustrate again with slurs, when a slur is used derogatorily, the EVALAUTIVE dimension is foregrounded; in contrast, when the slur has a neutral use, it is backgrounded, and others (descriptive ones) are foregrounded. Such a framework also has the capacity to deal with evaluation switches of the kind mentioned (which is also what happens when a slur is reclaimed, for example) by indexing the evaluation to a contextual *valence* parameter. One should of course pay close attention to the type of expression dealt with – not all evaluative expressions are nouns ("tasty", for example, is an adjective; etc.), and thus the meaning dimensions postulated should vary accordingly. However, variation along this line doesn't hinder the outcome of this application of the rich-lexicon framework, which yields a highly economical, unitary theory that accounts for the double aspect and the main uses of a large portion of evaluative expressions in natural language and is

based on a framework from lexical semantics that is widely accepted and has been proven theoretically very fruitful.