BOOK OF ABSTRACTS

The Dutch Research School of Philosophy (OZSW)
10th Annual Conference 2023

June 2-3, 2023
Leiden University

A change of perspective: Evaluating sequences of choice in behavioural paternalism

Måns Abrahamson

People make context-dependent choices, in the sense that what option they choose often depend on welfare-irrelevant contextual cues such as the salience of different options or whether a particular background music is playing when making a choice. This presents a challenge and an opportunity for paternalist interventions: (i) the context-dependence highlights that people's preferences are unreliable guides to what is in people's best interest, because they will support different welfare judgements over the same options depending on the context in which they are revealed; (ii) the context-dependence shows that paternalists have access to a (supposedly) non-coercive and cheap tool to shape people's behaviour: altering contextual cues in the choice architecture.

Behavioural paternalism (BP) draws on these insights. BP is concerned with altering the choice architecture in order to steer people in the direction of desired behaviour. These interventions are justified by an appeal to people's 'true', rather than revealed, preferences—the preferences people would act on under idealised conditions, without reasoning imperfections, failures of executive control, and faulty information. The assumption is that these true preferences are subjective, even though idealised, and that intervening in people's choices to make people better satisfy their true preferences will therefore be an innocent form of means paternalism.

Recent critiques have challenged this approach. The criticism holds that people's true preferences will sometimes be context-dependent in the same way as people's revealed preferences; therefore, true preferences cannot provide the evaluative standard BP is looking for when arbitering between context-dependent choices. The implication is that BP's appeal to true preferences when steering people in the direction of a particular option consists in nothing more than policymakers imposing their own view on the good life on people: BP amounts to objectionable ends paternalism and not means paternalism as suggested by the approach's proponents.

In this article, I challenge the scope of this criticism by contesting an underlying assumption of the critique: that BP interventions should be evaluated from the perspective of individual instances of choice. I argue that many of the choice domains BP is concerned with are not characterised by one-shot decisions but repeated choice. These domains are better evaluated from the perspective of people's sequences of choice. This is not only supported by the aims of policymakers, who are concerned with broader policy outcomes, but is philosophically more appealing given the role personal values are taken to play as fundamental building blocks of true preferences by both proponents and critics of BP. The logic of most personal values is that they can only be fulfilled over time through persistent effort, while individual deviations will not have significant impact on their fulfilment. By drawing on sequences of choice, behavioural paternalists can therefore be more discerning in their welfare judgements by excluding certain choice patterns as mistakes on subjective grounds. I further elaborate on how this change in perspective can be operationalised by eliciting people's stated preferences over choice architecture, which empirical evidence suggest is not context-dependent in the same ways as people's individual choices of options.

ID: 170 / **Panel 5-3-E: 1**
Individual paper
Topics: Philosophy of Science, None of the topics above (please include your topic under 'keywords')
Keywords: meta-philosophy; empirical philosophy; citation analysis; philosophical method

Are philosophers doing something different now? A citation analysis

Till Armbruster, **Paul Rehren**

Many philosophers agree that (like it or not), empirical research has become a much more prominent part of philosophical inquiry in recent decades. For example, in 2008, Prinz declared that philosophy was in the midst of "a methodological revolution," one marked by a sharp increase in the number of philosophers who "integrat[e] lessons from various branches of psychology [...], neuroscience [...], evolutionary theory, experimental economics, and other 'scientific' fields" into their work (189; also, see, Sytsma and Livengood 2015, 3–6). Similar stories have been told about various subareas of philosophy, like philosophy of mind (Knobe 2015), ethics (Doris 2010, 2-3), philosophy of technology (Verbeek 2022) and philosophy of language (Hansen 2014, 1).

Despite this agreement, however, there is little systematic evidence about when, where and in what way such a revolution took place—if indeed, it did. One exception is Knobe (2015), who investigated the role of empirical results in the philosophy of mind through a comparison of a sample of highly cited papers from the twentieth century with a similar sample from 2009-2013. While Knobe's study did find a clear trend towards more engagement with empirical results, leading Knobe to proclaim that "[p]hilosophers are doing something different now" (36), it suffers from major limitations, including limited scope (focussing only on the philosophy of mind), small and non-representative samples, coarse temporal resolution and a complete lack of information on what fields the empirical results philosophers engage with are from.

To provide a more detailed and complete picture of the role of empirical results on academic philosophy, we make use of the advantages of citation analysis (Zhao and Strotmann, 2015). We collected a large corpus of articles published in three high-profile generalist philosophy journals (Noûs; Philosophical Studies; Synthese) over a 50-year period, starting from 1971. Using a mix of automated and manual methods (Keil et al., 2023), we then extracted reference lists from these articles, and associated each reference with its discipline(s) based on the publication it appeared in (Zhang et al. 2021). We also tracked down the PhilPapers categories of each base article (where available), allowing us to compare citation practices across different subareas of philosophy.

Our results show that while there have been minor trends for a few individual fields (e.g., mathematics, neuroscience, psychology), overall, the rate at which philosophers cite articles from the sciences has not undergone any drastic change in the last 50 years. We also did not find clear signs that some subareas of philosophy (e.g., philosophy of mind) are outliers in this respect. Our findings indicate that "empirical philosophy" may represent a less disruptive turn than is often suggested, and that talk of a "methodological revolution" (both, Prinz 2008, 189) may have been overblown. We will discuss potential implications of these findings for the self-image of philosophy, and for philosophical methodology.

ID: 167 / Panel 4-3-C: 1
Individual paper
Topics: Intercultural philosophy, Social and political philosophy, Critical Theory
Keywords: Black Male Studies, Post-intersectionality, Intimate Violence, Decolonial Theory, Critical Race and Gender Theory

The unspeakable unspoken: Black male vulnerability to intimate violence

<u>Thomas Baruzzi</u>

What would make Erin Pizzey, founder of the world's first and largest domestic violence shelter for women in 1971, proceed to denounce feminism? She was promptly written out of its history including that of the shelter, which nowhere mentions her name. Similarly, what would lead Professor Emeritus of psychology, Donald G. Dutton, to completely rethink intimate partner violence in the early 2000s? In particular, given his past filing a government report demanding a more aggressive response to wife assault and, for about 15 years, providing therapy to male batterers in court-mandated treatment programs. For Erin Pizzey, it was that "women and men are both capable of extraordinary cruelty". Ex-senator Anne Cools who, following Pizzey's footsteps, ran the first domestic violence shelters in Canada, concluded idem.4 So did Prof. Dutton, writing that "women are as violent as males".

This perspective is diametrically opposed to the standard creed of the patriarchal, heteronormative relationship as a social hierarchy in which all (cisgender) males are the violent oppressors of all (trans)women and females the world over. According to this view, any even equivalent cross-over between the two sides is either explained away as exceptional or incorporated into the system. Yet, if such influential figures as those above, once strongly committed to feminist ideology, took such drastic turns in their convictions and careers, then the premises of mainstream gender theory must be taken to task. Assuming and proliferating them uncritically risks completely misrepresenting and thus failing to properly address the horrid phenomenon of domestic violence which implicates and affects both males and females, adults and children.

My paper takes domestic violence to be a women's issue, but not only. Indeed, gendered or not in its ontology, it has many gendered implications, even on the opposite camp. Specifically, using insights from the new paradigm of "Black Male Studies" coined by famed Africana professor, Tommy J. Curry, I make the case that intimate (familial/partner, physical/sexual) abuse has a significant effect not just on males generally, but particularly on subordinate-group, racialized (Black) men and boys. I cite extensive research in the U.S. and Latin America, where 'machismo' is deemed rampant. This, I show, has been actively censored, redacted, or ignored by gender theorists in their writings, and yet makes important revisions to the normative thread of gender and race studies theorizing Black women as the prime intersectional subjects. In doing so, I expose the Black male's vulnerability under white patriarchal-colonial societies to rape by both whites and Blacks, men and women, and on why this violence is not just unspoken, but maintained by modern gender and race theorists in normative theory as something unspeakable.

Extreme beliefs as morally serious, but purely epistemic wrongs

Lisa Bastian, Philip Robichaud

Epistemic appraisal of those with extreme beliefs can be straightforward. Consider agents who believe that Jews are controlling the media, or that women are morally inferior to men. They are holding beliefs that are false and, maybe, unjustified. However, a more complex and arguably more important kind of appraisal of extreme believers involves issuing moral verdicts. We may also want to say about the antisemite and the misogynist that their beliefs speak badly of them as moral rather than just epistemic agents. We may want to say that they are morally and not just epistemically blameworthy in virtue of what they believe.

The most straightforward way to establish that agents can be morally responsible for their extreme beliefs is provided by doxastic wronging (DW) views, which hold that, in at least some cases, it can be morally wrong to hold certain beliefs. An agent who believes morally wrongly without excuse or exemption is non-derivatively blameworthy and hence responsible.

More specifically, DW views hold that in the relevant cases, there can be moral reasons against holding a belief that might otherwise epistemically justified. Assuming they are correct, this leaves open the question of what to do next: should the agent form a different belief, or suspend judgment? Given the societal relevance of extreme belief cases, it is especially important that these views give some form of positive guidance that goes beyond merely prohibiting certain beliefs.

This leads to the first argument of this paper, namely that, given a few plausible assumptions, DW views are committed to the further claim that a moral reason against a certain belief also functions as a reason for another belief (and in some cases for suspending judgment).

This, however, raises another problem. A careful examination of how agents are able to respond to moral reasons for belief (or suspension) reveals a difference between forming a belief on the basis of an epistemic reason and forming a belief on the basis of a moral reason. While it is generally possible to be motivated to form a belief on the basis of an epistemic reason this need not be the case for moral reasons. In cases where we must choose between B(p) and B(not-p), and we have a moral reason against B(p), then, ceteris paribus, it is permissible to B(not-p). Yet, we are unable to be motivated to B(not-p) for the relevant moral reason. In this scenario, it is permissible but not possible to hold B(not-p) on the basis of the moral reason.

The remainder of the paper explores the options left open by this perplexing result and concludes that we are better off rejecting DW accounts. We then show that by keeping moral and epistemic appraisals separate we can offer a better account of how those who correct their extreme beliefs can be understood as taking responsibility for being morally better believers and how they can atone3 for their prior false beliefs.

ID: 193 / Panel 3-2-D: 1
Individual paper
Topics: Social and political philosophy, Ethics (general and applied), Metaethics, Moral Psychology
Keywords: human rights; climate change; environment; individualism

Diachronic Rights

Jelena Belic

Human rights are often said to be unsuitable to protect a range of important interests set back by climate change due to their individualistic, temporal, or anthropocentric limitations. To overcome these limitations, it is argued that rights need to be granted to a broader set of right holders including collectives, future generations, or nature as a whole (or at least some parts of it). Although the focus on right holders is an important avenue for rethinking the concept of rights, it falls short of questioning the normative foundations of human rights; that is, what is it that these rights protect. In this paper, I undertake such questioning by attempting to develop a concept of diachronic rights. The concept relates to the temporal dimension of the interests that are protected by human rights. I contend that a great deal of theory as well as practice of human rights assumes what some philosophers term temporal individualism (J. O'Neill 1993; Scheffler 2021). In its essence, temporal individualism amounts to taking a short-term perspective on individual interests by presupposing a separation between past, present, and future generations. In the context of human rights, temporal individualism has significant implications as it allows the interests of the members of the present generation to take priority over the interests of future generations (Gardiner 2013). Temporal individualism, however, need not be taken for granted as it presupposes a particular metaphysical account of personhood. This account has already been questioned and arguments have been made in favor of a wider conception of personhood according to which personhood extends beyond individual biological boundaries to encompass the concern for both past and future generations (Heyd 1992; Meyer 1997). The wider conception of personhood presupposes that individual interests have a diachronic dimension. As such, it can have important implications for rethinking human rights. A preliminary thought is that at least some human rights are diachronic (such as the right to a clean and healthy environment) and as such do not lend themselves to intergenerational conflicting claims in the way other rights do.

ID: 198 / Panel 5-1-G: 1
Individual paper
Topics: Social and political philosophy, Philosophy of Language
Keywords: Normative Concepts, Political Philosophy, Metaphilosophy, Conceptual Revision

Political concepts and normative overextension

Nadia ben Hassine

Normative concepts are distinguished by their evaluative role. We use normative concepts not only to describe and categorize, but also to appraise and condemn. The origins of such normativity is often suggested be a concepts' use: normative concepts play a certain evaluative role. Under this view, we can identify normative concepts by picking out the normative role with which they are semantically associated. Through these semantic ties with practical uses, the concepts' evaluative component is carried across a wide range of contexts. Whatever you ascribe the predicate "dismal" to, for example, will express something negative, because that is the identifiable normative role associated with this concept. In this paper I will be building on a concern with this view, namely the worry that taking a normative role to be "tied" to a concept will result in overlooking important divergences from the associated normative convention.

I will be focusing on the importance of normative flexibility for normative concepts in political philosophy. The normative component of political concepts such as "equality", or its inverse, "inequality", play a role in recommending societal interventions. This goes beyond political philosophy, with many institutions dedicating themselves to, for instance, the promotion of equality of opportunity. Yet, simultaneously, there are spaces where the concept "unequal" does not carry such a strong negative force, for instance in cases where local inequality is justified through its alleviation of wider social asymmetries. Think, for instance, of cases where quotas are enforced to increase accessibility of the job market for previously disadvantaged social groups, or suggestions to place higher taxes on those with higher incomes. Although these cases may fit certain descriptive features of a given equality-concept, they do not fit its commonly associated normative component. In this paper I will identify and map such phenomena as cases of normative overextension, these are cases where a straightforward extrapolation of the associated normative role of a concept do not generate a fully considered evaluation.

In developing an account of normative overextension as a particular form of overextension, I will first lay out a Wittgensteinian account of general conceptual overextension. Conceptual overextension is a philosophical phenomenon where a concept (or a cluster of concepts) is applied to cases beyond its reach, generating philosophical puzzlement and sometimes spurring conceptual revision. What often occurs in these cases is that a broader concept is identified with a narrower definition, and is subsequently used to understand a wide range of cases we associate with that concept. What arises are tensions which are seen as conceptual defects, which can lead the philosopher to think that their overall concept requires revision. In focusing on normative overextension, I am suggesting that even in cases where a concept is used which is fine-grained enough to capture the relevant descriptive features, normative overextension can arise. By laying out the idea of normative overextension and its implications, I will be developing the wider argument that a fixed normative role lacks the flexibility necessary to generate apt judgments.

ID: 122 / **Panel 1-2-H: 1**
Symposium paper
Topics: Ethics (general and applied), Metaethics, Moral Psychology, Phenomenology
Keywords: Artificial Intelligence, Moral Emotions, Emotional Practices, Healthcare, Emotional Attunement

AI mediation of moral emotions: How the introduction of artificial intelligence in the field of healthcare is mediating emotional practices

Eliana Bergamin

Moral emotions such as empathy, sympathy and compassion are key to healthcare practice. The relationship established between caretakers and caregivers in situations of vulnerability calls for an empathetic bond. However, recent implementations of data-focused research and methodologies may disturb traditional practices of empathy and compassionate care (Kerasidou, 2020). Techniques such as data driven diagnostics, clinical trial enrolment and AI risk, prognosis & data management are entering the healthcare field. By doing so, they are mediating the emotional practices between the different – and new – stakeholders involved and the implemented technology. In a world where Artificial Intelligence is permeating every field — including healthcare —, where human experience is mediated by it, is the emotional environment also being modified by this technology?

To answer this question, I build on the postphenomenological perspective of mediation theory to show how introducing AI-powered technologies in the healthcare field mediates the emotional practices of the different stakeholders involved. Drawing on insights from philosophy of emotions, specifically concerning the pivotal role of emotions for moral judgment, I aim at showing how implementing AI in healthcare can lead to a disruption of emotional practices. Using the notion of emotional attunement, I argue that changes in emotional practices determined by technological implementation can be detected when actors perceive a glitch – a friction – in the emotional surrounding they are used to experiencing.

Mediation theory affirms that technologies mediate the way human beings act in the world, which comprises its moral aspects (Verbeek, 2011). While serving as a lens to analyze how technologies have an influence on human practices, mediation theory focuses primarily on the perspective and experience of the individual. This can at times overlook the broader influence that technologies have on a broader societal and cultural level (Cressman, 2016). This research aims at covering the gap between technological moral mediation and the often-overlooked role of moral emotions in it. To do so, I introduce the concept of "emotional attunement". Building on the Heideggerian notion of attunement, and more specifically on its reinterpretation by Sareeta Amrute, I show how humans live and act in an attuned way to their emotional environment (Amrute, 2019). Subsequently, I merge this analysis with Amrute's characterization of 'glitchy attunements'. While Amrute defines this as a break in the technical system, I adapt this definition to become a break in the emotional environment – or emotional practice. I argue that these frictions, or glitches, can signal how the technology is modifying the way of doing healthcare. Through this analysis, I wish to highlight how this specific focus on the moral and emotional side of sociotechnical practices can be used as a hermeneutic tool to better understand AI implementation in healthcare.

Considering how fundamental the role of moral emotions is in the field of healthcare itself, I highlight how their role is still confined to the background of the debate on morality and technology, and how their potential reconsideration can provide insights on how AI implementation should be approached in near-future scenarios.

ID: 188 / Panel 5-3-F: 1
Individual paper
Topics: Social and political philosophy
Keywords: Legitimacy, Coercion, Political Philosophy

Neither Anarchism nor the Right to Rule

Justin Bernstein

Making, enforcing, and interpreting law often involves the use of physical force, threats to use physical force, or actions that are more indirectly backed by the threat to use physical force. Such actions constitute exercises of coercive power. But many of the ways in which government officials act would be impermissible if some other person or entity were to attempt to exercise coercive power in a similar fashion. The fact that government officials act in these ordinarily impermissible ways naturally leads to a fundamental question of political morality:

THE QUESTION OF GOVERNMENT COERCION: What (if anything) permits government officials, but not others, to exercise coercive power by making, interpreting, or enforcing law?

This paper articulates reasons for dissatisfaction with the two most prominent classes of answers to the Question of Government Coercion, and briefly suggests some alternative approaches to answering the question.

The first class of answers, those from philosophical anarchists, insists that government officials are only permitted to act in ways that ordinary citizens are not permitted to act. For instance, just as ordinary citizens are permitted to use coercive power to prevent physical assault, so are government officials. Yet this class of answers struggles to accommodate two widely held intuitions. First, many share the intuition that officials should enforce make, interpret, or enforce various laws that solve collective action problems— e.g., regulations concerning pollution or intellectual property rights—but ordinary citizens should not make or enforce such rules. Second, many think that officials should deliberate in very different ways from citizens; they should guide their deliberation by conventional rules rather than 'the moral law,' and they should strive to be impartial in contexts where citizens are not obligated to be impartial, or where citizens would do something morally wrong by acting impartially.

The second class of answers, those from the proponents of the 'right-to-rule framework,' contend that government officials enjoy special permissions to exercise coercive power in virtue of the state's right to rule. In states with the right to rule, government officials—and only government officials—are permitted to exercise this right. Yet I give reasons to think that a state possessing the right to rule is neither necessary nor sufficient for this special permission on the part of officials. The right to rule is not necessary for the special permission because political officials enjoy a special permission to make or enforce law—even when they act on behalf of states without the right to rule. The claim that the right to rule is sufficient for the special permissibility of particular actions encounters a dilemma: either the claim is manifestly implausible, given that government officials in states with the right to rule regularly perform impermissible actions, or attempts to avoid this implausible implication end up involving ad hoc modifications that deprive the right to rule of its explanatory power.

Before concluding, I briefly consider what it might look like to reject both philosophical anarchism and the right-to-rule framework when answering the Question of Government Coercion.

ID: 137 / **Panel 3-3-B: 1**
Individual paper
Topics: Social and political philosophy
Keywords: balancing, rationality, coherence, incommensurability, rights

The rationality of balancing

Laurin Berresheim

Decision-makers in law and politics often face situations in which different rights and interests are in conflict with each other. It is sometimes necessary to adopt a measure that interferes with one right in order to promote other rights or interests. In a judicial contexts, such a measure is commonly found to be justified if it is proportionate. To assess this, courts across the globe have established a proportionality test with three cumulative steps, which involves examining whether a measure is, first, suitable and, second, necessary to achieve the aim that it purports to achieve. If these conditions are fulfilled, the third step involves balancing the two rights or interests that are at stake.

The concept of balancing is however widely disputed. One of the main points of critiques, as voiced, for instance, by Jürgen Habermas, is that there are no rational standards for balancing competing rights and interests and the decision therefore is arbitrary. Related to this, various authors have criticized that balancing purports to be an objective method to compare principles, which in fact cannot be compared because they are incommensurable. It transfers an image borrowed from the physical world, where things can be measured by using a common metric and thus be made comparable, to the moral world, where there is no such common metric.

In my paper, I refute these arguments based on a conception of balancing that was developed by the legal philosopher Robert Alexy. I first highlight the rules of legal reasoning to which balancing is subject to challenge the argument that there are no rational standards. Second, I argue that a common measure is not necessary to compare competing principles, as long as the comparison is based on a common reference. This discussion makes evident recurrent conceptual confusions between rationality and certainty as well as between objectivity and certainty. I argue that balancing can be considered to be rational and objective despite not yielding a single, unequivocal answer for each case. Finally, the paper elaborates a concept of the legitimacy of law, which builds on several elements such as the rationality, the coherence and the impartiality of the application of law.

This paper is based on a draft chapter of my thesis project, which addresses the question of how to solve conflicts between the right to privacy and other competing rights and interests by means of balancing.

ID: 155 / Panel 2-2-E: 1
Individual paper
Topics: Social and political philosophy, Phenomenology
Keywords: Material Engagement Theory, Foucault, Mind, Body, Technology

Brain plasticity after Foucault: The political naivety of Material Engagement Theory

Víctor Betriu Yáñez

A blind person walks down a street without stumbling upon any obstacle thanks to the activity of sweeping the floor from side to side with a simple stick. The stick acts as the prosthesis which supplements the individual's lack of vision, thus informing her about her immediate surroundings. Some philosophers have invoked this situation to call into question the limits of perception, cognition, locomotion or self (cf. Polanyi, 1958; Merleau-Ponty, 1962; Bateson, 1973)—for it is difficult to tell where the blind person ends and where the world begins. Lambros Malafouris (2013) has taken this example as the starting point of his Material Engagement Theory (MET): a (re)conceptualisation of the relationship between cognition and material culture from a non-anthropocentric perspective, which entails theorising from an active abandonment of boundaries and centres (Malafouris, 2013, p. 36). This is because substantial approaches risk confounding the artificiality of analytic boundaries (e.g., mind/matter, cognition/world, brain/culture) with ontological reality (Malafouris, 2013, p. 208), which may result in a category mistake that generally bestows privilege upon the human in relation to the world due to the concealment of an epistemological issue—the difficulty of knowing for certain whether an entity has the property we say it has. Well informed by contemporary theories of the mind and the scientific notion of neuroplasticity, Malafouris grounds his MET on a relational ontology, which characterises entities and phenomena through the analysis of their mutual interaction instead of imposing inherent properties that define their essences, thus offering a view of the mind that is coextensive with the totality of beings involved in material engagement—the synergetic interaction among brains, bodies and things which is the condition of possibility of the mind (Malafouris, 2013, p. 17).

In this paper, I raise an objection to MET on the grounds that, albeit it is highly alluring for giving compelling and non-dualist explanations of the dynamics in brains-bodies-things interactions, it remains politically naïve for neglecting cases in which such engagements occur rather unconsciously and are indissociable from disciplinary power (Foucault, 1995). Through the rifle and the Panopticon examples that Foucault (1995) analyses in Discipline and Punish, and invoking other contemporary examples (e.g., Benjamin, 2002; Sennett, 1994), I show that there are situations of material engagement where the elements participating in cognition have different degrees of agency, so that some might be subdued to others. Thus, a paradox thrives within MET: (1) its political naivety nourishes the illusion that humans can have control over their surroundings, while (2) it risks, precisely for the illusory character of (1), turning its alleged non-anthropocentrism into submissive human existencethus making resistance to power impossible. This paradox demonstrates that, insofar as there are instances of oppressive synergetic interactions, then there is something like an oppressed mind, to which MET remains blind. In pursuit of an amendment, I conclude that MET could include a political dimension if it pays heed to Malabou's (2008) political reading of plasticity.

ID: 192 / **Poster session 2: 3**
Poster
Topics: Epistemology, Aesthetics and philosophy of art
Keywords: ideal landscape, spatial montage, artistic practice and research, moving images, epistemological inquiry

The ideal landscape – A visual narrative.

Ilaria Biotti

The research project "The Ideal Landscape - A Visual Narrative" explores the ideal landscape through and in moving images. Through and in moving images, the ideal landscape is thematised as a living contemporary imagery connoted by time and movement. In its pictorial dimension, this genre reflects a way of looking at landscape that emerged in Europe in the 17th century, influenced by classic canons of beauty and form, as well as a growing interest in the natural world. The project aims to critically examine the process of imaging a physical landscape as ideal in the contemporary and to deepen the understanding of the images of thought that lead to imagining such idealisation today.

Spatial montage is adopted as a method to explore a physical landscape imagined as an ideal landscape in the here and now. The term "spatial montage" refers to the process of combining spatial elements to visualise a composite space. The research project takes a critical approach to spatial montage, examining the epistemic relationships, power dynamics and the resulting implications of this process. Through this approach, it seeks to unravel the contemporary assumptions about the ideal inherent in today's physical landscape. It also aims to challenge these assumptions and their socio-political implications.

The research project addresses several questions, including the extent to which a critical approach to spatial montage can contribute to untangling the epistemological dimension of the imaging processes within the idealisation of the physical landscape today. It also explores the extent to which spatial montage as an artistic method can undo the relationship between image and imagination in physical landscape. Finally, through practice, it questions to what extent an artistic inquiry can contribute to unravelling contemporary assumptions about the ideal in relationship to landscape.

Despite the investigation happening through and in moving images, the outcome is not envisioned to be a film, but rather the spatial visualisation of moving images. By foreseeing this type of outcome, this research project wants to offer an open and contingent approach to understanding the complex relationship between humans and the environment. It wants to challenge contemporary assumptions about landscape and to delve into a deeper understanding of the ways in which we imagine and interact with the physical world around us, refracting within landscape the narratives and ideas of contemporaneity.

Overall, the research project "The Ideal Landscape – A Visual Narrative" draws on a visual practice driven approach to explore contemporary physical landscape. The attempt is to deepen an understanding of the complex relationship between humans and their agency on the environment. Future potential implications or applications of the research project could contribute to inspire a more reflected approach to landscape and the natural world. In view of how the focus of the project is on both artistic and critical inquiry, its developments could contribute to facilitate a dialogue between different disciplines, fostering a more holistic understanding of landscape.

ID: 181 / Panel 3-2-E: 1
Individual paper
Topics: Logic, Philosophy of Science, Philosophy of Language
Keywords: Frege, Logicism, Definition, Reference, Scientific concepts

Definition and Content in Frege's Logic

Rachel Boddy, Robert May

In Grundgesetze, Frege attempted to demonstrate that his logical language, the Begriffsschrift, is a fully referential language. Although Frege's proof of referentiality fails (Russell's Paradox), Frege's reasons for requiring referentiality remain of interest. We argue that for Frege the significance of the proof of referentiality resides in the role it plays in the justification of definitions. Thinking through this perspective is our purpose in this paper; doing so leads to a view of Frege's account on which definitions can isolate not only logical concepts, but can also, as a central purpose, afford the recognition of scientific concepts.

We begin by showing the textual evidence that Frege intended the Proof of Referentiality to be part of the groundwork for the introduction of definitions, before turning to the concern of how Frege's insistence that the relation of definiendum and definiens be conservative can allow that the definition of the cardinal number of a concept, definition (Z) of Grundgesetze, be a logical definition, yet introduce names for numbers. In the context of the logicist project, definitions are meant to afford the recognition of the core mathematical concepts, most notably the concept of number, and by doing so reveal logical propositions as expressing mathematical content. Here the key question is how definitions can be fruitful, and accordingly scientifically illuminating, if they are conservative over the logical language. In answering this question, a distinction is drawn between analytic definitions, justified by how they illuminate concepts, and proper definitions, justified by their utilization in proof.

That definitions are conceptual raises an apparent problem: In Frege's system, this property cannot be directly captured, not if definitions are to be logical. We show how Frege addresses this issue, and discuss why this requires that definitions introduce novel terms within a fully referential milieu. In particular, Frege addresses this issue by proposing that definitions can introduce referential terms for objects that represent concepts, specifically, to value-ranges of concepts. Formulated in this manner, the core mathematical definition, the definition of the number of a concept, establishes reference to numbers by introducing a referential term (definiendum) whose referent is the value-range of the concept of equinumerosity (definiens). Critical to this relation of concepts and values-ranges is definition (A) of Grundgesetze, which is intended to capture the predicativity of concepts (falling under) in terms of a relationship between objects and value-ranges (membership). Lastly, the discussion turns to Russell's Paradox. What the paradox narrowly shows is that the proof of referentiality fails. But its broader implication is that the canons of definition collapse, as the logical transition from a concept to its value-range, necessary for the specification of definitions, fails. Thus, in the context of Frege's logicist program, the lesson of Russell's Paradox is that it undermines the definition of number, and hence the scientific content of Frege's logicist project. We conclude with remarks on the status of Frege's logicist project as a semantic, as opposed to epistemic, project, and its implications for Frege's view of scientific knowledge.

ID: 141 / **Panel 2-2-B: 1**
Individual paper
Topics: Ethics (general and applied), Metaethics, Moral Psychology
Keywords: taking responsibility, responsibilities, obligation, prospective responsibility, forward-looking

Taking on responsibility: wrongly, badly, and unfairly

Dominik Boll

The phenomenon of taking forward-looking responsibility is familiar enough. I might, for instance, volunteer to organise a departmental birthday calendar. We do these sorts of things constantly; they are often optional, supererogatory, and praiseworthy; large parts of life depend on such structuring; and we fulfil ensuing responsibilities without obscurity. Yet, philosophical puzzles remain. It is frequently unclear who should take responsibility; taken responsibilities regularly, but not always, turn into obligations entailing blameworthiness; taking and rescinding responsibility often leaves important tasks unassigned; and poor execution can have disastrous consequences. These problems demonstrate the attention taking responsibility deserves, albeit currently lacking in the literature.

My analysis starts by considering three questions. First, I investigate the nature and content of responsibilities as opposed to tasks, duties, and obligations (Björnsson and Brülde 2017). Responsibilities may be distinct because they pertain to specific states of affairs (Goodin 1986), leave room for discretionary judgment (Feinberg 1988), solve cooperative problems (Alfano 2021), or involve (self-)assigned ownership of ends (Calhoun 2019). Second, I focus on different ways to acquire responsibilities, both implicitly and explicitly as well as formal and informal. I inquire into felicity conditions specific to taking responsibility, and the consequences of (not) succeeding (Austin 1962). Third, I analyse appropriate agents. Whether someone can take responsibility depends on the content and acquisition of responsibilities, but also on the agent's capacities or relation to the respective end. These further factors are complicated by the tension between the societal need for assignment (Richardson 1999) and the elective structure it has for individuals (Calhoun 2020).

I then draw on this analysis and recent insights from Calhoun (2019) to move towards an ethics of taking forward-looking responsibility, concentrating on three cases. First, an agent can take responsibility wrongly because they are ineligible or do so inappropriately. For instance, a malicious assaulter might not be in the right position to take responsibility for the victim's recovery. Second, an agent successfully acquiring responsibility might execute or rescind it badly because they discharge the involved tasks unsatisfactorily. For example, my organisation of the birthday calendar is optional, but stopping immediately might provoke justified frustration. The distinction between these two mistakes is often unclear and underappreciated, for instance in public discussions on collective action. If I take responsibility for improving my university's climate impact, but then engage in leisurely flying, others might have a grievance not only because I fail to promote the end, but also because I harm the cause promoting it. Depending on the agent and specific failure, the case might be described as either taking responsibility wrongly or badly. Third and additionally, agents can take responsibility unfairly because of their means or the resulting distribution (Miller 2001). For instance, I might exploit social expectations to fulfil certain optional tasks, thereby taking them away from colleagues. Thus, taking responsibility can be utilised in disrespectful and manipulative ways. I end by describing the specific badness involved in these cases and point to how this analysis yields insights into promising ways to improve social practices.

ID: 174
Full symposium
Topics: Social and political philosophy
Keywords: taxation, inheritance, tax avoidance, passive income

Philosophy of Taxation

Chair(s): **Huub Brouwer** (Tilburg University, Netherlands, The)

Taxation is an important way in which governments can implement policy measures to approximate philosophical ideals about the just distribution of income and wealth. Nevertheless, the topic of taxation has, for a long time, been ignored by philosophers, who mostly focused on how governments should spend the tax income they collect. This is a shame, because as the philosopher Leif Wenar puts it, "[i]n the absence of theories of distributive justice that have some traction on the public conscience, the battles to set tax policy will always be won by appeals to interest and efficiency" (2005, p. 286). Political philosophers have only recently started to work on philosophical justifications for various forms of taxation (Gaisbauer, Schweiger, & Sedmak 2015; Halliday 2018, 2020; Murphy & Nagel 2001, 2002; O'Neill & Orr 2018; Pedersen 2018, 2020, 2021). The aim of this symposium is to contribute to philosophers systematically reflecting on systems of taxation.

Presentations of the Symposium

Limiting lifetime inheritance and gifts

Ingrid Robeyns

Ethics Institute, Utrecht University

With rising wealth inequalities and diminished social mobility, bequest and inheritance taxation is back on the academic and political agenda. Welfare economists have typically proposed particular institutional schemes, but without a normatively sound justification. Applied ethicists and political philosophers typically prioritise in their analyses one value that is to be protected and base the justification of an institutional design on that value, or else provide pro tanto analyses. If one of the tasks of normative political philosophy is to provide guidence on which institutional proposal takes into account all values that citizens can reasonably endorse, then we need to move to synthetic normative political philosophy to analyse and justify any institutional proposals.

In this paper, I build on these two literatures to provide a defense of a lifetime limit to how much a person can inherit or receive as gifts. Any inheritances or gifts that a person would receive above that limit should be taxed at a rate of 100%. Institutionally, this is combined with a unconditional citizen's stake for young adults financed from those tax revenues. This specific institutional proposal follows from a discussion of how the core values at stake in the design of taxation on inheritances and gifts should be brought into equilibrium. These values are well-being, freedom, desert, equality of opportunity, family values, and political equality. I then consider and respond to three objections which are all in one form or another question about the feasibility or incentive-compatability of this proposal: that there is no democratic support for inheritance taxation in general, that the proposal does not properly distinguish between unproductive wealth and capital, and that it is unfeasible because of capital flight.

The fairness of tax avoidance

Bruno Verbeek

Centre for Political Philosophy, Leiden University

Tax avoidance, the legally permitted reduction of fiscal burdens, is a hot topic in circles of activists (e.g., the Tax Justice Network) and policy makers (e.g., the BEPS initiative of the OECD). The discussion is typically limited to multinational companies and their aggressive tax planning structures which often border on what is legally permissible. However, in this context it is crucial to distinguish tax avoidance from tax evasion on the one hand, and tax mitigation on the other. In this talk, I will try to give a characterization of tax avoidance, evasion, and mitigation, and try to argue when and why avoidance becomes problematic from a moral point of view. I will conclude that taxpayers, whether natural or legal persons, have fiscal duties that go beyond what the law requires.

Letting your money work for you

Huub Brouwer
TiLPS, Tilburg University

In a recent interview on national television, the director of the Dutch central bank, Klaas Knot, called the Netherlands a "rentier economy" (Buitenhof 2022). In the Netherlands, he went on to explain, labor income is taxed at a relatively high rate, whereas wealth income is taxed at a relatively low rate. He called for a revision of the Dutch tax system to change this, because the tax system should incentivize people actively earning income rather than passively receiving it.This raises the question: Should passive income be taxed at a different rate than active income? In this paper, I will investigate the distinction between active and passive income and ask whether it can be justified to tax these two types of income differentially.

This paper proceeds in three steps. First, I will point out that there is a long historical tradition of arguing against the receipt of passive income. These debates center around the notions of 'usury' (Aristotle, the Bible) and 'exploitation' (Marx). These historical condemnations of usury and exploitation stand in stark contrast with contemporary views on which passive income, in the form of venture capitalism or entrepreneurship, can be deserved on the basis of the (financial) risks involved in these activities and the enabling of contributions that would otherwise not have been possible (see Shapiro 2018; Mankiw 2013; Kershnar 2005). Second, I argue that the distinction between active and passive income is not as clear as it may seem. And certainly it does not seem like interest, rent, royalties, and dividends neatly fit into the active versus passive distinction. Some investors, for instance, receive interest for lending out their money to start-ups while being actively involved in their development. Other people just receive interest because they have put their money on a savings account and the bank loans it out / invests it for them. I distinguish three different types of passive income (that I call 'passive but active', 'active turning passive', and 'purely passive'). Third, I evaluate the case for taxing passive income at a lower rate than active income, using the notions of desert, equality, and freedom. I conclude that the desert-based and equality-based cases for the differential treatment of passive income fails. If anything, they tell in favour of taxing passive income at a higher rate than active income. Lower taxation of passive income can only be justified by appealing to freedom-based arguments - but these tell against any taxation beyond what is required for a nightwatchman state more generally.

The project of the paper is important, because many countries (not only the Netherlands) tax passive income at a lower rate than active income. Return on capital mainly consists of passive income. As Thomas Piketty has famously argued, the rate of return on capital is higher than economic growth ('r > g'), and, consequently, wealth inequality in many developed countries is rising (2014). As a result, it is crucially important to debate whether taxing passive income at a lower rate can be justified.

ID: 219 / **Panel 2-1-F: 1**
Symposium paper
Topics: Social and political philosophy
Keywords: taxation, inheritance, tax avoidance, passive income

Letting your money work for you

Huub Brouwer
TiLPS, Tilburg University

In a recent interview on national television, the director of the Dutch central bank, Klaas Knot, called the Netherlands a "rentier economy" (Buitenhof 2022). In the Netherlands, he went on to explain, labor income is taxed at a relatively high rate, whereas wealth income is taxed at a relatively low rate. He called for a revision of the Dutch tax system to change this, because the tax system should incentivize people actively earning income rather than passively receiving it.This raises the question: Should passive income be taxed at a different rate than active income? In this paper, I will investigate the distinction between active and passive income and ask whether it can be justified to tax these two types of income differentially.

This paper proceeds in three steps. First, I will point out that there is a long historical tradition of arguing against the receipt of passive income. These debates center around the notions of 'usury' (Aristotle, the Bible) and 'exploitation' (Marx). These historical condemnations of usury and exploitation stand in stark contrast with contemporary views on which passive income, in the form of venture capitalism or entrepreneurship, can be deserved on the basis of the (financial) risks involved in these activities and the enabling of contributions that would otherwise not have been possible (see Shapiro 2018; Mankiw 2013; Kershnar 2005). Second, I argue that the distinction between active and passive income is not as clear as it may seem. And certainly it does not seem like interest, rent, royalties, and dividends neatly fit into the active versus passive distinction. Some investors, for instance, receive interest for lending out their money to start-ups while being actively involved in their development. Other people just receive interest because they have put their money on a savings account and the bank loans it out / invests it for them. I distinguish three different types of passive income (that I call 'passive but active', 'active turning passive', and 'purely passive'). Third, I evaluate the case for taxing passive income at a lower rate than active income, using the notions of desert, equality, and freedom. I conclude that the desert-based and equality-based cases for the differential treatment of passive income fails. If anything, they tell in favour of taxing passive income at a higher rate than active income. Lower taxation of passive income can only be justified by appealing to freedom-based arguments - but these tell against any taxation beyond what is required for a nightwatchman state more generally.

The project of the paper is important, because many countries (not only the Netherlands) tax passive income at a lower rate than active income. Return on capital mainly consists of passive income. As Thomas Piketty has famously argued, the rate of return on capital is higher than economic growth ('r > g'), and, consequently, wealth inequality in many developed countries is rising (2014). As a result, it is crucially important to debate whether taxing passive income at a lower rate can be justified.

ID: 217 / Panel 2-2-F: 1
Symposium paper
Topics: Social and political philosophy
Keywords: taxation, inheritance, tax avoidance, passive income

Limiting lifetime inheritance and gifts

Ingrid Robeyns

Ethics Institute, Utrecht University

With rising wealth inequalities and diminished social mobility, bequest and inheritance taxation is back on the academic and political agenda. Welfare economists have typically proposed particular institutional schemes, but without a normatively sound justification. Applied ethicists and political philosophers typically prioritise in their analyses one value that is to be protected and base the justification of an institutional design on that value, or else provide pro tanto analyses. If one of the tasks of normative political philosophy is to provide guidance on which institutional proposal takes into account all values that citizens can reasonably endorse, then we need to move to synthetic normative political philosophy to analyse and justify any institutional proposals.

In this paper, I build on these two literatures to provide a defense of a lifetime limit to how much a person can inherit or receive as gifts. Any inheritances or gifts that a person would receive above that limit should be taxed at a rate of 100%. Institutionally, this is combined with a unconditional citizen's stake for young adults financed from those tax revenues. This specific institutional proposal follows from a discussion of how the core values at stake in the design of taxation on inheritances and gifts should be brought into equilibrium. These values are well-being, freedom, desert, equality of opportunity, family values, and political equality. I then consider and respond to three objections which are all in one form or another question about the feasibility or incentive-compatability of this proposal: that there is no democratic support for inheritance taxation in general, that the proposal does not properly distinguish between unproductive wealth and capital, and that it is unfeasible because of capital flight.

ID: 195 / Panel 3-2-G: 1
Individual paper
Topics: Metaphysics, Social and political philosophy
Keywords: Social Properties, Constraints on Behaviour, Hyperplans, Plan-Based Metaphysics, Social Ontology

A theory of social properties

Thomas Nicolaas Pieter Albert Brouwer

Being a philosopher; being guilty of tax fraud; being popular; being an artwork; being in fashion; being against the law – these are examples of social properties of persons, objects and actions. Understanding what social properties are, how they come about, and how they matter to us is an important task for social metaphysics; in this talk, I present a theory of them.

It is a familiar thought in social ontology that there is some special metaphysical connection between the nature of social properties and constraints on behaviour. That thought can be variously articulated, as Jenkins (2020) usefully surveys; here I articulate it using techniques from plan-based semantics, an approach which has travelled from metaethics into the philosophy of language, but hasn't yet made inroads into social ontology.

Plan-based semantics was introduced by Gibbard (2003) as a way of modelling the content of normative thought and language; it represents the normative constraints that an agent takes themselves to be under using sets of hyperplans – sets of maximally comprehensive and specific directives which can represent open practical possibilities much like sets of possible worlds can, for various theoretical purposes, represent alethic or epistemic possibilities. Hyperplans are a flexible theoretical tool (Yalcin 2019) and they've since been repurposed (Yalcin 2012) to model the practical dimension of a Stalnaker-style (1978) conversational common ground; MacFarlane (2016) has extended this approach to develop a context-sensitive semantics of gradable adjectives.

Building on this, I repurpose hyperplans again, in two ways. I propose, first, that hyperplan sets can be used to represent practical contents associated not with individuals or conversational contexts, but with entire communities. And I propose that hyperplan sets can be deployed beyond semantics for more directly metaphysical purposes, specifically those of social metaphysics.

Communities exhibit a variety of social mechanisms by which, for good or ill, they constrain and/or direct the behaviour of their members; among the more studied of these are conventions (Lewis 1969, and many others since) and social norms (Bicchieri 2006, Brennan et al. 2013). For a given community, we can represent – I'll argue – the totality of the behavioural constraints and directives so generated as a hyperplan set: call it the community plan.

I'll describe how we can, in various systematic ways, abstract more specific hyperplan constructs from such a community plan which I call treatments (cf. Williams 2016). I'll argue that we can metaphysically identify the fact that an object instantiates a certain social property with the fact that the community plan accords the object a certain treatment, and that we can extract grounding conditions for the instantiation of social properties from the trigger conditions that a community plan imposes for various treatments.

After articulating the proposal and confronting some devils lurking in its details, I discuss what its explanatory scope and strengths are, how it fits into larger research programmes of social ontology, and to what extent it can recapture elements of existing approaches in social ontology (e.g. Searle 1995, Guala & Hindriks 2015, Epstein 2015).

ID: 110 / Panel 4-2-H: 1
Individual paper
Topics: Ethics (general and applied), Metaethics, Moral Psychology, History of philosophy (from ancient to modern), Philosophy of Religion
Keywords: assisted dying, Cryonic Wager, cryothanasia, Pascal, suicide

Dying to live: transhumanism, cryonics, and euthanasia

Adam Buben

At first glance, it might seem that transhumanists would have little interest in receiving assistance in ending their lives. Such assistance is typically regarded as a last resort for relieving otherwise inescapable and intolerable suffering, so why would people characterized by tremendous optimism about what near-future technology might make possible—which includes hope for radical life extension—focus their attention on ways to end it all? This is not to suggest that transhumanists are unconcerned about suffering and all the ways it shows up in the world; in fact, concern about the misery ordinary humans have experienced throughout history is among the primary motivators for much of the transhumanist agenda. What sets them apart, however, is that most transhumanists think just-over-the-horizon technological developments will put a stop to much, if not all, of our misery without needing to "put us out of our misery." Of course, problems arise when that horizon turns out to be a bit farther away than they had initially hoped and predicted, and it is in response to these problems that a connection between transhumanism and various forms of assistance in dying comes into view.

For those of us unlikely to survive long enough to see the promised technological wonders, transhumanists think there might be a way to buy some extra time: cryonic preservation. As in the case of the procurement of functional organs for transplant, however, the preservation process works best if it is begun before serious deterioration of the tissue takes place, and this fact makes controlling the time and manner of death very attractive for transhumanists interested in cryonics. Building on a fairly small body of relevant existing literature, this paper argues that—at least under certain circumstances—these transhumanists might be justified in seeking assistance in dying (and perhaps even dying via the cryonics procedures themselves) in order to increase the quality of preservation. Interestingly, this argument is built around an analogy with Pascal's Wager.

His Wager suggests that where the potential benefits of making a finite sacrifice are "an infinite life of infinite happiness," it really does not matter how likely getting those benefits is; so long as there is even a remote possibility, making the sacrifice is sensible. Given that the potential (though unlikely to materialize) benefits of cryonic preservation for the transhumanist include indefinite life extension, one might suggest that the merely finite sacrifice/wager of time and money involved in pre-mortem preservation is similarly sensible. Although the analogy with Pascal's Wager is imperfect, the potential benefit of indefinite life extension means that so-called "cryothanasia" would compare quite favorably, even if there is a substantial divergence in probability of success, to other experimental last resorts (already allowed in many jurisdictions) that might only buy someone a few more years.

ID: 183 / **Poster session 1: 3**
Poster
Topics: Philosophy of Science, None of the topics above (please include your topic under 'keywords'), Philosophy of Language
Keywords: Philosophy of technology, explainable AI, tacit knowledge, large language models, understanding

Tacit knowledge for explaining how large language models work

Céline Evianne Budding

Large language models (LLMs), such as GPT-3 (Brown et al., 2020) and ChatGPT (OpenAI, 2022) are exhibiting increasingly impressive and arguably human-like behavior: they can translate and generate text and seem to engage in generally realistic dialogue with human users. This behavior has led to questions regarding the cognitive and linguistic capacities of these models. Language has long been associated with intelligence, think for example of the Turing test (Turing, 1950). Since we now have artificial systems that generate language, what does it mean for these systems to 'know' language and to what extent should we attribute any cognitive and linguistic capacities like knowledge and understanding to them?

In this contribution, I aim to defend two claims. Firstly, I argue that it is insufficient to merely consider behavior when determining whether LLMs should be attributed any cognitive or linguistic capacities. Instead, we should develop explanations that show how a network actually works by looking at the internal processing of the model (Block, 1981; Zednik, 2021), i.e. how it mediates transitions from inputs to outputs. This could show what regularities the network picks up from the input data, what representations are learned, and what structures underlie generalization to new inputs. These kinds of explanations seem necessary to evaluate whether LLMs have an internal structure that supports robust linguistic capacities (See e.g. Piantadosi & Hill, 2022) or are mere 'stochastic parrots' that cannot capture meaning (Bender et al., 2021; Bender & Koller, 2020).

Due to the opacity and scale of these networks, however, it is unclear how these explanations should be developed. Drawing on work from Davies (1987, 1990) and a recent paper by Lam (2022), my second claim is that tacit knowledge could be a promising theoretical posit for developing how-explanations of LLMs. In the context of neural networks, tacit knowledge is defined as implicitly represented rules that are used by the network to guide its behavior. To be ascribed tacit knowledge, Davies proposes that networks should, amongst others, meet the constraint of causal systematicity, which requires that there is a pattern of causal common factors that mediates transitions from similar inputs to similar outputs. The problem is, however, that this is not directly applicable to contemporary LLMs due to their distributed nature. To show that tacit knowledge might nevertheless be helpful for developing how-explanations, I critically reflect on Davies' conception of tacit knowledge and argue that it could in fact be used to explain the behavior of LLMs to some extent.

In summary, this contribution has the following aims: a) to introduce and nuance the current discussions regarding the potential cognitive and linguistic capacities of LLMs, b) to argue that attribution of such capacities requires a suitable internal structure, c) to propose that tacit knowledge is one way to conceptualize such a structure, and d) to critically evaluate and adapt Davies' account of tacit knowledge to make it suitable for contemporary LLMs.

ID: 177 / Panel 3-3-F: 1
Individual paper
Topics: Social and political philosophy, Ethics (general and applied), Metaethics, Moral Psychology
Keywords: corporate responsibility, responsibility for structural injustice, Iris Marion Young, business ethics

Corporate political responsibility for structural injustice: Taking Young seriously in business ethics

Barbara Bziuk

From human trafficking, exploitation and poverty to obesity and other food-related health problems, business corporations have been linked to and deemed responsible for many global problems, based on Iris Marion Young's account of structural injustice and responsibility. In the business ethics literature, Young's model of responsibility attribution gave rise to many arguments in favor of ascribing far-reaching, and often positive, obligations to business corporations, such as lobbying the government to remedy injustice or engaging with others to create public institutions, proactively protecting human rights, and many others. And yet, when corporations act on their political responsibilities in real life, they face serious criticism from civil society. In this paper, I argue that business ethics literature on corporate political responsibility for structural injustice cannot predict and explain that critique because it is not sufficiently sensitive to power. This limitation in the literature has problematic consequences for both, the adequacy of the normative analysis of structural injustice as well as of the practical implications of such analysis and the application of the social connection model to business corporations, or so I argue. Thus, the aim of the paper is twofold. First, I argue that the business ethics literature on corporate political responsibility is not sensitive to the power imbalance between corporations and other actors because, although it starts from Young's account of structural injustice, it does not fully embrace it. For Young, power and relations between various social positions, such as between private companies and their workers, are a critical element of identifying and defining structural injustice. I will demonstrate that the literature on corporate political responsibility misses this point. Overlooking power relations in defining structural injustice and responsibility for it leads to overlooking the relational power of business corporations. Thus, second, I argue that the mischaracterization of structural injustice in the business ethics literature has a further problematic consequence: it cannot be action-guiding. If the specific power dynamics are not included in the analysis of injustice, the practical conclusions of such an analysis cannot be, normatively and empirically, accurate for the world in which these power dynamics are at play, or so I argue. Thus, this paper contributes to the business ethics literature on corporate responsibility for structural injustice by incorporating Young's political-philosophical work, which has so far been neglected in that literature, and showing its consequences for thinking about corporate responsibility. The paper identifies a problem in the literature which is that structural injustice is not characterized sufficiently in terms of power and points to a challenge that arises when such characterization is provided. Thereby, the paper invites scholars concerned with corporate responsibility for structural injustice to consider different ways in which structural injustice to which business corporations are connected could be alleviated. More broadly, this paper fits into and brings forward the project of incorporating insights from political philosophy into business ethics.

ID: 166 / Panel 4-1-D: 1
Individual paper
Topics: Ethics (general and applied), Metaethics, Moral Psychology
Keywords: Immanuel Kant, F. W. J. Schelling, propensity to evil, possibility of evil

Kant and Schelling on the propensity to evil. Interpreting the Freedom Essay as a revision of Kant's Religion

Sebastian Cabezas

F. W. J. Schelling's Freedom Essay (1809), considered by many his main work, is usually seen in the tradition of Kant's writing Religion Within the Boundaries of Mere Reason (1793). At first sight, Schelling's philosophical elaborations in the Freedom Essay indeed take up important terminology from Kant's text and contain references to it. Specifically, this influence is mostly notorious in Schelling's doctrine of an intelligible deed and the thereby connected claim of a propensity to evil in human beings. Nonetheless, this seems to hide the fact that precisely at those places Schelling's philosophical position can be seen as diverging from the Kantian one in crucial points. In this paper, I shall advance the thesis that Schelling's position in the Freedom Essay concerning the aforementioned issues, in spite of the undeniable terminological proximity to Kant's Religion, is to be read rather as an alternative project to the one proposed by Kant.

Particularly, I will focus on the relevant topics of the Freedom Essay in which Kant's influence seems to be most apparent, namely the nature of the propensity to evil (Hang zum Bösen) in human beings. To support my claim, I shall develop two main observations. First, it is to be noted that even though the concept of a propensity to evil is in fact taken from Kant's Religion, Schelling modifies the function Kant assigns to it in such a way that he abandons a crucial aspect of Kant's doctrine. While for Kant the propensity to evil assumes the role of accounting for both the possibility and actuality of moral evil, Schelling takes it to account only for the latter, thereby dropping the claim that it also grounds the possibility of moral evil. Second, Kant's position implies an asymmetry regarding the possibility of moral good and moral evil: whereas the possibility of moral good is attributed to human nature itself, the possibility of moral evil is grounded by an act of freedom. This position, too, is abandoned by Schelling's account, who places both in the very nature of human beings.

I will conclude with an attempt to explain why Schelling is making those changes to Kant's doctrine by pointing out some difficulties implied by Kant's position which might have led Schelling to his revised version of it. These difficulties concern the tension arising between Kant's claims about the propensity to evil and his commitment to the possibility of moral betterment.

ID: 184 / **Panel 4-2-D: 1**
Individual paper
Topics: Ethics (general and applied), Metaethics, Moral Psychology, History of philosophy (from ancient to modern)
Keywords: Kant, moral philosophy, constitutivism, metaethics

Kant's constitutivism and the shmagency objection

Vinicius Carvalho

Many interpreters have argued for constitutivist readings of Kant's moral philosophy, yet there has been no account so far of whether David Enoch's (2006) shmagency objection poses a threat to Kant's alleged constitutivism. This paper addresses precisely this question. I argue that on current constitutivist readings Kant's theory is worryingly vulnerable to the objection. For instance, Formosa (2013) argues that for Kant the categorical imperative is binding because it is the constitutive principle of practical rationality. Reath (2018) submits to a similar claim, arguing that the normativity of the categorical imperative rests on the fact that it is the constitutive principle of rational volition. We might ask, however, why the fact that the categorical imperative is constitutive of practical rationality is supposed to have any normative significance, let alone an overriding significance, for imperfect rational beings. We can imagine a character akin to Enoch's shmagent raising the question: "why behave like a perfectly rational agent would? I'm fine acting as an imperfectly rational agent, giving precedence to whatever brings happiness instead of satisfying the demands of practical rationality". To insist that the moral law is the constitutive principle of one's practical rationality will not make a difference unless there is a reason to take it as normatively relevant.

I submit that Kant's argument for the normativity of the moral law does precisely that by relying on an intrinsically normative premise: rational beings are ends-in-themselves. In the Groundwork, Kant argues that a practical law must be grounded on something that is of unconditional worth in order to command categorically. Ultimately, he claims that the only thing capable of grounding the categorical imperative is rational nature itself. "Rational beings", he remarks, "are called persons because their nature already marks them out as an end in itself, that is, as something that may not be used merely as a means, and hence so far limits all choice (and is an object of respect)" (GMS 4: 428). Essentially, I argue that the moral law is normative for rational beings because it is the practical principle that is consistent with their standing as ends-in-themselves, a standing which they take to be normative.

On the one hand, that Kant does not ground the normativity of the moral law in its being a constitutive principle might make his version of constitutivism seem less attractive because less ambitious. On the other hand, that makes it immune to an objection that poses a threat to all other versions. Whether it is preferable depends on how well one is prepared to accept the claim that our nature as free and autonomous beings is the last step down the normative ladder.

Would Carnap have tolerated modern metaphysics?

Wouter Adriaan Cohen, Benjamin Marschall

Rudolf Carnap did not hide his dislike for metaphysics: he notoriously called metaphysicians 'musicians without musical ability'. The anti-metaphysical writings of Carnap and other logical positivists were influential and, for several decades, metaphysics was frowned upon in the leading circles of analytic philosophy. But times have changed. Today, metaphysics is considered one of the core areas of theoretical philosophy and self-consciously metaphysical papers fill the pages of the most highly respected journals.

What would Carnap have thought of this development? Despite his uncompromising opposition to metaphysics in his own time, the answer is not straightforward (the question is also addressed by Huw Price (2009) and Vera Flocke (forthcoming)). Metaphysics, after all, is done in many ways. What Carnap initially criticised in the 1920s and 1930s was the philosophy of German obscurantists like Martin Heidegger, who opposed formal logic and a scientific worldview. In contrast, many contemporary metaphysicians embrace formal methods and consider their methodology to be continuous with that of science. There is thus room for a conciliatory reading according to which the metaphysics of today is not threatened by Carnap's critique of the metaphysics of his own time.

I will develop a robustly anti-metaphysical interpretation of Carnap according to which he would have been critical of key parts of contemporary metaphysics. I cannot, of course, cover all of Carnap's many and sometimes subtle anti-metaphysical arguments here. My aim is rather to contrast his early meaning-theoretic critique—which aims to show that metaphysics is meaningless—with his later value-based critique—which aims to show that even when metaphysics is meaningful, it is, in many cases, not a worthwhile enterprise, and then to emphasise the latter. The value-based critique is not as well known as, yet arguably more compelling than, the meaning-theoretic strategy. It is also forceful against parts of modern metaphysics in a way that the meaning-theoretic critique is not.

The structure of the presentation is as follows. I first very briefly introduce Carnap's early meaning-theoretic critique of metaphysics, and note that it does not seem applicable to many modern metaphysical debates. I then draw attention to his Principle of Tolerance, according to which we must assess what Carnap calls 'linguistic frameworks' only by their pragmatic value. I note that this principle, which Carnap accepted only in 1932, is in tension with his early meaning-theoretic critique of metaphysics. Then, in the second half of the presentation, I argue that the Principle of Tolerance allows Carnap to develop a critique of metaphysics from a new angle. In particular, I argue that this new value-based critique of metaphysics applies to several contemporary metaphysical debates, namely those motivated by what I will call 'ontological anxiety'. These are debates that essentially revolve around reducing ontological commitments, such as fictionalism and nominalism.

References

Flocke, Vera forthcoming. "Carnap Is Not Against Metaphysics," in Adam Tuboly and Alan Richardson, eds., Interpreting Carnap: Critical Essays.

Price, Huw 2009. "Metaphysics After Carnap: The Ghost Who Walks?" in David Chalmers, David Manley and Ryan Wasserman, eds., Metametaphysics.

ID: 112 / Panel 3-1-C: 1
Individual paper
Topics: Ethics (general and applied), Metaethics, Moral Psychology, Philosophy of Language
Keywords: Thick terms; Evaluative language; Cancellability task; Cloze task; NLP; Polarity effects; Distributional semantics; Valence

Can Yvette be racist and good? An experimental and computational study on thick terms

Matteo Colombo

Thick terms, such as smart or racist, are at the same time evaluative and descriptive [1] unlike thin terms, e.g., good or right, which are merely evaluative, and purely descriptive terms, which express seemingly non-evaluative properties, e.g., Dutch.

In this work we clarify how thick terms combine evaluation and description using behavioral and computational approaches and ask three questions. First, do thick terms entail their evaluative content? Is the evaluative content of a thick term cancellable without contradiction? Second, what sources of information predict participants' felicity judgements in a cancellability task involving thick terms? Third, how well does a thick term's evaluative component predict expectations of upcoming words' polarity in a Cloze task?

We started by replicating the cancellability task by [2] to test their two conclusions that cancelling the evaluative content conveyed by a thick term in a sentence results in lower contradiction rating compared to cancelling the semantically entailed content of a non-thick term, and that positive thick terms behave differently from negative ones. Our results did not support these two conclusions. But, consistent with [2], the interaction between polarity of the target adjective and the thin attribute was reliable, with respondents indicating a prompt was more contradictory when the adjective's polarity matched the thin term's one.

Next, we used a Cloze task to explore whether the evaluative component of thick terms influences people when asked to complete a sentence. We asked participants to complete a sentence containing a thick term, manipulating its polarity and the conjunction (coordinating vs. adversative). We then fed the continuations to BERT's sentiment classifier [3], obtaining the support the model assigns to the positive and negative sentiment label. We observed a reliable interaction between the conjunction and the valence of the thick term. As expected, with a coordinating conjunction, the more positive the thick term, the more positive the continuation. The pattern was however reversed when the conjunction was adversative.

Overall, our work shows that the evaluative element of thick terms behaves like a semantic entailment; and we suggest that the stability of this evaluative element depends on robust patterns of affectively valenced associations acquired through experience of co-occurrent words within a shared cultural milieu.

References:

[1] Väyrynen, P. (2021). Thick Ethical Concepts. In Ed. N. Zalta (ed.), The Stanford Encyclopedia of Philosophy (Spring 2021 Edition).

[2] Willemsen, P. & Reuter, K. (2020). Separability and the Effect of Valence. Proceedings of the 42nd Annual Conference of the Cognitive Science Society, Austin, TX: Cognitive Science Society, pp. 794–800.

[3] Devlin, J., J., Chang, M.-W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. NAACL-2019.

ID: 191 / **Panel 3-2-C: 1**
Individual paper
Topics: Ethics (general and applied), Metaethics, Moral Psychology
Keywords: Envy, Emotion, Ethics

Benign envy?

Rob Compaijen

Amadeus (1984) is a vivid depiction of envy. In the film, court composer Salieri is torn apart by envy as soon as the extraordinarily talented Mozart enters the scene. Although his feelings towards Mozart are hostile from the outset, his envy gradually becomes so extreme that he begins to develop plans to murder him.

One answer to the question 'what is envy?' is not controversial. It is widely agreed that envy involves three elements: the envier (e.g., Salieri), the envied person (e.g., Mozart), and a good that the envier regards as constitutive for their self-worth (e.g., an extraordinary talent for writing music). Bringing these three elements together, we can understand envy more specifically as the painful experience of the envier for feeling inferior to the envied person, because the envied person possesses a good that the envier lacks and desires.

What is controversial, however, is a further issue: is envy necessarily vicious or can it also be benign? That is, does envy necessarily involve the desire that the envied person loses the good? (As is argued, for example, by D'Arms & Kerr (2008) and Fussi (2017).) Or is there also a benign form of envy, one that does not involve this desire? (As is argued, for example, by Rawls (1971), Neu (1980), Taylor (2006) and Protasi (2021).)

In this talk I will explore this controversial issue. More specifically, I will argue against the idea that there is benign form of envy. I will argue, that is, that envy necessarily involves hostility towards the envied person. In developing my argument, I will engage quite extensively with Sara Protasi's recent book The Philosophy of Envy (2021) in which she brings forward the most sophisticated defense of the existence of benign envy (she calls it 'emulative envy').

The most important point I seek to establish in my talk is this: defenses of purported cases of benign envy misrepresent the pain that is experienced by the 'envier'. If we look carefully at the cases that are presented as involving benign envy we will see that the pain the 'envier' experiences is about the lack of the desired good instead of the inferiority vis-à-vis the envied person.

The outline of this talk is as follows. In the first part of this talk I will develop a general account of envy, arguing that it involves three elements: (1) feeling inferior, (2) to a similar other, (3) with regard to a good that the envier regards as constitutive for their self-worth. In the second part, I will present an overview of the arguments of those who defend the existence of a benign form of envy, focusing specifically on a case presented by Protasi (2021). In the third and final part of the talk I will argue against the idea of benign envy and explain why I believe envy necessarily involves hostility.

The dark side of niche construction

Sabrina Coninx

In evolutionary biology, niche construction denotes the modification of the relation between organism and environment through a variety of activities other than selection, such as the modulation and structuring of environmental entities. While the concept of niche construction is originated in the study of phylogenetic processes (Aaby & Ramsey, 2019; Laland & Sterelny, 2006; Sterelny, 2010), it has been systematically broadened and applied to different areas of research, including cultural evolution (Flynn et al., 2013; Laland et al., 2000; Odling-Smee & Laland, 2011), ontogentic development (Colombetti & Krueger, 2015; Stotz, 2010; Sutton, 2010), and coordination with the environment in the here-and-now (Bertolotti & Magnani, 2017; Clark, 2005, 2006). The concept of niche construction has been used to better understand the physical constitution of humans as well as their cognitive, affective, and social abilities (Clark, 2005; Colombetti & Krueger, 2015; Sterelny, 2010).

The aim of this paper is to investigate in how far niche construction, which is commonly considered adaptive, can turn maladaptive, when different spatio-temporal scales are taken into account (Coninx & Stephan, 2021; Fabry, 2021; Sinha, 2015). Thus, of particular interest are those forms of niche construction which appear adaptive with respect to one spatio-temporal scale but maladaptive with respect to another. The paper primarily serves an explorative function indicating a useful conceptual tool to better understand central aspects of human life and the entangled contribution of evolutionary, socio-cultural, personal, and situational aspects. The focus of this paper will be on pain as an illustrative example. Pain lends itself to closer examination as it offers an evolutionary benefit but also represents in its chronification one of the greatest challenges to modern healthcare systems and concerned patients (Breivik et al., 2006). Further, the emergence and impact of chronic pain is linked to our modifications of environmental structures and lifestyle choices restricting activity and promoting aversive behavior (Büchel, 2021; de C Williams, 2016).

The paper proceeds as follows: In §2, I set the stage for the project by addressing what niche construction is and in which different manner it can be conceptualized. I introduce a distinction between four kinds of niche construction: phylogenetic, sociogenetic, ontogentic, and microgenetic. In that, the paper contributes to the general debate on niche construction by means of a systematic classification based on different grains of analysis. In §3, I introduce the concept of negative niche construction and discuss the most suitable normative criterion to distinguish between adaptation and maladaptation depending on the considered kind of niche construction. Although the idea of negative niche construction is not new, we are so far missing clear criteria to differentiate negative and positive kinds. In §4, I discuss how the concept of niche construction can help us to understand central aspects of modern life with a focus on human affectivity. The paper uses the example of pain to illustrate the potential clash of positive and negative aspects of niche construction across spatio-temporal scales. §5 summarizes these considerations and indicates further areas of application.

ID: 201 / Panel 1-2-G: 1
Individual paper
Topics: None of the topics above (please include your topic under 'keywords')
Keywords: Decision Theory, Practical Rationality, Hard Choices

Distinctively hard choices: creating reasons by choosing

Annalisa Costella

Choices such as those of one's career or whether to have a family are, often, a tall order. Under one family of accounts, individuals lack conclusive 'standard' reasons for choosing in these cases. Under another, individuals have reason to pick. Each family of accounts faces a challenge. Accounts of the first type need to explain how an individual should deliberate when she does not have a conclusive reason, what a non-standard reason is, and why it is appropriate for choosing. Call this the deliberative challenge. Accounts of the second type need to explain why picking in the face of distinctively hard choices does not hamper one's agency. Indeed, many choices that shape a person's identity are distinctively hard. Explaining how one ought to choose in these occasions by reducing the decision-making process to a random choice seems to degrade the individual to a wanton. A theory should thus be able to vindicate the idea that picking when faced with choices that may shape one's identity does not undermine a person's agency. Call this the arbitrariness challenge. I argue that current accounts cannot accommodate these challenges. As a remedy, I propose an account that draws on Chang's idea that individuals can create reasons. Contrary to her, I do not rely on the dubious assumption that one creates reasons by willing them. I argue, instead, that reason creation is located in the very act of choice.

ID: 201 / Panel 1-2-G: 1
Individual paper
Topics: None of the topics above (please include your topic under 'keywords')
Keywords: Decision Theory, Practical Rationality, Hard Choices

Distinctively hard choices: creating reasons by choosing

ID: 121 / **Panel 5-2-A: 1**
Individual paper
Topics: Social and political philosophy, Ethics (general and applied), Metaethics, Moral Psychology
Keywords: critical security studies, ethics, pluralism, security, value

The Ethics of Security: a value-pluralist Account

Josette Anna Maria Daemen

If decades of research on security have made one thing clear, it is that security has a tendency to become an overbearing objective. When a particular issue comes to be regarded as 'a matter of security', people accept that in public decision-making on this topic many other goals that normally carry significant weight in the political procedure are temporarily put aside. In exchange for security, apparently, we are generally prepared to pay a high price in terms of all other things we value.

Partly in response to this finding, critical security scholars in political science and IR have roughly taken two different tracks. Some have highlighted that security, in its tendency to become an overbearing goal, often significantly encroaches on the values of freedom and democracy. These scholars – most notably those belonging to the Copenhagen School, famous for its research on 'securitisation' – therefore see security as generally something negative to be avoided. Others have instead tried to use security's potential for overbearingness for good, hoping to direct its mobilising power to issues that they believe need urgent addressing, for example by connecting security to the relief of human suffering, the promotion of social justice, or the fight against climate change. These scholars – including those belonging to the Aberystwyth School, which redescribes security as 'emancipation', as well as those working with the concept of 'human security' – construct security as indeed something positive to strive for.

In this paper, I put forward a third avenue worth exploring when we analyse security from a normative point of view, which is rooted not so much in the fields of political science and IR, but rather in the tradition of ethics. Specifically, the account of security that I develop is inspired by the idea of value pluralism, as it has for example been defended by Isiah Berlin. On this view, there are several different moral values, which cannot be reduced to one 'supervalue'. In line with this notion, my account treats security as one value among many.

Although my perspective is distinct from the two branches of critical security studies mentioned earlier, it does combine insights from both. On the one hand, inspired by the contributions of securitisation scholars, my account acknowledges that security can conflict with, and come at the cost of, other widely held values such as liberty, equality, and democracy. On the other hand, in line with the work of scholars who try to import positive content into the concept of security so that it becomes a force for the better, my account conceptualises security not as something negative to be avoided but as something prima facie desirable. In short, on my view, security matters, but it is not the only thing that matters. By explicitly recognising this, my account of the ethics of security provides a helpful way of dealing with security's tendency to become overbearing, which is different from the two strategies familiar from existing literature in critical security studies.

ID: 135 / **Panel 1-1-F: 1**
Individual paper
Topics: Ethics (general and applied), Metaethics, Moral Psychology
Keywords: Bratman, institutional agency, collective moral responsibility

Shared Institutional Responsibility

Job de Grefte

One does not have to look far to find examples in recent newspapers where collective institutional agents are caught up in moral quandaries. For instance, in January 2021 the Dutch court ruled that Shell, inc. violated a duty of care towards various claimants in the Niger Delta. Many feel that the company not only is subject to a legal duty of care but to a moral one as well. Moreover, it is not just any subset of individuals within Shell that seems to be morally responsible but the company as a whole. Yet, it has turned out to be surprisingly difficult to make sense of moral responsibility of collective institutional agents (Sepinwall 2016).

Michael Bratman's new book Shared Institutional Agency (2022) might help us make progress in this regard. Here, Bratman constructs an account of collective agency that does not reduce to shared individual intentions. He explicitly stops short, however, of extending his argument to the ascription of moral responsibility to institutional agents. In this paper, I aim to continue where Bratman stops. I argue that we can use the resources his theory provides to construct a plausible account of institutional moral responsibility.

The basic idea is as follows. Bratman proceeds by a method of construction. Rather than aiming for necessary and sufficient conditions for institutional agency, he aims for a set of sufficient conditions. This allows him to provide a story of how an initial set of core capacities at the individual level, in Bratman's case our core capacity for planning, can give rise to a rich notion of collective institutional agency.

I propose we do the same for institutional moral responsibility. Concretely, I first go over some plausible criteria for institutional moral responsibility and show how they fit within Bratman's framework. Here I focus on capacities of rationality and moral judgment, as well as moral motions.

I argue that Bratman's framework is compatible with criteria of collective rationality and moral judgment, but not with criteria focusing on the capacity for moral emotion. But I argue this is a feature rather than a bug. Taking another leaf from Bratman's page, I argue that just as we should not expect agency to look identical on the collective level compared to the individual level, neither should we expect institutional moral responsibility to look exactly the same as individual moral responsibility. I close by suggesting some implications and directions for future research.

ID: 156 / **Panel 4-3-F: 1**
Individual paper
Topics: Social and political philosophy, Ethics (general and applied), Metaethics, Moral Psychology
Keywords: social privacy; genetic information; contextual integrity; criminal investigations

Crime-solving with Commercial DNA Databases: A Tyranny of the Minority?

Nina de Groot

Tens of millions of people have taken a 'for-fun' DNA test with a commercial company, driven by an interest in their genetic ancestry, distant relatives, or their athletic performance capabilities. Yet, the same commercial genetic data can also be used for an entirely different purpose: to identify criminal suspects. With this so-called 'investigative genetic genealogy' (IGG), the police uploads a crime-scene DNA profile in a commercial DNA database to identify a DNA test consumer who is a distant relative of the unknown suspect. Subsequently, one can examine where the family trees of the DNA test consumer and the suspect intersect to identify the suspect. This method has been used in hundreds of criminal investigations, mostly in the US. The debate tends to be reduced to the balancing of individual rights and interests (individual privacy, individual control over private information, and individual consent to participate in law enforcement access) versus the societal interests (the benefits of crime-solving). However, to consider privacy exclusively on the individual level leads to an undervaluing of privacy in utilitarian balancing – the prevailing way that policy-makers make decisions when conflicting interests are at stake (Solove, 2015).

In this paper, I propose to approach IGG through the lens of privacy's social value, in contrast to merely its individual value. First, I discuss the conceptualization of privacy as a social value. Next, I explore several issues of IGG that privacy's social value allows consideration for: the informational and decisional interconnectedness, the involvement of multiple sectors, the relationship between citizens and state, and the risk of a tyranny of the minority. In brief, with a tyranny of the minority, it only takes a relatively small amount of people to make a decision to share some of their data in order for third parties to infer that same information about others (Barocas & Nissenbaum, 2014). It is a concept predominantly discussed in Big Data contexts. I argue that it also occurs in IGG, because only 2% of the population has to be included in a DNA database accessible to law enforcement in order to identify almost anyone in that population.

I conclude that a social privacy approach offers a more fruitful perspective to evaluate the ethical desirability of IGG, evading the simplified dichotomy between individual privacy versus the security of society, in which the former will almost automatically lose. A focus on privacy's social value recognizes the effects for society on both sides of the balance.

References

Barocas, S., & Nissenbaum, H. (2014). Big data's end run around anonymity and consent. Privacy, big data, and the public good: Frameworks for engagement, 44–75.

Solove, D. J. (2015). The meaning and value of privacy. Social dimensions of privacy: Interdisciplinary perspectives, 71–81.

Persistence for the Permanentist

Catharine Elizabeth Diehl

There is a standard distinction in the philosophy of time between presentists–those who think that everything is present–and eternalists, those who think that there are non-present past and future things (Deasy 2019). Recently, however, Timothy Williamson has argued that this distinction—along with the modal distinction between actualism and possibilism—is misguided, because any attempt to clarify the notion of being present fails (Williamson 2012: 24–25). Williamson defines two new views: permanentism holds that it is always the case that everything is always something ($A\forall xA\exists yy = x$) and temporaryism, the negation of permanentism. He argues that we should abandon the old distinction between eternalism and presentism (along with the more prominent distinction between actualism and possibilism) and instead focus on the distinction between permanentism and temporaryism (Williamson 2012: 25).

This talk takes Williamson up on this challenge and asks what kind of theory the permanentist should have concerning the persistence of objects. I shall argue that permanentism sits badly with a popular theory concerning how objects persist through change, perdurantism. Perdurantism claims that objects persist by having different temporal parts located at different times. A four-dimensional object is composed of its temporal parts in much the same way that an object in space is composed of its extended parts. The perdurantist gambit is to use this strategy to explain changes of property: I have the property of sitting and the property of not-sitting, because one temporal part of me sits, while another temporal part of me stands. But I argue that this reduction of change is not available to the permanentist. There are cases in which an object changes property that are not grounded in differences in properties among the temporal parts that compose the object. In particular, changes in whether the object is concrete and whether it overlaps with an instantaneous temporal part of itself cannot be explained by the perdurantist schema, because the required temporal parts are not available.

While permanentism is thus compatible with the existence of temporal parts, these parts cannot serve the role in explaining away change that is re- quired by the perdurantist gambit, so change remains unreduced. Since the promise of perdurantism was to provide a reductive explanation of change, it seems that the permanentist should not endorse perdurantism. This leaves endurantism as the option of choice for the permanentist. This is not necessarily bad news for the permanentist, but it is surprising. B-theorists are frequently perdurantists, but since B-theorists are permanentists, then the popular combination of B-theory and perdurantism turns out to be unattractive.

ID: 154 / Panel 3-2-F: 1
Individual paper
Topics: Social and political philosophy, Ethics (general and applied), Metaethics, Moral Psychology
Keywords: Competition, Cooperation, Markets

What's wrong with competition?

Savriël Dillingh

Recent years have seen a marked resurgence of philosophical research into the concept of competition. Though where earlier philosophers have tended to treat competition, if properly governed, as a morally neutral mode of interaction, modern analyses favor the view that all forms of competition have the general tendency to bring about morally undesirable outcomes. On these accounts, competition variously harms, undermines self-esteem, or otherwise sets back the individual interests of at least one competitor. These undesirable outcomes are subsequently taken to be competition's chief moral defect. In contrast, the present paper argues that this type of consequentialist analysis is largely on the wrong track.

I demonstrate that while competition certainly may —and often does— bring about harmful outcomes, this particular feature is not unique to competition, nor what makes some forms of competition morally wrong. More specifically, I show that modern analyses tend to overestimate the negative effects competition may have ceteris paribus because they tend to rely on an overly individualistic conception of competition. Instead, I show how specifying competition as a mode of cooperation contradicts this supposed 'zero-sum' quality. Competition is, so I argue, characterized by rule-following; it is a type of social institution that governs competitors' interactions towards a mutually beneficial end.

Subsequently and notably, I argue that the very same specification clarifies the manner in which competition must by definition be morally inferior to other modes of cooperation. That is, a competitive mode of cooperation necessarily entails that participants are themselves reduced to a competitive quality and that their actions are quantified. Competitors are instrumentalized such that their competitive qualities are subordinate to the competition's end. In other words, competition as a mode of cooperation can only treat participants as means to an end and not as ends in themselves.

Crucially, I show that it is this very feature of competition that makes it an especially efficient mode of cooperation. Certain types of beneficial outcomes may even be achieved by a competitive mode of cooperation alone. Thus, modern critics have it exactly the wrong way around: while harmful outcomes cannot delegitimize competition, sufficiently beneficial outcomes may instead legitimize it.

Finally, an important corollary of this account is that if a different mode of cooperation is able to secure a similar good or similar amount of goods as competition, then that non-competitive mode will always be morally preferable. Thus, ultimately joining many modern critics of competition, albeit for vastly different reasons, I argue that it follows we should avoid distributing certain primary goods via competitive modes of cooperation such as the free market.

ID: 185 / **Panel 4-1-F: 1**
Individual paper
Topics: Social and political philosophy, Ethics (general and applied), Metaethics, Moral Psychology
Keywords: Criminal Justice, Respect, Social Agents, Rehabilitation

Rehabilitative justice: respect for persons as social agents

Emma Dore-Horgan

In perusing the philosophical literature that focuses on societal responses to criminal wrongdoing, we might be forgiven for thinking that the rehabilitation of those who have offended has nothing to do with justice. Justice (as per the received wisdom) involves criminal justice institutions 'putting things right' following the commission of crime, either by dispensing proportionate punishment or by requiring that those who have offended make some form of reparation to their victim(s). Rehabilitation, in contrast, involves criminal justice institutions striving to make society safer and/or evincing care and compassion for those who have offended: goals that are clearly rational and noble but which lack any connection – so the bulk of the literature would have us think – to justice or 'putting things right'. This paper argues, contra the received view, that the provision of rehabilitation is intimately connected to the achievement of justice. I assume, firstly, that criminal justice (like justice more generally) involves treating people in a manner that is appropriate given the kinds of beings that we are and given the relationship in which we stand to others. I contend that treating humans as the kinds of beings that we are requires that we respect the fact of our social agency: viz., that we are ultra-social beings capable and fundamentally in need of highly cooperative, socially integrated forms of living. I then argue that respecting the social agency of those who have offended requires offering rehabilitation following these individuals' conviction of crime – a demand that exists alongside, but independently of, any other requirement(s) of justice that crime occasions. I am consequently going one step further than those who suggest justice requires rehabilitation when individuals have been subjected to socially exclusionary and/or degrading punishments. I argue that rehabilitation is required when people are convicted of crime, regardless of the fact and manner of our punishment practices.

My discussion has four parts. Section one defends the thought that justice entails that we respect people as social agents, and claims, following Brownlee (2020), that respecting people's social agency requires that we strive to protect and preserve their social resources (i.e., their social abilities, social connections and social opportunities). Section two argues that respecting the social agency of those who have offended requires the offer of rehabilitation because a) conviction of criminal wrongdoing risks jeopardising one's social resources and/or serves as an indicator of already compromised social resources; and b) offering rehabilitation serves to acknowledge the fundamental need that convicted persons have, qua social agents, for an adequate set of social resources. Section three addresses a number of objections, including concerns that those who have offended forfeit their moral claim to respect as social agents; that the offer of rehabilitation is patronising and hence not a route by which we can evince respect for people as social agents; and that respecting convicted persons' social agency merely generates negative duties (e.g., that we refrain from meting out socially exclusionary punishments) and not positive duties of the kind I am envisaging. Section four concludes.

ID: 200 / **Panel 1-3-A: 1**
Individual paper
Topics: Epistemology, Social and political philosophy
Keywords: Democracy, Deliberation, Inquiry, Zetetic

Dance me to the end of love? The limits of political deliberation

<u>**Miguel Egler**</u>

Suppose there is a public debate in your country on whether vaccination against a disease should be mandatory. You attend a collective consultation session, ready to argue in favour of mandatory vaccination. There you are paired with a persistent political interlocutor (ppi), who constantly asks that you justify every claim you make, including your empirical evidence on the matter, arguments, and political views you hold.

While it may seem tempting to dismiss the ppi's requests for justification, matters are not so simple. Many normative theories of democracy claim you must justify your claims to the ppi, as no democratic citizen should have political decisions imposed on them. More schematically:

Qualified Acceptability Requirement (QAR): The exercise of political power is legitimate only if it is justifiable in terms acceptable to all qualified points of view (where "qualified" is equivalent to "reasonable" or some such thing).

However, there is a problem. Given the spread of misinformation, fake news, and polarisation that characterises the current political scenario, catering to the ppi may require engaging in a sort of follie à deux: a collective madness where you must consider even the most abstruse of the ppi's concerns. So what should you do?

In the first part of this paper, I look at ways of diagnosing the ppi's actions as problematic so that you can legitimately dismiss their requests for justification. I consider three proposals. The first claims that the ppi is not reasonable, which would violate the QAR. A second proposal is that the reasons motivating the ppi's inquiry are not public reasons—making them (politically) illegitimate. And lastly, I consider the idea that the ppi's requests for justification would lead to non-cognitivism about politics (where people vote for views they do not believe).

I argue that all these proposals fail. First, I use recent work on conspiracy theories to argue that the ppi can be entirely reasonable in his requests for justification. In a similar vein, I argue that it is possible to conceive of a ppi as driven only by public reasons. And against the third proposal, I argue that non-cognitivism about politics is not so troublesome as some would have us believe. This lays the groundwork for the second part of the paper, where I offer my own diagnosis of what is wrong with the ppi's actions.

My main contention is that we should reconceive democratic decision-making as a form of collective inquiry that is governed by zetetic norms: i.e., norms for how we should conduct inquiry. I first motivate this proposal by showing how it aligns with common sense views of democratic decision-making. I then argue that the persistent political interlocutor violates zetetic norms. In their repeated requests for justification, they stall inquiry—which amounts to an anti-democratic attitude. I conclude by considering how this proposal impacts in normative theorising about democracy.

ID: 214
Full symposium
Topics: Social and political philosophy
Keywords: climate

The role of philosophers in the climate crisis

Chair(s): **Thomas Fossen** (Leiden University)

Recent scientific research paints an increasingly bleak picture of the state of our climate and ecosystems. Meanwhile, governments globally seem unable or unwilling to tackle climate change effectively. Faced with the terrifying prospect of social and ecological catastrophe, many academics today struggle with their role. Increasing numbers take to the streets and engage in activism not just as citizens, but as scientists and scholars.

What is the role for philosophers in the climate crisis? How (if at all) should we reform our own research and teaching practices? Do philosophers have a special responsibility to get involved in climate activism?

Ingrid Robeyns (UU), Gerrit Schaafma (UvA), Tom Wells (Leiden), and Harriët Bergman (University of Antwerpen) take the lead in a discussion with the audience.

Presentations of the Symposium

Ingrid Robeyns
Utrecht University
Contribution to the discussion

Gerrit Schaafsma
University of Amsterdam
Contribution to the discussion

Tom Wells
Leiden University
Contribution to the discussion

Harriët Bergman
Universiteit Antwerpen
Contribution to the discussion

ID: 158 / **Panel 4-2-E: 1**
Individual paper
Topics: Epistemology
Keywords: Knowing-how, Knowing-that, Gettier problem, Science-Technology Relation

Is knowledge-how fundamentally different from knowledge-that? The argument from Gettier cases

Maarten Franssen

The notion of knowledge-how was introduced by Gilbert Ryle, in The concept of mind, as underlying intelligence. In contrast to knowledge-that, knowledge-how is non-propositional and builds on ability. It has remained controversial, however, whether something can, on the one hand, count as a form of knowledge but can, on the other hand, not be articulated. This debate took a new turn with the publication of the (2001) paper by Stanley and Williamson, who argued, to a large extent through reference to how the concept figures in our language, that there cannot be such a thing as knowledge-how as conceived by Ryle, because all knowledge is knowledge that. According to Stanley and Williamson, for some person to know how to Φ is for that person to know, of some way W, that W is a way to Φ. But if this is correct, it should be possible to construct Gettier cases of knowledge-how as we know them for knowledge-that: cases where all requirements for a person's knowing how to do something are satisfied but by a coincidence of accidents, such that this person cannot really be said to know how to do it. Stanley and Williamson acknowledged this and proposed a Gettier case of knowing-how in their paper. This case, however, did not convince sceptics and failed to settle the matter in favour of their claim. It has been argued both by authors who reject Stanley-and-Williamson's position on the identity of knowledge-how and knowledge-that (Poston 2009, Norström 2015) and by authors who are sympathetic to it (Cath 2011) that their Gettier case as well as other cases proposed since misfire because in all these cases the person still seems to have genuine knowledge-how. This strongly suggests that knowledge-how behaves differently from knowledge-that.

In this paper I argue that this apparent difference between knowledge-that and knowledge-how is illusory. I argue, first, that for knowledge-how there is no universally accepted fixed analogue to the 'justified true belief' conception of knowledge-that and that different conceptions of knowledge-how show differences in how easy it is to construct Gettier cases. Second, I argue that the seeming difference between knowledge-that and knowledge-how in the susceptibility to being 'Gettiered' is actually due to a singular-vs-universal difference in the nature of what is known that has little to do with the contrast between knowledge-that and knowledge-how. In support of this I present an example of knowledge-how for which a Gettier case can be constructed that is more plausible than any of the cases proposed so far, as well as an example of knowledge-that which resists being 'Gettiered' to the same extent as has been claimed is the case for knowledge-how.

References:

Cath (2011), 'Knowing how without knowing that', in Bengson & Moffett eds., Knowing how, OxfordUP, 113-135.

Norström (2015), 'Knowing how, knowing that, knowing technology', Philosophy and Technology 28: 553-565.

Poston (2009), 'Know how to be gettiered?', Philosophy and Phenomenological Research 79: 743-747.

Stanley & Williamson (2001), 'Knowing how', Journal of Philosophy 98: 411-444.

ID: 139 / Poster session 1: 2
Poster
Topics: Ethics (general and applied), Metaethics, Moral Psychology
Keywords: Ethics of technology, robot ethics, social robots, humanoid robots

On the very idea of humanoid robots

Cindy Friedman

When one thinks about the term "humanoid robots", the very first image that appears in our minds may be the likes of futuristic robots seen in the television series Westworld, or from movies such as Ex Machina or Artificial Intelligence. These are highly sophisticated technologies, the likes of which we do not yet have today. The humanoid robots we have today are a little more rudimentary. For example, Sophia the robot is impressive in that she can give unique verbal responses in different social contexts, and can also "recognize human faces, see emotional expressions, and recognize various hand gestures" (Hanson Robotics 2022). However, she doesn't come near the robots from science fiction just mentioned (Sharkey 2018). Despite this, however, there are still some very real ethical concerns that are coming to the fore, with regard to the development and proliferation of humanoid robots in society. Moreover, there is a big incentive to advance their development even further (Telving 2022), so we may not be as far off from more sophisticated robots as we may initially think.

Given this, robot ethics is a rapidly growing field, in which people engage with questions such as whether robots pose a risk to society, whether the use of certain robots is problematic (Sparrow 2007), whether certain designs are problematic (Cave & Dihal 2020), whether they can be moral agents (Wallach & Allen 2009) and moral patients (Gunkel 2018), amongst others.

Humanoid robots, however, are particularly ethically interesting. They are developed to have a realistically human-like appearance, and simulate human-like behaviour, which has the potential to elicit certain social (and even emotional) responses from those who interact with them. Moreover, they have the potential to fulfil various roles in society. For example, Bina is a humanoid robot that has co-taught an intro to ethics course, thereby fulfilling the role of a teacher (Houser 2018). And Harmony is a sex robot, which has the potential to fulfil the role of a human partner.

Within the literature, the development of humanoid robots has been met with mixed reactions: some authors argue that we should not create humanoid robots at all; whereas others argue that they are good as means to certain ends. Some others, on the other hand, think that creating them might be an end in itself. The aim of this presentation is to offer a general take on how we may go about approaching the very idea of humanoid robots, both in terms of defining what humanoid robots are, as well as in terms of how to go about assessing them from a normative point of view. Through discussing these varied opinions about humanoid robots, and bringing my own argument to the fore, I highlight how important it is that we engage with the topic of what humanoid robots are, and think carefully about how (or even whether) we should move forward with their development. In doing so, I engage with, and respond to, authors such as Kathleen Richardson, Joanna Bryson and Robert Sparrow.

ID: 173 / **Panel 3-3-D: 1**
Individual paper
Topics: Social and political philosophy
Keywords: climate change, exchange, innovation

Is exnovation policy a socially acceptable mechanism for sustainability

Lukas Fuchs, **Rafael Ziegler**

Innovation scholars increasingly argue that innovation does not merely have a rate, but also a direction, such as "inclusive" or "sustainable" innovation. As a result, sustainability can be understood not merely as a side-condition of successful innovation, but as a new socio-economic paradigm that is the goal of innovation driving economic change. Inversely, various innovation discourses claim a central place of transformation towards sustainability.

However, in the light of persistent pro-innovation bias - innovation is good, always and should be accelerated - it can be asked if this discourse shift is sufficient for reversing ongoing unsustainability tendencies. As a result, there are also calls for discussing exnovation, which is the deliberate termination of existing (infra)structures, technologies, products and practices. For example, the EU has announced it will ban the sale of petrol or diesel cars by 2035. Without termination policies, sustainability innovations (for example in renewables) might simply add to existing uses (of non-renewables), but not actually reduce, or sufficiently reduce unsustainable resource use and pollution production. Moreover, the anticipation of future prohibition (as in the EU example) puts pressure on market actors to adapt and shift to different technologies (electric cars), making the establishment of a new attractive mass market more likely. Innovation and exnovation are not just about introducing and removing combinations in the market but may also transform other economic distribution mechanisms and non-economic aspects of public life, such as public service provision and social practices. This paper will focus on the social acceptability of exnovation.

First, what is the relationship between innovation and exnovation as mechanisms for changing direction? A key factor in the attractiveness of innovation is that it promises to bring about improvements through the forces of evolutionary economic development. New combinations in the form of new products, ways of production or distribution are perceived as qualitatively better or simply cheaper compared to previous products. Economic change can occur without the issuing of prohibitions, placing of incentives on new products or otherwise steering the behaviour of consumers, sellers and producers. By contrast, exnovation seeks to bring about change by prohibiting a product or a service, which raises classical concerns about the legitimacy of the state interfering in market processes.

Second, what follows from the response to the first question for the the requirements of a successful implementation of exnovation? Not all kinds of exnovation may be designed equally well. Exnovation seems to require a time horizon that is long enough for market actors to adapt and to explore alternative technologies and markets. Special consideration must be given to the likely effects of exnovation on the most vulnerable citizens in society (this is the prime concern in the context of the German Energiewende, where the phasing out of nuclear energy has affected the least well off disproportionately). Formulating the key requirements for socially acceptable exnovation will be a crucial step in understanding the role and limits of this policy mechanism for transformation towards sustainability.

ID: 131 / Panel 4-1-E: 1
Individual paper
Topics: Epistemology
Keywords: knowledge, justification, vigilance, future, teleology

The future of Knowledge

Victor Gijsbers

In contemporary analytic epistemology, knowledge is usually defined as a state of a subject at a time. Whether the subject is in that state at time t depends at minimum on the state of both subject and world at time t: the subject must be in a state of belief, and the world must be such that it makes the belief true. Most philosophers also acknowledge a dependence of knowledge on the past, for instance through a past-sensitive justification condition or through the condition that the belief must have been produced by a reliable process. So whether a subject has knowledge at time t depends on what happens at or before t. None of the prevalent analyses acknowledges a dependence on what happens after t.

My greater project is to claim that the basic idea of epistemology should be the quest for knowledge, which is a goal-directed project aimed at an indefinite improvement of our grasp of the world; and that other epistemological categories can only be understood in the context of this larger quest. In this talk, I will focus on the idea that individual items of knowledge are successful moves in this goal-directed project and are thereby dependent on a teleological orientation towards the future.

Suppose that S has a justified true belief that p, but holds this belief in a dogmatic way such that they wouldn't give it up even in the face of recalcitrant evidence. Can S be said to know p? It would seem that their epistemic state falls too far short of perfection for that. In fact, it seems that there is a normative symmetry between, on the hand, sensitivity to reasons in the past (justification) and, on the other hand, sensitivity to reasons in the future (which we can call epistemic vigilance).

I then work out the relation between justification and vigilance more systematically by turning to the value problem. Why is justified true belief better than mere true belief? I argue that being connected to other beliefs through links of justification is essential for making a belief capable of playing an active role in the project of knowledge. When, in the future, a belief is confirmed or disconfirmed by experience, this has immediate repercussions for those beliefs that are connected to it, and sometimes also for the higher-order principles that connect them. Thus, if my non-justified belief about the road to Larissa turns out to be false, I merely arrive in the wrong place. If my justified belief about the same turns out false, I arrive in the wrong place, but I can also improve my estimation of the reliability of the old map I bought, and thereby of all the other beliefs that I gained from looking at it. Thus justification is crucial for the quest for knowledge; more crucial, in a sense, than truth itself. But justification can play this role only when it is accompanied by vigilance. Justification and vigilance are two sides of one coin.

ID: 178 / **Panel 1-3-F: 1**
Individual paper
Topics: Social and political philosophy, Critical Theory
Keywords: Housing, segregation, egalitarianism, marginalization, spatial justice

Housing allocation and spatial inequality: the exclusionary city space

ELISABETTA GOBBO

Housing protests are unraveling in most big European cities at the beat of chants such as "Huizen voor mensen niet voor winst" (woonopstand.nl) and "Wohnen enteignen" (dwenteignen.de). These slogans highlight that citizens perceive an injustice in the way in which housing is allocated (based on one's economic availability to buy or rent). But what, if anything, is wrong about it? Regrettably, philosophers have not paid enough attention to this relevant problem to be able to ground protesters' claim on a theory of housing justice.

In this essay I wish to amend this lack by focusing on one aspect of spatial justice in the urban context: justice in housing, and in particular on just housing allocation. I contend that governing the allocation of housing according to households' income availability is problematic because it creates spatial inequality (specified into segregation and marginalization), which undermines the conditions that citizens of a democratic state need to live in if they are to relate as social equals. In a nutshell: if we are to inhabit a space as social equals, the allocation of living facilities in that space cannot discriminate based on income.

I will develop my argument as follows. First, I will focus on the problem of segregation. Citizens with lower income levels can only afford to live in certain areas of the city, the outskirts, while the richer residents inhabit the city center. Whilst it is common to have intuitively unproblematic segregation in other instances (access to luxurious restaurants and holiday resorts), I contend that segregation in living space is normatively relevant. I will clarify under which conditions and why income-dependent segregation, which is seemingly not directly enforced by public and private institutions with discriminatory goals, is problematic.

Second, I will focus on income-based marginalization. I contend that when housing is allocated based on income, there is inequality in the access to advantages in city-space. Indeed, the higher the income, the better the quality of housing, the safer the location they can live in, the more the private (cafes, theaters, gyms) and public investments (infrastructure, schools) they can benefit from in the area they live in. When households are shielded from these advantages because of their lower income, they do not relate as equals with the other city dwellers. One one side, the disadvantaged are placed in a self-reinforcing position of disadvantage. On the other, when the inequality of investment is state-led or state-facilitated, lower-income households are disrespected by public institutions which do not value their living conditions as highly as the ones of their richer co-citizens. In this section I will contribute to the existing philosophical literature on gentrification and housing by bringing the mechanisms of capitalist space production to the fore.

Finally, I will consider some proposed solutions to the problem of housing allocation. I suggest that ultimately, an egalitarian take on housing justice might entail rather radical changes in housing allocation and the distribution of investments in city-space.

ID: 213 / Panel 1-1-H: 1
Individual paper
Topics: Philosophy of Science, None of the topics above (please include your topic under 'keywords')
Keywords: AI, Affective computing

Biased affective computing and A.I.

Kris Goffin

Affective computing is on the rise. Affective computing is a term that refers to all forms of ways that computers are used to analyze emotions. A prominent example is emotion recognition software. Different algorithms and AI's can track given an input (facial expressions, vocal expressions) which emotions one is feeling. This is often used by corporations in generating personal advertisements; as in educational software.

In this paper, I will focus on bias in affective computing and identify strategies for mitigating bias and ensuring responsible use of A.I.

More specifically, I will focus on the following controversial applications:

1. Advertisement: Emotion recognition software used by companies that create algorithms for targetted advertisement.

2. Bias in affective computing for psychiatric applications: An interesting application is the use of affective computing in psychiatric conditions for autistic and neurodiverse individuals, such as software that is used to help autistic people recognizing emotions. I will investigate how these applications can combat (and possibly also reinforce) social injustices.

3. Border control technology: emotion recognition software is used in technology checks whether people form a threat. For instance immigrants and asylum seakers facial expressions are recorded and scanned.

Most of these affective computing A.I.'s rely on Basic Emotion Theory (BET). BET states that people only have a limited set of (universal) basic emotions. Each emotion correspond to a characteristic facial expression.

Constructivists, such as Lisa Feldman Barrett, criticize BET-based affective computing by stating that emotions lack universal markers. They argue that thare or no basic emotions; there is no fixed set of emotion, which is universal. Also there is no one to one correspondence from facial features to emotional kinds.

In this paper, I present an in-between account in which I state that emotion recognition software is biased. Some biases are due to cultural differences. Others are due to more insidious racialized and gender-based stereotyping. So it not the case that there is a crystal clear one to one correspondence between facial expressions and emotional kinds. Nor is it the case that "anything goes". It is rather the case that emotion recognition software is more or less able to track different emotional kinds, but it can be biased in various ways. This will lead to (sometimes very harmful) errors.

I will investigate emotions and emotional recognition software with an eye on questions regarding social injustices and identify strategies for ensuring responsible use of A.I.

ID: 127 / **Panel 1-1-D: 1**
Individual paper
Topics: Critical Theory, Hermeneutics, Feminist philosophy
Keywords: Canon, deconstruction, postcolonial-feminism, postcritique, hermeneutics

Beyond the Canon? Decanonization after Critique and Postcritique

Lucas Gronouwe

This paper encourages academic philosophers in the Netherlands to reconsider three questions: What do we study and teach? How do we study and teach? And why do we study and teach? The answer to the first question is no longer quite self-evident. In the past decades, the philosophical canon has come under siege for being too male, too white, and too Western. Polarization prevails between those who defend the canon as it stands, emphasizing the value of tradition, and those who call for its diversification, stressing the value of diversity. Meanwhile, some have drawn our attention to the limits of diversification and have instead challenged the very idea of a canon for philosophy. To move this conversation forward, one needs to do justice to the concerns of each of these parties. We cannot simply do away with Kant, for instance, but we cannot close our eyes to his sexism and racism either. Hence, we have to reconsider how we study and teach particular philosophers. And if we want to move beyond the canon, we also have to reevaluate why we study and teach certain texts in the first place. After a brief introduction, I turn to these two questions.

First, how do we study and teach? This paper confronts three approaches that provide an answer to that question. The first is deconstruction: a way of reading, initiated by the French philosopher Jacques Derrida, that focusses on contradictory and marginal aspects of philosophical texts. The second is postcolonial-feminism: a critical movement within the humanities, that demonstrates how the androcentric, ethnocentric, and Eurocentric biases of the philosophical tradition lead to the exclusion of certain social identities. The Indian scholar Gayatri C. Spivak has critically adapted deconstruction for these purposes. Although both of these approaches focus on the dynamics of inclusion and exclusion in canonical philosophical texts, they recognize that one cannot do away with tradition: we still have to read Kant; we just have to read him in a non-traditional way. In this part of the paper, I provide a brief outline of what such a non-traditional way might entail.

Deconstruction and postcolonial-feminism are confronted with a third approach, which is called "postcritique." This is both a theory and method of reading and interpretation, coined by literary scholar Rita Felski, that resists critical or suspicious forms of reading, and instead seeks to reveal how texts can be seen as actors that create various kinds of attachments in present-day readers. What if we would translate this theory within the discipline of philosophy? That would certainly force us to reconsider why we study and teach certain texts in the first place. In this final part of the paper, I consider alternatives for the justification criterium of canonicity, such as the capacity of philosophical texts to change how we think and feel, to enchant or shock us, or to assist us in addressing problems we currently care about. I conclude by discussing how these criteria might reconcile the different parties in the canon debate.

ID: 211 / **Panel 5-1-H: 1**
Individual paper
Topics: Social and political philosophy
Keywords: Dewey, Bennett, new materialism, vitalism

A Dialogue between Jane Bennett and John Dewey: An inquiry into the possibility of a vital materialist understanding of public

<u>Irem Guven</u>

In her Vibrant Matter, Jane Bennet establishes a 'flattened' ontology to explain what public is. According to her, how a public starts existing cannot be solely explained by human will. Humans belong to the public defined as 'political ecologies'. These political ecologies are shaped by a common experience of harm defined as a problem. With a flat and relational ontology, Bennett aims at defining the social from an ecological point of view, and the individuals (or "things") in their relationality to other individuals (human or non-human) that make up the social. In doing this, she references John Dewey's The Public and Its Problems and claims that a public is not formed as an act of will, but in reference to a problem.

Dewey's text is from a hundred years ago, written at a time when humanities were affected by Darwin's thesis on evolution. This was a period, similar to ours, when human agency was problematised extensively. It is indeed true that Dewey mentions the affinity between animals' "herds, packs and swarms" and human society. He even mentions that this attraction to come together may be found in the inanimate world, alluding to chemistry and physics. So it is not surprising that Bennett, in an attempt to locate the human within its environment, employs Dewey as a reference for her political theory.

In searching for an answer for how a public is formed, Dewey argues that focusing on the phase of human action "to which direct causative power is attributed" amounts to creating a mythology. Instead, one should focus on the consequences of human actions. And these consequences differ in accordance with the "material culture" of a given society, which points at a situated approach to human beings. Hence the public is formed as a consequence of the responses given to the problems.

In this paper, I would like to problematize Bennett's employing Dewey in a vital materialist framework that relies on a flat ontology. Does the critique of liberal individualism in Dewey amount to a parallelism with Bennett's vital materialism? Is Bennett's flat ontology compatible with Dewey's definition of the public as a response to problems?Is a flat ontology the best way to make Dewey philosophically relatable today?

ID: 196
Full symposium
Topics: Intercultural philosophy, None of the topics above (please include your topic under 'keywords'), Critical Theory
Keywords: Diversifying Philosophy

Making spaces for new faces: recent approaches in diversifying philosophy

Chair(s): **Annemie Halsema** (Vrije Universiteit, Netherlands, The), **Lilith Lee** (Vrije Universiteit)

Within Anglophone academia recently, increasing efforts are seen to address the traditional lack of addressing diversity in the research and teaching of philosophy. In line with it, in Dutch universities diverse initiatives are taken to destabilize and open up the Western philosophical canon. In some cases, such as in Leiden, a complete Bachelor and Master program Global and Comparative Philosophy is offered to students. In other cases, individual teachers aim at redesigning their own courses in order to address the demand to diversify the curriculum.

The aim of this symposium is to facilitate the exchange of experiences between universities with diversifying philosophy, to share strategies and combine fruitful perspectives. We invited staff members and students of 5 universities to discuss how they aim at diversifying philosophy, the problems they face and the opportunities they envision.

The speakers and talks, in no order of significance, are as follows (abstracts further below):

1. dr. Katrine Smiet (Radboud University), "Site for Unlearning: Philosophy Department"

2. dr. Ruth Rebecca Tietjen (Tilburg University), "Disrupting the philosophical canon in the mode of playful earnestness"

3. dr. Lilith W. Lee (VU Amsterdam), "Reading the Cracks and Detritus of a Frozen Core: An Instance of Decolonising and Diversifying History of (Political) Philosophy"

4. Fabius Schoendube and Clara Mendes Pereira (Leiden University), "Giving a space to affects"

5. dr. Caroline Suransky (University for Humanistics): Title TBA + abstract follows soon.

After the talks (with short pauses in between each to facilitate consolidation and articulation of thoughts), speakers are invited to respond to each others' talk, before discussion is then opened to the floor.

Presentations of the Symposium

Site for unlearning: philosophy department
Katrine Smiet
Radboud University

The standard philosophical canon and curriculum is marked by eurocentrism, ethnocentrism and androcentrism. Increasingly, teachers and students call for a more diverse curriculum and more inclusive teaching. However, in order to realize that, simply introducing some new/different texts into existing syllabi is not enough. Instead, drawing on the work of postcolonial philosopher Gayatri Spivak (1990, 1999), we posit that a process of 'unlearning' is necessary: to question the epistemic norms and disciplinary conventions that underpin this curriculum, and to cultivate receptivity for different perspectives and traditions of thought. The project 'Site for Unlearning: Philosophy Department' at Radboud University takes on this challenge. A collective consisting of students, teachers and education support staff form the motor for a collective exploration of the potential and challenges of unlearning. The collective will set out on a collaborative inquiry, asking:

1. (Why) is unlearning relevant for philosophy and philosophy education?

2. What do we need or want to unlearn in order to realize more diverse and inclusive education in philosophy?

3. How can we (start to) unlearn that?

In this action research project, action and reflection go hand in hand. The collective will identify 'sites for unlearning', and propose concrete interventions and changes, as well as reflecting on the ups and down of the process, the resistances encountered and the challenges it poses.

Disrupting the philosophical canon in the mode of playful earnestness

Ruth Rebecca Tietjen
Tilburg University

In my contribution, I defend the idea that we should disrupt the philosophical canon in the mode of playful earnestness. I do so by proceeding in three steps. First, I offer a critical reflection on what a or the philosophical canon is, thereby drawing on the contemporary debate about "the" philosophical canon and its critique (e.g., Alcoff 2017; Dotson 2012; Rée 2002; Waithe 2015; Westphal 1993). Particularly, I explore how, as members of institutionalized academic philosophy, we undergo processes of subjectification that subject us to, make us adopt, and reify certain (narrow and oppressive) ideas of what philosophy is, how it should be done and presented. It is these implicit rules and norms that constitute what I take to be "the philosophical canon." These ideas in turn culminate in our image of "the philosopher." Second, I explore the question of who and what gets excluded in this process (specific genres, topics, groups of people) and show how this exclusion is aesthetically, epistemically, and politically problematic. For instance, there is hardly room for surprise when already beforehand we clearly have to/want to state what will happen in our talk, our lecture, our paper. This is incompatible with the idea that philosophy is and should be an emancipatory and dialogical enterprise – and activity or event rather than the result of an activity. Third, in the practical spirit of this symposium, I offer some insights into my own attempts to disrupt my and our habitualized forms of philosophical thinking and acting (writing, speaking, learning, and teaching) with performative and other creative tools. I argue that playful earnestness is the right mode for this kind of enterprise because, on the one hand, we need to take the task to change ourselves, (academic) philosophy, and our society at large existentially and politically seriously and, on the other, we need freedom and creativity (i.e., playfulness) for our liberatory practice to flourish. (If you are sensing a tension between what I say and how I say what I say, you are on the right track; would a conference abstract be accepted that consists of nothing but an image?)

References

Alcoff, Linda Martín (2017). Philosophy and Philosophical Practice: Eurocentrism as an Epistemology of Ignorance, The Routledge Handbook of Epistemic Injustice, ed. by James Kidd, José Medina, Gaile & Pohlhaus, Routledge, 397–408.

Dotson, Kristie (2012). How is this paper philosophy? Comparative Philosophy 3 (1), 3–29.

Rée, Jonathan (2002). Women philosophers and the canon, British Journal for the History of Philosophy, 10 (4), 641-652.

Waithe, Mary Ellen (2015), From Canon Fodder to Canon-Formation: How do we get there from here?, The Monist 98 (1), 21–33.

Westphal, Merold (1993). The Canon as Flexible, Normative Fact, The Monist 76 (4), 436–449.

Reading the Cracks and Detritus of a Frozen Core: An Instance of Decolonising and Diversifying History of (Political) Philosophy

Lilith Lee
Vrije Universiteit

Geologists reconstruct past climates with collections of cores extracted from ice mass, allowing humanity understanding of the global implications of climatic events that have occured, are occurring, and are to occur. Frozen cores with cracks are limited in educating us less on our past, present, and future and more on how the core was produced. The traditional core curricula of history of philosophy, taken from a 'Western' field reaching down to Thalesian waters and enshrined in hallowed

halls, are not only often the only samples available to understand intellectual climates of the past couple of centuries, but have also cracks that tell more of the colonial production of these cores than of a philosophy fit for humanity (cf. Mills 2002, Wynter 2006, Park 2013, Cantor 2022).

For critically-oriented historians of philosophy, the multiculturalist imperative to "diversify or die" (Van Norden 2017) commands us to (i) move beyond speculating about the traditional collection and produce new cores not only in the 'Western' field (e.g. Black, queer) but also in new fields altogether (e.g. Pre-Columbian, Sinophone). At the same time, we must also (ii) trace the cracks in the existing

cores: to caution us against reproducing the same epistemic and ethico-political failures in new cores.

Further, under the auspices of the owners of those hallowed halls, we continually must (iii) negotiate the cores to be replaced in an industrial complex structured by traditional ones. Under such conditions, therefore, the imperative rather commands a tensive triad of directions: (i) diversify, (ii) decolonise, or (iii) declutter.

I begin this presentation by observing that, for the same educator, the tensive command to diversify philosophy commands not only that their students, many of whom would not remain in academia, (i) be introduced to an enriched understanding of history of philosophy from cores gathered from varied intellectual climates. They have to also (ii) be equipped with literacy concerning intellectual coring practices and the colonial cracks in, and implications of, our existing curricula—with the latter premised on students (iii) being familiarised with the traditional curricula that all this fuss is being made over.

I offer for critique and solidarity an experiment in following the tensive triad in a recent introductory history of philosophy course in VU Amsterdam's PPE programme, attending specifically to the design and delivery of the course content. The course design (i) drew from a diverse intellectual geography (Chinese, Anglo-European, Indian, and Caribbean), (ii) focusing not only on critical concepts of

ideology, colonialism, and genealogy, but also (iii) traditional conceptualisations of tradition, polis, law.

The course (i) required students to collectively engage with the complete range of material and (ii) encouraged them to trace the cracks in, and implications of, traditional conceptualisations (e.g. from Aristotle's eudaimonist justification for slavery to Fanon's materialist critique of colonialism), (iii) while nevertheless working closely with thinkers and concepts of traditional core curricula (e.g. Plato, Aristotle, Aquinas, Kant, Mill). I conclude by discussing student assignments and feedback, as well as possible future directions for this course.

Giving a space to affects

Fabius Schoendube, **Clara Mendes Pereira**
Leiden University

With the push to make philosophy more diverse, educators and students have called for revalorizing the emotional core of experience that originates thought. As students and heads of a study association, we observe this push all around us in the early attempts of students to articulate their own thoughts in an often colorful array of ideas, influences and experiences. While the diversity at the core of philosophy is evident in this excitement, we find, simultaneously, a tendency in these young thinkers to disregard the value of their thinking, resulting in constant attempts to justify themselves with abstract vocabulary within academically sanctioned debates. Such a contrast yields a limitation that not only restricts the types of thinking students pursue, but more broadly directs discussions to already well trodden territory hostile to critical re-articulations by groups that have historically been negated as knowledge producers. The challenge is further compounded with another complexity: when we begin with a diverse set of thinkers that are impelled to draw from their complicated lived experiences at the intersection of identities to do philosophy, we find that their ideas quickly take the form of culture war echoes that, with the emotional load of justifying one's own existence daily, tire them out

and make actual novel thinking difficult.

We contend that study associations can function as playgrounds that help in diversifying the communicative tools that enable student's full involvement as members of the philosophy community. We are in a privileged position to form a strong student community that facilitates encounters between thinkers at various intersections of identity, while also serving as a bridge for productive conversations between tutees and

faculty. By committing to the idea of student input as epistemically valuable, our

association provides constant feedback to our pedagogues, promotes socio-academic involvement at our varied events such as cinema events, creative writing sessions, podcasts, student-led courses and conferences that work in tandem to fill the gaps in themes and approaches that our curricula naturally have with.

Zooming in, the project "Affect at the Intersection", developed in collaboration with

professor Jingjing Li, nudges our peers to attend to the dimension of emotion underlying embodied experience in order to think critically through the socio-political dimensions of their lives. Analyzing the particular anxieties facing this generation, we aim to reveal the ways in which a vast diversity of possible thinking is normalized by structures that are determined to induce nothing but nihilism in their participants. Facing these difficult and complicated emotions with an intersectional philosopher's toolkit that closely attends to the treasure trove that is lived experience, we work to afford both ourselves and our peers the space to creatively rethink oppressed identities and hope so loudly that we regain the strength to affirm our lives against political hopelessness.

Diversifying Philosophy – wrestling with the canon in Humanistic Studies

Caroline Suransky
University of Humanistics

A canon presupposes and encapsulates a theory of history, a moral theory, a conception of society and assumptions around exemplary science (Burawoy, 2021). Furthermore, a canon is foundational – as it grounds the production and sharing of knowledge; it is relational as the selected concepts, people and their ideas do not stand in isolation but speak to each other and it is historical, as it is challenged and changed over time (Burawoy, 2021).

The University for Humanistic Studies profiles itself as an inter- or transdisciplinary university that is inspired by the worldviews and traditions of humanism. It focusses on contemporary issues with a normative component towards developing humane and caring societies. Rather than orienting itself to a singular disciplinary canon, it wrestles with diverse canons and is challenged to figure out what it concretely means when it seeks inspiration from humanist worldviews and traditions.

Currently, two such challenges stand out. Firstly, the challenge of anthropocentrism and the development of ecohumanism in the Anthropocene. Flowing from the era of Enlightenment-Modernity, independence and autonomy became important basic values of humanism. However, the contemporary challenges posed by the global ecological crisis are inextricably connected with fundamental questions about being human in relation the earth. This requires us to reexamine our foundational humanistic philosophies as basis for how we think and act. It requires a paradigm shift in which the ideal of the autonomous human being makes way for a paradigm of connectedness in the context of complex ecological processes.

Secondly, the University of Humanistic Studies is still a very white institution, almost all academic staff and students are white. While not denying diversity within white communities, we may conclude that Western humanism, flowing from Modernity has been rather blind to the composite notion of modernity-coloniality which refers to the way in which the two concepts are inseparable (Quijano, 2007, Mignolo, 2018). Enlightenment thinking became a normative ambition when modern Europe presented itself as the antithesis of the rest of the world, which they qualified as pre-modern, 'primitive' and 'barbaric'. While the West advocated universal equality and freedom for 'humanity', it showed a very different face in the colonies. How to address this legacy is an important question while Humanistic Studies wrestles with competing canons and paradigms.

And so we search new interlocutors with whom the 'humanistic canon' could be critically reexamined in the light of the criteria, theories and orientations to the canon as foregrounded by Burawoy.

ID: 234 / **Panel 2-3-A: 1**
Symposium paper
Topics: Intercultural philosophy, None of the topics above (please include your topic under 'keywords'), Critical Theory
Keywords: Diversifying Philosophy

Reading the Cracks and Detritus of a Frozen Core: An Instance of Decolonising and Diversifying History of (Political) Philosophy

Lilith Lee
Vrije Universiteit

Geologists reconstruct past climates with collections of cores extracted from ice mass, allowing humanity understanding of the global implications of climatic events that have occured, are occurring, and are to occur. Frozen cores with cracks are limited in educating us less on our past, present, and future and more on how the core was produced. The traditional core curricula of history of philosophy, taken from a 'Western' field reaching down to Thalesian waters and enshrined in hallowed

halls, are not only often the only samples available to understand intellectual climates of the past couple of centuries, but have also cracks that tell more of the colonial production of these cores than of a philosophy fit for humanity (cf. Mills 2002, Wynter 2006, Park 2013, Cantor 2022).

For critically-oriented historians of philosophy, the multiculturalist imperative to "diversify or die" (Van Norden 2017) commands us to (i) move beyond speculating about the traditional collection and produce new cores not only in the 'Western' field (e.g. Black, queer) but also in new fields altogether (e.g. Pre-Columbian, Sinophone). At the same time, we must also (ii) trace the cracks in the existing

cores: to caution us against reproducing the same epistemic and ethico-political failures in new cores.

Further, under the auspices of the owners of those hallowed halls, we continually must (iii) negotiate the cores to be replaced in an industrial complex structured by traditional ones. Under such conditions, therefore, the imperative rather commands a tensive triad of directions: (i) diversify, (ii) decolonise, or (iii) declutter.

I begin this presentation by observing that, for the same educator, the tensive command to diversify philosophy commands not only that their students, many of whom would not remain in academia, (i) be introduced to an enriched understanding of history of philosophy from cores gathered from varied intellectual climates. They have to also (ii) be equipped with literacy concerning intellectual coring practices and the colonial cracks in, and implications of, our existing curricula—with the latter premised on students (iii) being familiarised with the traditional curricula that all this fuss is being made over.

I offer for critique and solidarity an experiment in following the tensive triad in a recent introductory history of philosophy course in VU Amsterdam's PPE programme, attending specifically to the design and delivery of the course content. The course design (i) drew from a diverse intellectual geography (Chinese, Anglo-European, Indian, and Caribbean), (ii) focusing not only on critical concepts of

ideology, colonialism, and genealogy, but also (iii) traditional conceptualisations of tradition, polis, law.

The course (i) required students to collectively engage with the complete range of material and (ii) encouraged them to trace the cracks in, and implications of, traditional conceptualisations (e.g. from Aristotle's eudaimonist justification for slavery to Fanon's materialist critique of colonialism), (iii) while nevertheless working closely with thinkers and concepts of traditional core curricula (e.g. Plato, Aristotle, Aquinas, Kant, Mill). I conclude by discussing student assignments and feedback, as well as possible future directions for this course.

ID: 235 / Panel 3-1-A: 1
Symposium paper
Topics: Intercultural philosophy, None of the topics above (please include your topic under 'keywords'), Critical Theory
Keywords: Diversifying Philosophy

Giving a space to affects

Fabius Schoendube, **Clara Mendes Pereira**

Leiden University

With the push to make philosophy more diverse, educators and students have called for revalorizing the emotional core of experience that originates thought. As students and heads of a study association, we observe this push all around us in the early attempts of students to articulate their own thoughts in an often colorful array of ideas, influences and experiences. While the diversity at the core of philosophy is evident in this excitement, we find, simultaneously, a tendency in these young thinkers to disregard the value of their thinking, resulting in constant attempts to justify themselves with abstract vocabulary within academically sanctioned debates. Such a contrast yields a limitation that not only restricts the types of thinking students pursue, but more broadly directs discussions to already well trodden territory hostile to critical re-articulations by groups that have historically been negated as knowledge producers. The challenge is further compounded with another complexity: when we begin with a diverse set of thinkers that are impelled to draw from their complicated lived experiences at the intersection of identities to do philosophy, we find that their ideas quickly take the form of culture war echoes that, with the emotional load of justifying one's own existence daily, tire them out

and make actual novel thinking difficult.

We contend that study associations can function as playgrounds that help in diversifying the communicative tools that enable student's full involvement as members of the philosophy community. We are in a privileged position to form a strong student community that facilitates encounters between thinkers at various intersections of identity, while also serving as a bridge for productive conversations between tutees and

faculty. By committing to the idea of student input as epistemically valuable, our

association provides constant feedback to our pedagogues, promotes socio-academic involvement at our varied events such as cinema events, creative writing sessions, podcasts, student-led courses and conferences that work in tandem to fill the gaps in themes and approaches that our curricula naturally have with.

Zooming in, the project "Affect at the Intersection", developed in collaboration with

professor Jingjing Li, nudges our peers to attend to the dimension of emotion underlying embodied experience in order to think critically through the socio-political dimensions of their lives. Analyzing the particular anxieties facing this generation, we aim to reveal the ways in which a vast diversity of possible thinking is normalized by structures that are determined to induce nothing but nihilism in their participants. Facing these difficult and complicated emotions with an intersectional philosopher's toolkit that closely attends to the treasure trove that is lived experience, we work to afford both ourselves and our peers the space to creatively rethink oppressed identities and hope so loudly that we regain the strength to affirm our lives against political hopelessness.

ID: 232 / **Panel 2-1-A: 1**
Symposium paper
Topics: Intercultural philosophy, None of the topics above (please include your topic under 'keywords'), Critical Theory
Keywords: Diversifying Philosophy

Site for unlearning: philosophy department

Katrine Smiet
Radboud University

The standard philosophical canon and curriculum is marked by eurocentrism, ethnocentrism and androcentrism. Increasingly, teachers and students call for a more diverse curriculum and more inclusive teaching. However, in order to realize that, simply introducing some new/different texts into existing syllabi is not enough. Instead, drawing on the work of postcolonial philosopher Gayatri Spivak (1990, 1999), we posit that a process of 'unlearning' is necessary: to question the epistemic norms and disciplinary conventions that underpin this curriculum, and to cultivate receptivity for different perspectives and traditions of thought. The project 'Site for Unlearning: Philosophy Department' at Radboud University takes on this challenge. A collective consisting of students, teachers and education support staff form the motor for a collective exploration of the potential and challenges of unlearning. The collective will set out on a collaborative inquiry, asking:

1. (Why) is unlearning relevant for philosophy and philosophy education?

2. What do we need or want to unlearn in order to realize more diverse and inclusive education in philosophy?

3. How can we (start to) unlearn that?

In this action research project, action and reflection go hand in hand. The collective will identify 'sites for unlearning', and propose concrete interventions and changes, as well as reflecting on the ups and down of the process, the resistances encountered and the challenges it poses.

ID: 236 / **Panel 3-2-A: 1**
Symposium paper
Topics: Intercultural philosophy, None of the topics above (please include your topic under 'keywords'), Critical Theory
Keywords: Diversifying Philosophy

Diversifying Philosophy – wrestling with the canon in Humanistic Studies

Caroline Suransky
University of Humanistics

A canon presupposes and encapsulates a theory of history, a moral theory, a conception of society and assumptions around exemplary science (Burawoy, 2021). Furthermore, a canon is foundational – as it grounds the production and sharing of knowledge; it is relational as the selected concepts, people and their ideas do not stand in isolation but speak to each other and it is historical, as it is challenged and changed over time (Burawoy, 2021).

The University for Humanistic Studies profiles itself as an inter- or transdisciplinary university that is inspired by the worldviews and traditions of humanism. It focusses on contemporary issues with a normative component towards developing humane and caring societies. Rather than orienting itself to a singular disciplinary canon, it wrestles with diverse canons and is challenged to figure out what it concretely means when it seeks inspiration from humanist worldviews and traditions.

Currently, two such challenges stand out. Firstly, the challenge of anthropocentrism and the development of ecohumanism in the Anthropocene. Flowing from the era of Enlightenment-Modernity, independence and autonomy became important basic values of humanism. However, the contemporary challenges posed by the global ecological crisis are inextricably connected with fundamental questions about being human in relation the earth. This requires us to reexamine our foundational humanistic philosophies as basis for how we think and act. It requires a paradigm shift in which the ideal of the autonomous human being makes way for a paradigm of connectedness in the context of complex ecological processes.

Secondly, the University of Humanistic Studies is still a very white institution, almost all academic staff and students are white. While not denying diversity within white communities, we may conclude that Western humanism, flowing from Modernity has been rather blind to the composite notion of modernity-coloniality which refers to the way in which the two concepts are inseparable (Quijano, 2007, Mignolo, 2018). Enlightenment thinking became a normative ambition when modern Europe presented itself as the antithesis of the rest of the world, which they qualified as pre-modern, 'primitive' and 'barbaric'. While the West advocated universal equality and freedom for 'humanity', it showed a very different face in the colonies. How to address this legacy is an important question while Humanistic Studies wrestles with competing canons and paradigms.

And so we search new interlocutors with whom the 'humanistic canon' could be critically reexamined in the light of the criteria, theories and orientations to the canon as foregrounded by Burawoy.

ID: 233 / Panel 2-2-A: 1
Symposium paper
Topics: Intercultural philosophy, None of the topics above (please include your topic under 'keywords'), Critical Theory
Keywords: Diversifying Philosophy

Disrupting the philosophical canon in the mode of playful earnestness

Ruth Rebecca Tietjen

Tilburg University

In my contribution, I defend the idea that we should disrupt the philosophical canon in the mode of playful earnestness. I do so by proceeding in three steps. First, I offer a critical reflection on what a or the philosophical canon is, thereby drawing on the contemporary debate about "the" philosophical canon and its critique (e.g., Alcoff 2017; Dotson 2012; Rée 2002; Waithe 2015; Westphal 1993). Particularly, I explore how, as members of institutionalized academic philosophy, we undergo processes of subjectification that subject us to, make us adopt, and reify certain (narrow and oppressive) ideas of what philosophy is, how it should be done and presented. It is these implicit rules and norms that constitute what I take to be "the philosophical canon." These ideas in turn culminate in our image of "the philosopher." Second, I explore the question of who and what gets excluded in this process (specific genres, topics, groups of people) and show how this exclusion is aesthetically, epistemically, and politically problematic. For instance, there is hardly room for surprise when already beforehand we clearly have to/want to state what will happen in our talk, our lecture, our paper. This is incompatible with the idea that philosophy is and should be an emancipatory and dialogical enterprise – and activity or event rather than the result of an activity. Third, in the practical spirit of this symposium, I offer some insights into my own attempts to disrupt my and our habitualized forms of philosophical thinking and acting (writing, speaking, learning, and teaching) with performative and other creative tools. I argue that playful earnestness is the right mode for this kind of enterprise because, on the one hand, we need to take the task to change ourselves, (academic) philosophy, and our society at large existentially and politically seriously and, on the other, we need freedom and creativity (i.e., playfulness) for our liberatory practice to flourish. (If you are sensing a tension between what I say and how I say what I say, you are on the right track; would a conference abstract be accepted that consists of nothing but an image?)

References

Alcoff, Linda Martín (2017). Philosophy and Philosophical Practice: Eurocentrism as an Epistemology of Ignorance, The Routledge Handbook of Epistemic Injustice, ed. by James Kidd, José Medina, Gaile & Pohlhaus, Routledge, 397–408.

Dotson, Kristie (2012). How is this paper philosophy? Comparative Philosophy 3 (1), 3–29.

Rée, Jonathan (2002). Women philosophers and the canon, British Journal for the History of Philosophy, 10 (4), 641-652.

Waithe, Mary Ellen (2015), From Canon Fodder to Canon-Formation: How do we get there from here?, The Monist 98 (1), 21–33.

Westphal, Merold (1993). The Canon as Flexible, Normative Fact, The Monist 76 (4), 436–449.

ID: 164 / **Panel 1-1-G: 1**
Individual paper
Topics: Intercultural philosophy, Ethics (general and applied), Metaethics, Moral Psychology, History of philosophy (from ancient to modern)
Keywords: Buddhism, Santideva, Virtue, Patience

Śāntideva and the Virtue of Patience (kṣānti)

Stephen Harris

This talk draws on the characterization of the virtue of patience (kṣānti) by the 8th century CE Mahayana philosopher, Śāntideva, to argue for the cross-cultural applicability of the category of virtue theory to Buddhist ethical thought. I begin by responding to two objections to using the concept of virtue to engage with Buddhist thinkers. First, I consider the concern that Buddhist metaphysical commitments to selflessness and radical impermanence entail they cannot be understood as developing a rich account of virtuous character. In response, I argue that the Abhidharma Buddhist account of reliably repeating casually connected mental states provides a sufficient metaphysical basis for theorizing conventionally existing virtuous habitual dispositions, for authors such as Śāntideva. Second, I consider the objection that Śāntideva's thought is best understood as a consequentialism, committed to the impersonal maximization of happiness. Such characterizations, however, are controversial, and risk superimposing a foreign conceptual framework on Śāntideva's thought.

Instead, I argue that we can engage more naturally with Buddhist moral philosophers by using the broader, more inclusive category of virtue theory, as characterized by Julia Driver. For Driver, a virtue theory is any systematic account of the virtues and their role in moral life. Significantly, any moral theory may develop a virtue theory; for instance, a universal consequentialist might hold that virtues are those qualities that help the agent maximize good consequences, and a deontologist can theorize the importance of virtue in performing one's duty or following the relevant rules. Using this concept, therefore, allows us to consider the insights of Buddhist thinkers like Śāntideva without situating them in relation to ongoing debates as to whether consequentialists, deontologists or virtue ethicists provide the most adequate theory of right action.

In illustrating these points, the talk takes as its case study Śāntideva's understanding of the virtue of patience, the disposition to remain mentally tranquil in times of difficulty. I show how Śāntideva develops patience to benefit its possessor, by eliminating anger, which he argues always damages its possessor, as well as by enabling the virtuous person to endure any amount of physical pain without mental suffering. The refinement of desire is also central to Śāntideva's account of patience, given that he claims that anger can only be fully overcome when selfish desires have been transformed into compassion for all beings.

ID: 212 / **Panel 4-1-H: 2**
Individual paper
Topics: Ethics (general and applied), Metaethics, Moral Psychology
Keywords: Cryonics, Life-extension, Death, Mourning, Bioethics

Should We Warm Up to the Idea of Being Frozen? Reconsidering the Permissibility and Desirability of Cryonics

Robin Hillenbrink

Cryonics is a procedure that uses the extremely low temperature of liquid nitrogen to pause all biological processes in the body. The cooling and storing of a human body is performed in the hope of future resurrection. As of today, approximately 500 patients have been cryopreserved. Additionally, around 5000 people worldwide, from all continents, are currently signed up to be cryopreserved in the future. Overall, cryonics is growing into a substantial venture that an increasing amount of people is opting for as a chance to a 'second' life. Members and patients of these organizations – 'cryonicists' – believe that true irreversible death does not occur simultaneously with legal death, and consider cryonics as an extreme potential form of life-extension therapy. In the wake of this relatively new technology, philosophers have considered its present and future implications, and its moral permissibility and desirability. Most objections to cryonics have either claimed that the procedure and its possible outcomes are undesirable or worrisome for the patient themselves, or immoral towards others in society at large. These arguments are, inter alia, considered in David Shaw's "Cryoethics: Seeking Life After Death". In this paper, after careful consideration of arguments against cryonics, Shaw nevertheless arrives at the conclusion that the "Cryonic Wager" is overwhelmingly attractive for the 'rationalist humanist'. That is, cryonicists have not much to lose and plenty to gain. However, philosophical literature on this topic is limited. For this reason, I will argue that his list of practical and ethical objections is incomplete. In addition, I will claim that, in literature on cryonics, the wellbeing of one group of stakeholders is especially underrepresented: surviving loved ones. Inter-personal objections considering this group of people are connected to the human mourning process. Unlike practical and ethical arguments, inter-personal objections are relevant regardless of the success or failure of cryonics.

This paper is structured as follows: First, I present a short introduction to the practice of cryonics, and the main arguments presented in debates examining its moral permissibility and desirability. Second, I argue that in the ethical 'waste of resources' objection, the time and money spent educating doctors that will later dedicate their time and research to cryonics must be considered when considering the distribution of resources too. Third, I consider the conceivable undesirable practical and moral consequences of potential failure in the future thawing process. Fourth, I argue that, while cryopreservation could be desirable for the patient per se, it might not be desirable to them regarding their loved ones, since the mourning process for a person that cannot be cremated or buried, nor can be considered fully deceased but merely 'paused from life', might be disrupted entirely. To illustrate this objection, I examine a recent case study, presented in the documentary Hope Frozen. Finally, I conclude that, while cryonics is not morally impermissible, it is not desirable for a patient because it is potentially harmful to their loved ones. Thus, there is more to lose in the Cryonic Wager than is presented in current literature.

ID: 132 / Panel 1-1-B: 1
Individual paper
Topics: Aesthetics and philosophy of art, Ethics (general and applied), Metaethics, Moral Psychology, History of philosophy (from ancient to modern)
Keywords: music, sexuality, morality, Søren Kierkegaard, Leo Tolstoy

Ethics and aesthetics in Kierkegaard and Tolstoy: music, sexuality, and religious morality

Sophie Höfer

This paper seeks to explore the question whether aesthetic phenomena can have an effect on ethical judgements; specifically, whether music can be a source of moral corruption. It does so by drawing on Søren Kierkegaard's pseudonymous essay 'The Immediate Erotic Stages or The Musical–Erotic' from Either/Or, and Leo Tolstoy's novella The Kreutzer Sonata. The aim is to draw attention to some of the thematic overlap in Kierkegaard's and Tolstoy's philosophies of music – most notably around the topics of sexuality, morality, and religiosity – as well as to add their views to ongoing debates around the moral dimension of music. In 'The Musical–Erotic', Kierkegaard's pseudonym 'A' offers an aesthetic theory that presents music as the most abstract of all art forms. For a perfect unity, what should be expressed in music should be the most abstract of all ideas. From this arises the idea of the sensuous-erotic as the absolute theme of music. Ultimately, 'The Musical-Erotic' suggests that music is erotic-demonic in character and cannot express moral or religious ideas. Similarly, Tolstoy's The Kreutzer Sonata paints a grim picture of music by depicting it as a source of sexual immorality and moral corruption to the point of murder. In the novella, musical performance essentially functions as a socially acceptable form of erotic encounter, and reveals the hidden immoral potential in those involved with it.

Thus, both works juxtapose the aesthetically beautiful with the morally good, and discuss music as an erotic force related to immoral ideas and behaviour. This understanding causes tension with the Christian backgrounds of both Kierkegaard and Tolstoy. At first glance, music as the erotic seems to be a purely aesthetic activity and, as such, neither suited for a reflective ethical life, nor a religious existence. However, both authors ultimately seem to agree that music can and should go beyond its erotic-demonic elements. Even though music in the way it is discussed in 'The Musical–Erotic' and The Kreutzer Sonata seems to be a possible threat to both Kierkegaard's and Tolstoy's religious-ethical project, further writings by the two authors, specifically Kierkegaard's journal entries and Tolstoy's book What is Art?, suggest that music is ultimately reconcilable with what they consider a meaningful human existence. For Kierkegaard, music seems to problematic only if we assign to it a more important role in our lives than it should have, and thus deify it. In Tolstoy's thought, the immoral potential of music can be countered if a piece is composed and performed in the spirit of Christian ethics, that is, love.

ID: 108 / **Panel 5-2-E: 1**
Individual paper
Topics: History of philosophy (from ancient to modern), Philosophy of Language
Keywords: Early modern philosophy, computational history, scientometrics, semantic strategies

Early modern scientometrics? Extracting semantically disruptive works in natural philosophy.

Hugo Dirk Hogenbirk

Scientometrics is the quantitative study of the scientific enterprise. By analyzing the meta-data that has been made readily available in recent decades, structural features of the scientific enterprise can be found and investigated. For example, one can use citation-analysis in order to find works that are particularly disruptive. By finding works that 'replace' particularly many of the works it itself has cited we find works that disrupt the cumulative status quo of the scientific enterprise (Funk & Owen-Smith, "A dynamic network measure of technological change", 2017). This can be levied for both the investigation of particular works' amount of disruptiveness or for seeing general movements in the overall disruptiveness of corpora, or even science as a whole (Park et. al., "Papers and patents are becoming less disruptive over time", 2023).

But, what if the historical material one is interested in does not a) have citation practices needed for this sort of research and b) for the citations that are there, there is no centrally organized database? In this presentation I propose using algorithms that model the semantic contents of terms in specific books to develop a measure of 'semantic disruption' that mimics the disruption measures of scientometrics. By considering whether works within a corpus are particularly semantically continuous with their past and future, I extract those works that are particularly discontinuous with their past and continuous with their future.

This measure is applied to a corpus of early modern natural philosophers. This corpus has been constructed in the context of the [reference redacted for peer review] project. By considering words that are particularly central to natural philosophy, I extract the most disruptive works from this corpus and investigate for different words whether they are generally topics of disruption and innovation. A number of interesting results will be presented: 1. There is a correlation between authors that are deemed canonical by modern annotators and works in the corpus that are particularly disruptive 2. Newtonian authors have a strategy of conceptual development that allows them to disrupt certain terms but remain stable on most of their vocabulary 3. The "popularizing" Newtonian authors like Petrus van Musschenbroek and Willem 's Gravesande are heavily represented in the top of the disruptive works (DuCheyne, "s Gravesande's Appropriation of Newton's Natural Philosophy" & "Petrus van Musschenbroek and Newton's 'vera stabilisque Philosophandi methodus", 2014/15). Point 2 and 3 suggest that the strategy of conceptual development of the Newtonians exhibits "façade-like" behavior as introduced by Mark Wilson (Wilson, Wandering Significane, 2006). Point 1 might be of interest for the wider debate of canon reconstruction.

ID: 157 / **Panel 2-1-G: 1**
Individual paper
Topics: Philosophy of Science
Keywords: scientific understanding, biomedical sciences

Making X-ray images intelligible for understanding pulmonary tuberculosis: A community achievement

Linda Holland

In this paper, I will discuss how scientists made X-ray images intelligible for understanding pulmonary tuberculosis after X-ray imaging was introduced in the late 1890s as a new technique to assist in diagnosing the disease. I will use this case to study scientific understanding in the biomedical sciences. Insights from the case study will be compared with Sabina Leonelli's account (2009) of scientific understanding of biological phenomena.

According to Leonelli, understanding is a cognitive achievement of individual scientists. It can be realized by the coordination of both theoretical skills and embodied skills, allowing the scientist to reason about a phenomenon and/or materially intervene on a phenomenon. Still, in order to turn the understanding of individual scientists into scientific understanding, the understanding needs to be shared with other scientists. This means that scientists need to participate in scientific communities, which also requires social skills. These theoretical, embodied and social skills are partly articulate and partly tacit (i.e., inarticulate), and can only be learned by participation in scientific communities.

In my case study, scientists did already have some understanding of the phenomenon pulmonary tuberculosis, but they tried to increase this understanding by adding a new technique to the diagnostic armamentarium. Yet, it was not immediately clear to scientists what the shadows seen in the X-ray images they produced were shadows of. This shows that X-ray images had to be made intelligible for diagnosing pulmonary tuberculosis and increasing understanding of the phenomenon in this way. Scientists achieved to do this by relating the X-ray images to knowledge obtained from other experimental sources that were generally accepted by the scientific community. Part of this knowledge was already established, and part needed to be obtained specifically for this purpose. Relating the X-ray images to the generally accepted knowledge of the scientific community required skills that were partly tacit and partly articulate. The resulting conceptualization of the X-ray shadows allowed scientists to reach a consensus on which shadows were indicative of pulmonary tuberculosis. Thus, scientists determined together what shadows were signs of pulmonary tuberculosis, at the same time changing the understanding of pulmonary tuberculosis within the community.

My discussion of the case study agrees with Leonelli's view that understanding involves both tacit and articulate skills. However, while Leonelli argues that scientific understanding is an achievement of individual scientists, in the case study it becomes clear that this is an achievement of scientific communities. Scientists determined together what made X-ray images intelligible for understanding pulmonary tuberculosis. Moreover, on Leonelli's view, understanding is ascribed to scientists in a binary way: a scientist either has understanding or not. Defining understanding as a community achievement allows to meet an important aim of philosophers working on scientific understanding: making sense of degrees of understanding (i.e., ascribing comparatively more or less understanding to scientists). By defining the standard of maximal understanding at a community level, the degree of understanding of (groups of) scientists within the community can be evaluated in relation to the maximal understanding in the community.

ID: 146
Full symposium
Topics: Ethics (general and applied), Metaethics, Moral Psychology
Keywords: ethics of technology; technomoral change; value change; socially disruptive technologies; moral progress;

Varieties of technomoral change

Chair(s): **Jeroen Hopster** (Utrecht University, Netherlands, The)

This symposium brings together research on value change and moral progress in relation to technology. For many years, research in the Netherlands has been at the vanguard of the field of technology ethics. A recent focus of research in this field, especially at Dutch universities, have been theories of (techno)moral change. Research undertaken in the ERC-project "Value Change" (TU Delft), the ERC-project "Progress" (Utrecht University), as well as the interuniversity NWO Gravitation-project "Ethics of Socially Disruptive Technologies", has given impetus to various novel proposals about types of value change, mechanisms technomoral change, and the normative and meta-ethical implications of such change. The aim of this symposium is to foster knowledge exchange at the intersection of these closely related topics. Through presentations and a panel discussion, the symposium aspires to yield a mapping of different kinds and modes of technomoral change, a comparison of methodologies / approaches for studying technomoral change, and an explication of how the study of value change relates to the study of moral progress.

The symposium consists of 4 contributions: 3 presentations and a panel discussion (abstracts and panel listed below). Requested duration of the symposium is 150 minutes (30 minutes for each of the presentations and 60 minutes for the panel). We have a strong preference for the symposium to take place on Friday June 2nd (the organiser can only be present this day).

Presentations of the Symposium

A pluralistic model of value change and technology

Philip Nickel

Eindhoven University of Technology

This paper develops a pluralistic model of value change, with particular attention to how technology might bring about such change. The pluralistic model contrasts with two other models or pictures that remain largely implicit in current scholarship: a standard model according to which value change consists in a change within most individual members of a given society from one dominant value scheme to another; and a bypass model according to which we do not need an account of value change at an individual level in order to explain its causes at the social level. Consider the transition from a value scheme (VS1) including chastity and fidelity to another (VS2) including values of sexual freedom and expression. Within a standard process of value change, both internal forces (the need for practical consistency, rejection of oppressive systems) and external forces (the role of women in wartime production, the birth control pill, religion's loss of authority) might be said to cause the dominant values in a given society to shift from VS1 to VS2. The standard model understands this in terms of value replacement in individuals across society: the majority of individuals give up old values and adopt new ones. The bypass model, focusing on external forces, takes no particular view of what happens in individuals when societal values change.

I argue that by adding emergent and differential value change to the standard model, we stand to gain explanatory power in two ways: first, the heterogeneity of value change at the individual level helps to differentiate distinctive population-level disruptive effects involving disagreement and collective uncertainty; and second, individual level change is worth explaining in its own right because of the individual-level disruptions that are sometimes involved. Emergent value change is defined as a situation where a substantial group of individuals within society adopts values specific to a new activity or context. These values do not displace values that the relevant individuals possessed earlier. Differential value change is where two or more groups of individuals within society undergo value change in ways that differ substantially from one another, due to factors affecting groups differentially.

In the paper, I set out evidence that the standard and bypass models guide current philosophical thought about value change in English-language literature, even if they are not fully explicit. I then give empirical and theoretical reasons for distinguishing these other kinds of value change from the standard variety. Finally, I argue that varieties of value change at the individual level are interesting and important for their own sake, but also for being associated with distinctive patterns of social disruption, controversy, and uncertainty, particularly in relation to novel technologies.

Mechanisms of technomoral change: a new taxonomy

Jeroen Hopster
Utrecht University

In the philosophy of technology there are various theories that study the interactions between morality and technology, albeit from somewhat different theoretical angels. This body of scholarship – which I collectively refer to as work on "technomoral change" – has yielded important insights on the mechanisms by which emerging technologies transform social morality. However, the diversity of theoretical frameworks employed, and the multiplicity of mechanisms of change that scholars have identified, has obscured an integrative understanding. Such an understanding can benefit the philosophy of technology, as well as practical philosophy more generally, as many practical philosophers frequently engage with normative issues in which emerging technologies play a salient role. Accordingly, in this presentation I aim to make recent findings on technomoral change accessible to a larger audience of practical philosophers, and to develop a taxonomic understanding of the mechanisms of technomoral change. I do so by providing an overview of key mechanisms that have been outlined in the recent literature. Subsequently, I articulate a taxonomy that incorporates these findings, but also goes beyond them, as it outlines levels of technomoral mediation which have been overlooked in previous work.

My paper starts with providing an overview of mechanisms of technomoral change that have been identified in the recent literature (esp. Swierstra 2013; Hopster et al. 2022; Danaher and Sætra 2022, 2023). Departing from the work of Danaher and Sætra 2023, which arguably offers the most comprehensive taxonomy to date, I go on to discuss shortcomings of existing typologies, and propose how these can be amended. I argue that two levels at which technology mediates morality are not properly accounted for in current work: conceptual mediation and epistemic mediation. Technology mediates morality through concepts: for instance, some concepts and conceptions acquire new moral significance in response to technological changes (e.g. explainability), whereas other concepts lose such significance (e.g. chastity). Technology also mediates morality at an epistemic level, in particular at an informational level: emerging technologies enable new ways of knowledge acquisition, and change relationships in social access to knowledge. I argue, versus Danaher and Sætra, that a sufficiently comprehensive typology of technomoral change should explicitly incorporate these two types of technomoral mediation. I close by discussing the explanatory and predictive value of the resulting taxonomy, and by addressing some potential misgivings, such as the idea that the taxonomy implies a deterministic view of how technology transforms society.

Key-words: Techno-moral change, morally disruptive technologies, ethics of technology, taxonomy, mechanism

Literature:

Danaher, J., & Sætra, H. S. (2022). Technology and moral change: the transformation of truth and trust. Ethics and Information Technology, 24(3). https://doi.org/10.1007/s10676-022-09661-y

Danaher & Saetra (2023). The Mechanisms of Techno-moral Change: A Taxonomy and Overview. Under review.

Hopster, J. K. G., Arora, C., Blunden, C., Eriksen, C., Frank, L. E., Hermann, J. S., Klenk, M. B. O. T., O'Neill, E. R. H., & Steinert, S. (2022). Pistols, pills, pork and ploughs: the structure of technomoral revolutions. Inquiry (United Kingdom), 1–33. https://doi.org/10.1080/0020174X.2022.2090434

Swierstra, T. (2013). Nanotechnology and technomoral change. Etica e Politica, 15(1), 200–219.

How Technologies Change Morality by Changing our Concepts and how we can Achieve Moral Progress

Guido Löhr
Eindhoven University of Technology

New technologies are currently disrupting not just our beliefs and practices. They are disrupting and arguably even changing our most basic moral concepts like friendship, personhood, rights, or responsibility (Hopster, 2021). What does it mean for conceptual disruption to change our morality and how can technologies be used and regulated so that they generate moral progress as opposed to regress? I will first present a novel account of robust moral progress that significantly differs from current approaches (e.g., Hermann, 2017; Buchanan & Powell, 2018; Eriksen, 2020; Sauer, et al., 2021; Hopster, et al., 2022). Second, I will explain in detail how techno-conceptual moral changes can contribute to a progressive agenda. I will build on relational ethics literature (Darwall, 2013; Wallace, 2019) as well as my work on conceptual

disruptions and socially disruptive technologies (Löhr, 2023). Finally, I propose a general method for doing normative ethics of concepts in light of disruptive technologies.

Literature:

Buchanan, A., & Powell, R. (2018). The evolution of moral progress: A biocultural theory. Oxford University Press.

Darwall, S. (2013). Morality, authority, and law: Essays in second-personal ethics I. OUP Oxford.

Eriksen, C. (2020). Moral Change. Springer International Publishing.

Hermann, J. (2017). Possibilities of moral progress in the face of evolution. Ethical Theory and Moral Practice, 20(1), 39–54.

Hopster, J. K. G., Arora, C., Blunden, C., Eriksen, C., Frank, L. E., Hermann, J. S., ... & Steinert, S. (2022). Pistols, pills, pork and ploughs: the structure of technomoral revolutions. Inquiry, 1-33.

Hopster, J. (2021). What are socially disruptive technologies?. Technology in Society, 67, 101750.

Löhr, G. (2023). Do socially disruptive technologies really change our concepts or just our conceptions? Technology in Society. Volume 72, 102160.

Sauer, H., Blunden, C., Eriksen, C., & Rehren, P. (2021). Moral progress: Recent developments. Philosophy Compass, 16(10), e12769.

Wallace, R. J. (2019). The moral nexus (Vol. 8). Princeton University Press.

Panel: building bridges between "progress", "value change", and "ethics of socially disruptive technologies"

Jeroen Hopster[1], **Ibo van de Poel**[2], **Charlie Blunden**[1], **Philip Nickel**[3], **Guido Löhr**[3]
[1]Utrecht University, [2]Delft University of Technology, [3]Eindhoven University of Technology

Contributors: [Anonymised. Panel consists of one full professor, one assoc. professor, one postdoctoral researcher and one PhD-candidate, representing the different research projects]

Moderation: [Anonymised]

Contents:

This panel brings together representatives of three major research programmes in the Netherlands on (techno)moral change and progress: the ERC-project PROGRESS, the ERC-project VALUE CHANGE, and the NWO Gravitation-project ESDIT. The panel addresses questions at the intersection of these projects and seeks to identify points of convergence and dispute.

Discussion questions include the following:

• What are the key intellectual traditions / theoretical paradigms that inform the research of your respective projects?

• What is the state of the art of the research in your project? What have been core findings so far?

• What methods and approaches do you employ, or would propose, to study value change, social disruption, and moral progress?

• How do the notions of moral progress, value change, and social disruption relate to each other?

• What do you regard as key mechanisms of value change? How should the notion of "mechanism" be understood, in this context?

• How to make normative assessments of moral change? Specifically, how to make normative assessments when the change in question is radical (e.g., when it is a moral revolution)?

• What can the different projects learn from each other? What are relevant areas of knowledge exchange?

• What are specific research questions that you would like to see addressed in the near future?

ID: 229 / **Panel 3-1-H: 1**
Symposium paper
Topics: Ethics (general and applied), Metaethics, Moral Psychology
Keywords: ethics of technology; technomoral change; value change; socially disruptive technologies; moral progress;

Mechanisms of technomoral change: a new taxonomy

Jeroen Hopster
Utrecht University

In the philosophy of technology there are various theories that study the interactions between morality and technology, albeit from somewhat different theoretical angels. This body of scholarship – which I collectively refer to as work on "technomoral change" – has yielded important insights on the mechanisms by which emerging technologies transform social morality. However, the diversity of theoretical frameworks employed, and the multiplicity of mechanisms of change that scholars have identified, has obscured an integrative understanding. Such an understanding can benefit the philosophy of technology, as well as practical philosophy more generally, as many practical philosophers frequently engage with normative issues in which emerging technologies play a salient role. Accordingly, in this presentation I aim to make recent findings on technomoral change accessible to a larger audience of practical philosophers, and to develop a taxonomic understanding of the mechanisms of technomoral change. I do so by providing an overview of key mechanisms that have been outlined in the recent literature. Subsequently, I articulate a taxonomy that incorporates these findings, but also goes beyond them, as it outlines levels of technomoral mediation which have been overlooked in previous work.

My paper starts with providing an overview of mechanisms of technomoral change that have been identified in the recent literature (esp. Swierstra 2013; Hopster et al. 2022; Danaher and Sætra 2022, 2023). Departing from the work of Danaher and Sætra 2023, which arguably offers the most comprehensive taxonomy to date, I go on to discuss shortcomings of existing typologies, and propose how these can be amended. I argue that two levels at which technology mediates morality are not properly accounted for in current work: conceptual mediation and epistemic mediation. Technology mediates morality through concepts: for instance, some concepts and conceptions acquire new moral significance in response to technological changes (e.g. explainability), whereas other concepts lose such significance (e.g. chastity). Technology also mediates morality at an epistemic level, in particular at an informational level: emerging technologies enable new ways of knowledge acquisition, and change relationships in social access to knowledge. I argue, versus Danaher and Sætra, that a sufficiently comprehensive typology of technomoral change should explicitly incorporate these two types of technomoral mediation. I close by discussing the explanatory and predictive value of the resulting taxonomy, and by addressing some potential misgivings, such as the idea that the taxonomy implies a deterministic view of how technology transforms society.

Key-words: Techno-moral change, morally disruptive technologies, ethics of technology, taxonomy, mechanism

Literature:

Danaher, J., & Sætra, H. S. (2022). Technology and moral change: the transformation of truth and trust. Ethics and Information Technology, 24(3). https://doi.org/10.1007/s10676-022-09661-y

Danaher & Saetra (2023). The Mechanisms of Techno-moral Change: A Taxonomy and Overview. Under review.

Hopster, J. K. G., Arora, C., Blunden, C., Eriksen, C., Frank, L. E., Hermann, J. S., Klenk, M. B. O. T., O'Neill, E. R. H., & Steinert, S. (2022). Pistols, pills, pork and ploughs: the structure of technomoral revolutions. Inquiry (United Kingdom), 1–33. https://doi.org/10.1080/0020174X.2022.2090434

Swierstra, T. (2013). Nanotechnology and technomoral change. Etica e Politica, 15(1), 200–219.

ID: 230 / **Panel 3-2-H: 1**
Symposium paper
Topics: Ethics (general and applied), Metaethics, Moral Psychology
Keywords: ethics of technology; technomoral change; value change; socially disruptive technologies; moral progress;

How Technologies Change Morality by Changing our Concepts and how we can Achieve Moral Progress

Guido Löhr

Eindhoven University of Technology

New technologies are currently disrupting not just our beliefs and practices. They are disrupting and arguably even changing our most basic moral concepts like friendship, personhood, rights, or responsibility (Hopster, 2021). What does it mean for conceptual disruption to change our morality and how can technologies be used and regulated so that they generate moral progress as opposed to regress? I will first present a novel account of robust moral progress that significantly differs from current approaches (e.g., Hermann, 2017; Buchanan & Powell, 2018; Eriksen, 2020; Sauer, et al., 2021; Hopster, et al., 2022). Second, I will explain in detail how techno-conceptual moral changes can contribute to a progressive agenda. I will build on relational ethics literature (Darwall, 2013; Wallace, 2019) as well as my work on conceptual disruptions and socially disruptive technologies (Löhr, 2023). Finally, I propose a general method for doing normative ethics of concepts in light of disruptive technologies.

Literature:

Buchanan, A., & Powell, R. (2018). The evolution of moral progress: A biocultural theory. Oxford University Press.

Darwall, S. (2013). Morality, authority, and law: Essays in second-personal ethics I. OUP Oxford.

Eriksen, C. (2020). Moral Change. Springer International Publishing.

Hermann, J. (2017). Possibilities of moral progress in the face of evolution. Ethical Theory and Moral Practice, 20(1), 39–54.

Hopster, J. K. G., Arora, C., Blunden, C., Eriksen, C., Frank, L. E., Hermann, J. S., ... & Steinert, S. (2022). Pistols, pills, pork and ploughs: the structure of technomoral revolutions. Inquiry, 1-33.

Hopster, J. (2021). What are socially disruptive technologies?. Technology in Society, 67, 101750.

Löhr, G. (2023). Do socially disruptive technologies really change our concepts or just our conceptions? Technology in Society. Volume 72, 102160.

Sauer, H., Blunden, C., Eriksen, C., & Rehren, P. (2021). Moral progress: Recent developments. Philosophy Compass, 16(10), e12769.

Wallace, R. J. (2019). The moral nexus (Vol. 8). Princeton University Press.

ID: 228 / **Panel 3-3-H: 1**
Symposium paper
Topics: Ethics (general and applied), Metaethics, Moral Psychology
Keywords: ethics of technology; technomoral change; value change; socially disruptive technologies; moral progress;

A pluralistic model of value change and technology

Philip Nickel

Eindhoven University of Technology

This paper develops a pluralistic model of value change, with particular attention to how technology might bring about such change. The pluralistic model contrasts with two other models or pictures that remain largely implicit in current scholarship: a standard model according to which value change consists in a change within most individual members of a given society from one dominant value scheme to another; and a bypass model according to which we do not need an account of value change at an individual level in order to explain its causes at the social level. Consider the transition from a value scheme (VS1) including chastity and fidelity to another (VS2) including values of sexual freedom and expression. Within a standard process of value change, both internal forces (the need for practical consistency, rejection of oppressive systems) and external forces (the role of women in wartime production, the birth control pill, religion's loss of authority) might be said to cause the dominant values in a given society to shift from VS1 to VS2. The standard model understands this in terms of value replacement in individuals across society: the majority of individuals give up old values and adopt new ones. The bypass model, focusing on external forces, takes no particular view of what happens in individuals when societal values change.

I argue that by adding emergent and differential value change to the standard model, we stand to gain explanatory power in two ways: first, the heterogeneity of value change at the individual level helps to differentiate distinctive population-level disruptive effects involving disagreement and collective uncertainty; and second, individual level change is worth explaining in its own right because of the individual-level disruptions that are sometimes involved. Emergent value change is defined as a situation where a substantial group of individuals within society adopts values specific to a new activity or context. These values do not displace values that the relevant individuals possessed earlier. Differential value change is where two or more groups of individuals within society undergo value change in ways that differ substantially from one another, due to factors affecting groups differentially.

In the paper, I set out evidence that the standard and bypass models guide current philosophical thought about value change in English-language literature, even if they are not fully explicit. I then give empirical and theoretical reasons for distinguishing these other kinds of value change from the standard variety. Finally, I argue that varieties of value change at the individual level are interesting and important for their own sake, but also for being associated with distinctive patterns of social disruption, controversy, and uncertainty, particularly in relation to novel technologies.

ID: 179 / **Panel 3-3-E: 1**
Individual paper
Topics: Metaphysics, History of philosophy (from ancient to modern), Phenomenology
Keywords: Franz Brentano, Intentionality, Phenomenology, Philosophy of Mind

Intentionality without Objects

Carlo Ierna

Intentionality without objects

Carlo Ierna

There is no such thing like an "intentional object" and we should avoid using the term "object" in discussions about intentionality, because it is extremely misleading. Specifically, the use of terminology like "intentional object" leads to ontologizing whatever it is that it is (mistakenly) applied to, whether this is something immanent or transcendent. In order to argue for this admittedly radical view, I will turn to the philosopher who re-introduced this originally scholastic concept in contemporary debates: Franz Brentano. As it turns out, on closer inspection of his theory of intentionality, we could very well stick to his original account, while abandoning all misleading terminology.

The traditional interpretation of Brentano's theory of intentionality has been distorted by the inflationary use of one single quotation from his 1874 Psychology from an Empirical Standpoint. By considering other sources we can see more clearly what his account really amounted to. In his lectures and letters we find passages that identify the immanent object, the intentional object, and the correlate of an act. These turn out to be merely perspectival distinctions, not separate parts or independent things. The fundamental perspective in Brentano is that of internal perception. In internal perception it is clear that the phenomena I experience, as phenomena, are part of my mental acts, which are then mistaken from another perspective as independent objects. This mistake leads to ontologizing either the transcendent object or the immanent object and taking it as something it is not. All we have in intentional acts is a content, there cannot be any independent objects (in the proper sense of the word) inside or outside of us.

Intentionality then is not a relation between two independent objects or substances, but rather a correlation of a mental act and its dependent, immanent correlate: the content. This leads to the seemingly radical conclusion that there are no mind-independent objects. Transcendently directed acts conceive their contents as external objects, but "objects" are always "objects of an act" and exist only insofar they are part of and wholly dependent on the mind. Without loss of meaning we can use a less misleading expression, such as content, instead. This avoids any talk of "objects" existing either inside or outside the mind and stresses that we do not straightforwardly perceive "objects" at all, but merely interpret bundels of perceptions as "objects".

The conclusion will be that all objects are intentional in the proper sense, i.e. contents of an act, i.e. correlates of an act, i.e. dependent on the mind. They are therefore not at all what is usually understood by "object" - as opposed to "subject" - namely something that exists separately and independently of consciousness. There is no such thing like an "intentional object" and we should avoid using the term "object" in discussions about intentionality, because it is extremely misleading.

ID: 133 / Panel 5-3-C: 1
Individual paper
Topics: Ethics (general and applied), Metaethics, Moral Psychology
Keywords: Early risk factor detection, Prevention, Exposome, Ethics, Scoping Review

Early detection of disease risk factors in the exposome: A scoping review of ethical themes

Sammie Jansen, Bart Kamphorst, Bob Mulder, Irene van Kamp, Peter van den Hazel, Sandra Boekhold, Marcel Verweij

Research into the "human exposome" aims to collect and map large quantities of environmental exposures (both physical and social) and their biological manifestations over time to reveal their combined and cumulative impact on health and disease. By utilizing recent advances in computational and biomedical sciences to study complex, cross-temporal interactions between (clusters of) risk factors, exposome research promises to uncover previously unknown pathways to disease. One major practical aim of such initiatives is to benefit disease prevention strategies by improving the precision and accuracy of detecting early risk factors and designing new methods for identifying people and populations at risk of future disease (Wild, 2012).

However, the increasingly early detection of risk factors, as well as the practice of making predictions of future diseases based on these risk factors, raises ethical concerns (Dawson & Verweij, 2007). As the current literature on this subject is fragmented, and no systematic efforts have been taken to capture the overarching ethical considerations of early detection of disease risk factors, we performed a scoping review following the Arksey & O'Malley (2005) methodological framework. Searches in the Scopus and Embase databases were performed to identify ethics literature on the early detection of disease risk factors, resulting in 52 included articles. From those articles, eight ethical themes were extracted: (1) The (un)reliability and (un)certainty of early detection, including methodological and epistemological issues; Ethical considerations related to (2) autonomy and (3) privacy; Discussions of (4) beneficence and non-maleficence; (5) Possible unintended harms to well-being, including physical, psychological, and social harms; (6) Justice concerns, including discrimination and stigmatization; (7) The normative significance of medicalization and changing conceptions of health and disease; and (8) Individual and societal responsibilities.

We will present a thematic overview of the most salient ethical considerations and provide further reflections on three tensions in particular. The first tension we discuss pertains to the availability of disease risk information. It is sometimes suggested that having such information can help people improve their lifestyles and avoid environmental risk factors. However, multiple authors have expressed doubts about people's capacities to change these health behaviors on their own and caution that risk information might also cause stress and promote risk-increasing behaviors, especially in already vulnerable groups. The second tension we wish to highlight is that empowering people to take control over health-related aspects of their lives (supporting self-determination) can also overburden individuals with unrealistic responsibilities. Moreover, this shift towards individual responsibility for health may direct attention away from environmental and societal risks to ill health and the corresponding responsibilities of governments, institutions, and corporations to address these issues. Lastly, the (preventive) health benefits of detecting a wide variety of early risk factors in healthy people is arguably in tension with the prospect that essentially everyone is turned into patients with various levels of 'to-be-sickness' within their bodies, to be monitored and tested throughout their lifetime. While none of these tensions can be straightforwardly resolved, we conclude with some remarks that may guide future discussions about these implications of exposome research.

ID: 104 / **Panel 4-1-G: 1**
Individual paper
Topics: Ethics (general and applied), Metaethics, Moral Psychology
Keywords: Moral Error Theory – Now What Question – Revisionary Metaethics – Abolitionism – Prudential Value

Revising Revisionary Metaethics

Wouter Floris Kalf

Revisionary metaethical theories determine what we should do with our moral discourse. Error-theoretic versions of such theories are well-known. They answer the moral error theorist's now what question. Should error theorists abolish, keep, or modify their moral discourse? There are also success-theoretic revisionary metaethical theories. These theories tell success theorists whether they should abolish, keep, or modify their moral judgments even though these judgments are not systematically false. I argue that both the extant error-theoretic and success-theoretic revisionary metaethical theories make the same mistake. They only change the form of the new moral judgments whereas they should also consider changing the content of the new moral judgments. This is a mistake because for revisionary metaethicists' most important purposes—maximizing the realization of prudential value for moral error theorists, and maximizing the realization of moral value for moral success theorists—it often pays to change both the form and the content of some moral judgments. For instance, it may pay both to fictionalise moral judgments and to accept that patriotism is morally bad even though our current morality promotes patriotism. I conclude that revisionary metaethicists should consider revising both the form and the content of their new moral judgments.

ID: 168 / **Panel 1-2-D: 1**
Individual paper
Topics: Epistemology, Philosophy of Science, Feminist philosophy
Keywords: Fundamentalism; Genealogy; Conceptual engineering

Fundamentalism as a Contested Concept

Nora Kindermann

Social sciences increasingly study politically contested phenomena, such as extremism, terrorism, and fundamentalism. It is hotly and persistently debated what exactly these phenomena are, how we should conceptualize and theorize about them, and who has a say in these matters. In this presentation, I zoom in on the notion of global fundamentalism and develop a strategy to clarify and deal with these persistent debates. In the literature on fundamentalism, there is a persistent debate on whether or not 'global fundamentalism' is a valid category. We can distinguish two broad camps in research on fundamentalism: (1) academics who defend the concept and theories of global fundamentalism; and (2) academics who are critical of the concept and ways of theorizing fundamentalism, and either want to thoroughly restrict the concept's extension or eliminate it altogether. I will argue that genealogical approaches to these questions are fruitful in improving our grasp of these concepts, and of unquestioned and implicit assumptions that underpin conceptualizations and theories of extremism, terrorism, and fundamentalism (see for example Bötticher 2015; Corbett 2015; Erlenbusch-Anderson 2018; Taylor 2017). In zooming in on fundamentalism, I critically assess how a genealogical approach to the concept of fundamentalism can improve our conceptualizations and theories of the phenomenon.

Genealogical approaches help us locate conceptual practices and the development of theories of fundamentalism in broader socio-political structures. This, in turn, helps to uncover unquestioned and implicit assumptions in our conceptual practices and manners of theorizing about fundamentalism. As an example, genealogies of the concept of fundamentalism have uncovered that the term has been used pejoratively from the beginning (e.g. Desjardins 2017). This has influenced early fundamentalism scholars' approach to the phenomenon, and conceptualization and theorizing thereof. Other scholars who take seriously the historical development of the notion voice the worry that it is biased and Westerncentric (e.g. Corbett 2015, Taylor 2017).

In this paper, I argue that a firm grip on the concept's genealogy helps us uncover unquestioned and implicit assumptions in our concepts and theories (cf. Erlenbusch-Anderson 2018), such as implicit negative (or even pejorative) evaluations, biases and the influence Western centric interests and outlooks. Taking into account the concept's historical development, I first assess the arguments against and criticisms of the notion of global fundamentalism. I then scrutinize (genealogically informed) avenues to develop more fruitful conceptualizations and theories of fundamentalism (cf. Duthil Novaes 2020, Srinivasan 2019).

ID: 165 / Panel 2-1-B: 1
Individual paper
Topics: Social and political philosophy, Ethics (general and applied), Metaethics, Moral Psychology, Feminist philosophy
Keywords: taking responsibility, agency, sense making, complicity, indoctrination

Taking responsibility as agential orientation

Naomi Kloosterboer

In the philosophical literature on blame and responsibility, taking responsibility is seen as an extension of a person's responsibility, as something beyond the things for which a person can strictly be held responsible. This can be conceptualized as, first, an act where an agent adopts an end when someone needs to do so (Calhoun 2019). For instance, someone cleans the litter on their street, not because they ought to and could be blamed for omitting but because someone needs to do it and they elect to take care of it. A second view is that taking responsibility is a way of owning the normative consequences of one's agency (Enoch 2012; Mason 2019; Sliwa forth.). When an agent, say, inadvertently hurts someone, the idea is that she is not liable to blame for what happened, but she can still take responsibility for her (involuntary, unknowing) involvement by offering an apology (see also Wolf 2001).

In this paper, I develop a different account of taking responsibility, in which I bring together approaches in feminist epistemology, the existentialist tradition, and the debate on agential self-knowledge (Kloosterboer 2022; Larmore 2007; Moran 2001). The account focuses on the agent's orientation towards the situation in which they find themselves. The basic idea is that the agent can be active towards a situation through critical sensemaking. I will argue that this requires a shared practice of cultivating openness towards the meaning of the situation, how it relates to values and norms, the relevant lived experience, and different perspectives, especially friction between them (Haslanger 2021; Medina 2011). Taking responsibility in this sense connects to becoming aware of the social construction of one's perspective, challenging preconceived norms and interpretations, and contribute to making sense of a shared political reality.

After developing the account of taking responsibility as orientation, I will address its importance in thinking about responsibility in conditions of complicity (Knowles 2021); and in closed-minded groups such as fundamentalist groups. In general, I show that indoctrination, whether or not self-induced (Ranalli, manuscript), undermines the agent's capacity of taking responsibility as having the right orientation.

ID: 130 / Panel 2-2-H: 1
Individual paper
Topics: Philosophy of Science
Keywords: Explainable AI; Machine Learning Models; Pragmatics; Erotetic inference

Pragmatics for Explainable AI

Daniel Kostic

Machine Learning Algorithms (MLA) are often used in data-driven decision-making without an explanation of how they make those decisions, which is often referred to as the opacity problem of Artificial Intelligence (AI) (Burrell 2016). The growing field of Explainable Artificial Intelligence (XAI) has produced many promising research programs, which mostly address the aspect of the opacity problem stemming from the complexity of MLA models (Doran et al. 2017; Langer et al. 2021).

However, the explanatory opacity of AI is only partly a result of the profound complexity of the MLAs. There is also a pragmatic aspect to it, because on the one hand, different stakeholders have different interests, and on the other, different approaches to XAI suggest different explanatory norms.

In the absence of an account of pragmatics of XAI, that could provide relevance criteria for connecting stakeholders' interests and appropriate explanatory norms in XAI, it seems unintelligible why would one set of explanatory norms be relevant for an MLA rather than the other.

I aim to fill this gap by providing an account of pragmatics of XAI, based on erotetic logic (Hintikka 1981; Wiśniewski 1996). This account regiments the idea that a set of propositions, that encapsulates stakeholders' interests, determines explanatory relevance of an answer to a XAI question just in case that set of propositions erotetically implies the XAI question, e.g.:

Explanandum: a is G rather than G'.

Example: A network of symptoms in one's medical history has a small-world rather than regular topology.

There are some topological properties X and X' such that:

1. a is X (but not X')

Example: One's medical history is a small-world (but not regular) topology.

2. Had a been X' (rather than X), a would have been G' (rather than G).

Example: Had one's medical history been regular (rather than small-world), one would have been diagnosed with cancer (rather than simple inflammation).

WHY-QUESTION: Why is a G (rather than G')?

Example: Why is one diagnosed with simple inflammation rather than with cancer?

In my talk, I will use several different examples of stakeholders' contexts to flash out in much more detail the account of pragmatics of XAI outlined above.

References:

Burrell, Jenna. 2016. "How the Machine 'Thinks': Understanding Opacity in Machine Learning Algorithms." Big Data & Society 3 (1): 205395171562251. https://doi.org/10.1177/2053951715622512.

Doran, Derek, Sarah Schulz, and Tarek R. Besold. 2017. "What Does Explainable AI Really Mean? A New Conceptualization of Perspectives." arXiv. http://arxiv.org/abs/1710.00794.

Hintikka, Jaakko. 1981. "On the Logic of an Interrogative Model of Scientific Inquiry." Synthese 47 (1): 69–83. https://doi.org/10.1007/BF01064266.

Langer, Markus, Daniel Oster, Timo Speith, Holger Hermanns, Lena Kästner, Eva Schmidt, Andreas Sesing, and Kevin Baum. 2021. "What Do We Want from Explainable Artificial Intelligence (XAI)? – A Stakeholder Perspective on XAI and a Conceptual Model Guiding Interdisciplinary XAI Research." Artificial Intelligence 296 (July): 103473. https://doi.org/10.1016/j.artint.2021.103473.

Wiśniewski, Andrzej. 1996. "The Logic of Questions as a Theory of Erotetic Arguments." Synthese 109 (1): 1–25. https://doi.org/10.1007/BF00413820.

ID: 222 / **Panel 3-1-D: 1**
Symposium paper
Topics: Social and political philosophy, Ethics (general and applied), Metaethics, Moral Psychology
Keywords: Rights of Nature, Environmental Ethics, Political Philosophy, Philosophy of Law, Ethics

Rights of Nature: An adequate Response to the environmental Crisis?

Eric Boot
University of Amsterdam

No one disputes the fact that the current environmental crisis requires a response. The subject of this paper is a relatively novel response that is gaining more and more traction, both in public discourse and in law: the Rights of Nature (RoN) approach. Over the past 17 years, an increasing number of countries has legally recognized 'Rights of Nature' (RoN). Article 71 of Ecuador's constitution, for example, states that "Nature, or Pacha Mama… has the right to integral respect for its existence and for the maintenance and regeneration of its life cycles, structure, functions and evolutionary processes."(1) Bolivia has adopted a Law of the Rights of Mother Earth, conferring upon Mother Earth rights to, amongst others, life, water, clean air, and freedom from contamination.(2) New Zealand has recognized the Whanganui River(3) and Mount Taranaki(4) as legal persons with all the rights and duties that accompany such personhood. In total, as of January 2021, there are at least 178 legal provisions recognizing RoN worldwide (these vary from constitutions and laws to court rulings).(5)

In a nutshell, proponents of RoN claim that, by granting legal personhood and fundamental rights to nonhuman nature, these entities (trees, rivers, ecosystems) are recognized to have rights of their own that can be enforced against other legal persons. Nonhuman nature will then no longer be an object to use as we please, but a subject with rights. According to the proponents of RoN, if we only would recognize that natural objects and ecosystems are not things with which we may do as we like, but in fact are subjects of rights that place limits on the scope of permissible actions, then we might have a shot at addressing the environmental crisis in a satisfactory manner. This paper sets out to assess this claim, concluding that RoN is morally problematic, conceptually flawed and inefficacious in practice. Two alternative approaches – human rights and environmental virtue ethics – are discussed that combined could secure fundamental rights for humans regarding the environment, cultivate the appropriate character traits in them, and protect nature in the process.

(1) Constitution of the Republic of Ecuador, October 20, 2008. English translation available online at: https://pdba.georgetown.edu/Constitutions/Ecuador/english08.html.

(2) Ley de Derechos de la Madre Tierra [Law of the Rights of Mother Earth] No. 71, art. 7 (2010) (Bolivia).

(3) Te Awa Tupua Act (Whanganui River Claims Settlement) Act 2017, §§ 14–15 (New Zealand).

(4) Ngā Iwi o Taranaki and the Crown: Record of Understanding for Mount Taranaki, Pouākai and the Kaitake Ranges 2017, § 5 (New Zealand).

(5) Craig M. Kauffman & Pamela L. Martin, The Politics of Rights of Nature: Strategies for Building a More Sustainable Future (MIT Press 2021)

Alex Putzer, Tineke Lambooy, Ronald Jeurissen, and Eunsu Kim. 2022. "Putting the Rights of Nature on the Map. A Quantitative Analysis of Rights of Nature Initiatives across the World." Journal of Maps 18 (1): 89–96.

ID: 120
Full symposium
Topics: Social and political philosophy, Ethics (general and applied), Metaethics, Moral Psychology
Keywords: Rights of Nature, Environmental Ethics, Political Philosophy, Philosophy of Law, Ethics

Rights of Nature: an ongoing Discussion in Political Science and Legal Studies – when will Philosophers join?

Chair(s): **Matthias Kramm** (Wageningen University & Research, Mexico)

After their inclusion in the Ecuadorian constitution in 2008, Rights of Nature have been incorporated in legal frameworks of various other countries such as Aotearoa New Zealand, Bolivia, Canada, Colombia, Mexico, and the United States.(1) In September 2022, the first ecosystem in Europe – the Spanish Mar Menor – was assigned legal personhood.(2)

While these legal frameworks differ from one another in many regards, the common denominator is that Nature or specific ecosystems are empowered in legal terms. For court procedures, it is no longer necessary that claimants prove that their own interests could be harmed by a business proposal or development project. It is sufficient if there is a risk that Nature or specific ecosystems could be negatively affected. In these cases, Nature (via representatives) can initiate legal proceedings.

Up to now, the discussion of Rights of Nature has mostly been led by legal scholars and political scientists. This symposium addresses the opportunities and challenges of Rights of Nature from a philosophical and ethical perspective, while drawing on legal scholarship and political science. As a recent legal phenomenon, Rights of Nature can be related to ongoing debates in legal philosophy, environmental ethics, and philosophy of technology.(3)

Our symposium consists of four presentations (15 minutes) that will be followed by a short period of Q&A (5-10 minutes). After these four presentations, there will also be space for a more informal discussion on the topic in general.

The first presentation deals with the question to what extent Rights of Nature are still connected to the Western liberal rights paradigm and whether this relationship is morally problematic or not. The second presentation discusses the way in which Nature is conceptualized in different Rights of Nature approaches and to what extent Nature might be used as a proxy for other normative goals such as inclusion, resilience, and measurability. The third presentation provides a moral assessment of Rights of Nature and compares them to alternative proposals such as human rights and environmental virtue ethics approaches. The fourth presentation analyses different (ecocentric and anthropocentric) ways of justifying Rights of Nature and compares two different legal models and their political, epistemic, and practical advantages and disadvantages.

With our symposium on this recent legal phenomenon, we aim to provide a philosophical and ethical analysis of the topic, launch a highly necessary debate in the field, and raise awareness for the underlying questions among Dutch and non-Dutch philosophers and political theorists.

(1) Craig M. Kauffman & Pamela L. Martin, The Politics of Rights of Nature: Strategies for Building a More Sustainable Future (MIT Press 2021)

(2) Sam Jones, Endangered Mar Menor lagoon in Spain granted legal status as a person, The Guardian, 21-Sep-2022, https://www.theguardian.com/world/2022/sep/21/endangered-mar-menor-lagoon-in-spain-granted-legal-status-as-a-person

(3) Patrick Baard. 2021. "Fundamental Challenges for Rights of Nature." In Rights of Nature: A Re-Examination, edited by Daniel P. Corrigan and Markku Oksanen, 156–75. London, New York: Routledge.

Visa A. Kurki. 2019. A Theory of Legal Personhood. Oxford: Oxford University Press.

Presentations of the Symposium

The liberal Rights Paradigm and Rights of Nature

Adelheid van Luipen
Maastricht University

In current times, it becomes abundantly clear that humans leave a big mark on the deterioration of our planet, causing many environmental crises, while at the same time our interconnectedness and dependencies on sometimes unpredictable natural phenomena, come to the fore.(1) As some legal theorists argue, one of the most pressing questions philosophers face is to imagine a profound legal response to

climate crises and to rethink laws foundations. Critical legal scholarship, building upon new materialist and post-humanist philosophies, have urged for radical imagination in the legal realm to legally accommodate relations in and to a more-than-human world.(2) A legal path with momentum is Rights of Nature theory (RoN).(3) Yet, RoN have also raised numerous questions and critiques, particularly with regards to their emphasis on rights. A question is whether RoN are a mere stop along the way towards a more radical legal approach when they confirm or merely slightly question the dominant Western liberal rights paradigm.

This article aims to bring together different strands in (mainly) legal scholarship on the relation between the liberal legal rights paradigm and RoN. Two questions will be central. Firstly, on the relation between liberal rights and RoN. To what extent do RoN coincide with the liberal rights paradigm? The second part revolves around the question whether, and if so, why this presumed relation is problematic. In particular, this article focuses on the possible problematic character of this relation in light of the relations a more-than-human world would require. This might lead to insights on RoN's potential to legally accommodate relations in a more-than-human world and/or which (other) roads to explore.

(1) IPCC 2022: Climate Change 2022: Impacts, Adaptation and Vulnerability. Contribution of Working Group II to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change.

(2) See for instance A. Grear, 'Towards new legal futures? In search of renewing foundations', in: A. Grear and E. Grant (ed.), Thought, Law, Rights and Action in the Age of Environmental Crisis, Cheltenham: Edward Elgar Publishing 2015, p. 284-289; E. Jones, 'Posthuman international law and the rights of nature', in: A. Grear e.a. (ed.), Posthuman Legalities: New Materialism and Law Beyond the Human (Journal of Human Rights and the Environment 2021 (12), special issue), Cheltenham: Edward Elgar Publishing, p. 93-94.

(3) We follow the policy of various scholars and organizations, including the United Nations Harmony with Nature Programme, which writes Nature in uppercase to establish it also semantically as a subject (See, for instance, footnote 1 of the UN GA Resolution A/75/266). This policy also emphasizes the contrast to a lower-case, anthropocentric object vision of nature.

Rights of Nature proxies

Alex Putzer
Sant'Anna School of Advanced Studies

By promoting a non-anthropocentric understanding of the environment, the Rights of Nature emphasize the importance of inclusive ethics and justice.(1) Rather than considering nature as an object to be exploited and dominated, it is perceived as a subject, to be respected and cherished, with intrinsic value independent of human interests. A river, for instance, is protected for its own sake, rather than its usefulness for fishing, generating electricity, or as a waterway. While this appears to be a rather straightforward approach, the parts of nature which are specifically included or excluded, remain either un- or ill-defined. The Whanganui river in Aoetearoa New Zealand, for instance, represents one of the most recognizable legal Rights of Nature initiatives. The Whanganui River Settlement Act - passed by the national government in 2017 - however, does not protect the entirety of the river, but only the riverbed.(2) This approach was chosen to avoid the complicated alteration of existing claims to the entity's remaining parts. Notwithstanding the practical reasoning, the theoretical one seems to be missing. In this contribution, I delve into the theories behind why certain initiatives use very specific and others very general approaches to define nature.

For this, I look at the major ways nature is conceptualized and compare it to existing Rights of Nature initiatives. My claim is that these initiatives use natures as proxies for other goals. As such, perceiving the environment in a holistic way represents a proxy for inclusion, focusing on biodiversity emphasizes resilience, and a physical account underlines measurability. The inclusion of indigenous peoples and their philosophies is a proxy for justice and inclusion. As such, this latter proxy seems to be among the visions that are closest to what the overall movement might want to stand for. This, however, does not make it the perfect candidate. Through several additional reflections, it becomes clear that different natures and different proxies lead to different rights holders as well as to different outcomes in specific contexts. An awareness of the differences and especially the contradictions can help to create more targeted as well as more efficient environmental protection efforts.

(1) Christopher Stone. 1972. "Should Trees Have a Standing? Towards Legal Rights for Natural Objects." Southern California Law Review 45: 450–501.

(2) Toni Collins and Shea Esterling. 2019. "Fluid Personality: Indigenous Rights and the "Te Awa Tupua (Whanganui River Claims Settlement) Act 2017" in Aotearoa New Zealand. Melbourne Journal of International Law 20 (1): 197–220.

Rights of Nature: An adequate Response to the environmental Crisis?

Eric Boot

University of Amsterdam

No one disputes the fact that the current environmental crisis requires a response. The subject of this paper is a relatively novel response that is gaining more and more traction, both in public discourse and in law: the Rights of Nature (RoN) approach. Over the past 17 years, an increasing number of countries has legally recognized 'Rights of Nature' (RoN). Article 71 of Ecuador's constitution, for example, states that "Nature, or Pacha Mama… has the right to integral respect for its existence and for the maintenance and regeneration of its life cycles, structure, functions and evolutionary processes."(1) Bolivia has adopted a Law of the Rights of Mother Earth, conferring upon Mother Earth rights to, amongst others, life, water, clean air, and freedom from contamination.(2) New Zealand has recognized the Whanganui River(3) and Mount Taranaki(4) as legal persons with all the rights and duties that accompany such personhood. In total, as of January 2021, there are at least 178 legal provisions recognizing RoN worldwide (these vary from constitutions and laws to court rulings).(5)

In a nutshell, proponents of RoN claim that, by granting legal personhood and fundamental rights to nonhuman nature, these entities (trees, rivers, ecosystems) are recognized to have rights of their own that can be enforced against other legal persons. Nonhuman nature will then no longer be an object to use as we please, but a subject with rights. According to the proponents of RoN, if we only would recognize that natural objects and ecosystems are not things with which we may do as we like, but in fact are subjects of rights that place limits on the scope of permissible actions, then we might have a shot at addressing the environmental crisis in a satisfactory manner. This paper sets out to assess this claim, concluding that RoN is morally problematic, conceptually flawed and inefficacious in practice. Two alternative approaches – human rights and environmental virtue ethics – are discussed that combined could secure fundamental rights for humans regarding the environment, cultivate the appropriate character traits in them, and protect nature in the process.

(1) Constitution of the Republic of Ecuador, October 20, 2008. English translation available online at: https://pdba.georgetown.edu/Constitutions/Ecuador/english08.html.

(2) Ley de Derechos de la Madre Tierra [Law of the Rights of Mother Earth] No. 71, art. 7 (2010) (Bolivia).

(3) Te Awa Tupua Act (Whanganui River Claims Settlement) Act 2017, §§ 14–15 (New Zealand).

(4) Ngā Iwi o Taranaki and the Crown: Record of Understanding for Mount Taranaki, Pouākai and the Kaitake Ranges 2017, § 5 (New Zealand).

(5) Craig M. Kauffman & Pamela L. Martin, The Politics of Rights of Nature: Strategies for Building a More Sustainable Future (MIT Press 2021)

Alex Putzer, Tineke Lambooy, Ronald Jeurissen, and Eunsu Kim. 2022. "Putting the Rights of Nature on the Map. A Quantitative Analysis of Rights of Nature Initiatives across the World." Journal of Maps 18 (1): 89–96.

A Tale of two Models of Rights of Nature

Matthias Kramm

Wageningen University & Resarch

During the last decades, Rights of Nature (RoN) have become an important legal device to protect vulnerable ecosystems.(1) After their introduction in the Ecuadorian constitution in 2008, Rights of Nature have been incorporated in legal frameworks of various other countries such as Aotearoa New Zealand, Bolivia, Canada, Colombia, Mexico, India, and the United States at the national, regional, and local level.

Rights of Nature can be justified based on three types of arguments – a metaphysical, a metaethical, and a consequentialist-contractarian argument –, which I will introduce in the first section of the article.(2) While the first two types of arguments are based on legal realism and take an ecocentric stance according to which human law should be adapted to the Nature in which we are living, the third type of argument remains anthropocentric and considers law a tool that human beings use to balance the legal claims between human, social, and economic agents and Nature. Consequently, Rights of Nature can be supported by both ecocentric and anthropocentric arguments.

In the second section, I compare the two main models in which Rights of Nature have been conceptualized, the Nature's Rights-model and the Legal Personhood-model.(3) I argue that the second model has significant advantages compared to the first model. First, permanent legal representation facilitates the

initiation of legal proceedings or even the prior conduct of meditation proceedings. Second, the protection of specific ecosystems allows the regulatory framework to be adapted to the particular characteristics of the ecosystem. And third, the connection between local ecosystems and local people leads to an identification between Nature and its protectors and enables further institutional formats such as co-management systems or the organization of ecosystems as commons.

However, there are a variety of additional external factors that have to be considered before choosing one model over the other. I provide an analysis of these external factors in the third and last section of the article. The list of these factors comprise the questions of (a) whether there is weak or strong rule of law in a particular country, (b) whether environmental activism is protected, (c) which environmental legislation and property system are in place, and (d) whether Rights of Nature are supported by the populations as a whole or rather by specific communities.

(1) Rafi Youatt. 2017. "Personhood and the Rights of Nature: The New Subjects of Contemporary Earth Politics." International Political Sociology 11: 39–54.

(2) Patrick Baard. 2021. "Fundamental Challenges for Rights of Nature." In Rights of Nature: A Re-Examination, edited by Daniel P. Corrigan and Markku Oksanen, 156–75. London, New York: Routledge.

Visa A. Kurki. 2019. A Theory of Legal Personhood. Oxford: Oxford University Press.

(3) Craig M. Kauffman, Pamela L. Martin. 2021. The Politics of Rights of Nature: Strategies for Building a More Sustainable Future. Cambridge MA: MIT Press.

ID: 223 / **Panel 2-1-D: 1**
Symposium paper
Topics: Social and political philosophy, Ethics (general and applied), Metaethics, Moral Psychology
Keywords: Rights of Nature, Environmental Ethics, Political Philosophy, Philosophy of Law, Ethics

A Tale of two Models of Rights of Nature

Matthias Kramm

Wageningen University & Resarch

During the last decades, Rights of Nature (RoN) have become an important legal device to protect vulnerable ecosystems.(1) After their introduction in the Ecuadorian constitution in 2008, Rights of Nature have been incorporated in legal frameworks of various other countries such as Aotearoa New Zealand, Bolivia, Canada, Colombia, Mexico, India, and the United States at the national, regional, and local level.

Rights of Nature can be justified based on three types of arguments – a metaphysical, a metaethical, and a consequentialist-contractarian argument –, which I will introduce in the first section of the article.(2) While the first two types of arguments are based on legal realism and take an ecocentric stance according to which human law should be adapted to the Nature in which we are living, the third type of argument remains anthropocentric and considers law a tool that human beings use to balance the legal claims between human, social, and economic agents and Nature. Consequently, Rights of Nature can be supported by both ecocentric and anthropocentric arguments.

In the second section, I compare the two main models in which Rights of Nature have been conceptualized, the Nature's Rights-model and the Legal Personhood-model.(3) I argue that the second model has significant advantages compared to the first model. First, permanent legal representation facilitates the initiation of legal proceedings or even the prior conduct of meditation proceedings. Second, the protection of specific ecosystems allows the regulatory framework to be adapted to the particular characteristics of the ecosystem. And third, the connection between local ecosystems and local people leads to an identification between Nature and its protectors and enables further institutional formats such as co-management systems or the organization of ecosystems as commons.

However, there are a variety of additional external factors that have to be considered before choosing one model over the other. I provide an analysis of these external factors in the third and last section of the article. The list of these factors comprise the questions of (a) whether there is weak or strong rule of law in a particular country, (b) whether environmental activism is protected, (c) which environmental legislation and property system are in place, and (d) whether Rights of Nature are supported by the populations as a whole or rather by specific communities.

(1) Rafi Youatt. 2017. "Personhood and the Rights of Nature: The New Subjects of Contemporary Earth Politics." International Political Sociology 11: 39–54.

(2) Patrick Baard. 2021. "Fundamental Challenges for Rights of Nature." In Rights of Nature: A Re-Examination, edited by Daniel P. Corrigan and Markku Oksanen, 156–75. London, New York: Routledge.

Visa A. Kurki. 2019. A Theory of Legal Personhood. Oxford: Oxford University Press.

(3) Craig M. Kauffman, Pamela L. Martin. 2021. The Politics of Rights of Nature: Strategies for Building a More Sustainable Future. Cambridge MA: MIT Press.

ID: 221 / **Panel 2-3-D: 1**
Symposium paper
Topics: Social and political philosophy, Ethics (general and applied), Metaethics, Moral Psychology
Keywords: Rights of Nature, Environmental Ethics, Political Philosophy, Philosophy of Law, Ethics

Rights of Nature proxies

Alex Putzer

Sant'Anna School of Advanced Studies

By promoting a non-anthropocentric understanding of the environment, the Rights of Nature emphasize the importance of inclusive ethics and justice.(1) Rather than considering nature as an object to be exploited and dominated, it is perceived as a subject, to be respected and cherished, with intrinsic value independent of human interests. A river, for instance, is protected for its own sake, rather than its usefulness for fishing, generating electricity, or as a waterway. While this appears to be a rather straightforward approach, the parts of nature which are specifically included or excluded, remain either un- or ill-defined. The Whanganui river in Aoetearoa New Zealand, for instance, represents one of the most recognizable legal Rights of Nature initiatives. The Whanganui River Settlement Act - passed by the national government in 2017 - however, does not protect the entirety of the river, but only the riverbed.(2) This approach was chosen to avoid the complicated alteration of existing claims to the entity's remaining parts. Notwithstanding the practical reasoning, the theoretical one seems to be missing. In this contribution, I delve into the theories behind why certain initiatives use very specific and others very general approaches to define nature.

For this, I look at the major ways nature is conceptualized and compare it to existing Rights of Nature initiatives. My claim is that these initiatives use natures as proxies for other goals. As such, perceiving the environment in a holistic way represents a proxy for inclusion, focusing on biodiversity emphasizes resilience, and a physical account underlines measurability. The inclusion of indigenous peoples and their philosophies is a proxy for justice and inclusion. As such, this latter proxy seems to be among the visions that are closest to what the overall movement might want to stand for. This, however, does not make it the perfect candidate. Through several additional reflections, it becomes clear that different natures and different proxies lead to different rights holders as well as to different outcomes in specific contexts. An awareness of the differences and especially the contradictions can help to create more targeted as well as more efficient environmental protection efforts.

(1) Christopher Stone. 1972. "Should Trees Have a Standing? Towards Legal Rights for Natural Objects." Southern California Law Review 45: 450–501.

(2) Toni Collins and Shea Esterling. 2019. "Fluid Personality: Indigenous Rights and the "Te Awa Tupua (Whanganui River Claims Settlement) Act 2017" in Aotearoa New Zealand. Melbourne Journal of International Law 20 (1): 197–220.

ID: 220 / **Panel 2-2-D: 1**
Symposium paper
Topics: Social and political philosophy, Ethics (general and applied), Metaethics, Moral Psychology
Keywords: Rights of Nature, Environmental Ethics, Political Philosophy, Philosophy of Law, Ethics

The liberal Rights Paradigm and Rights of Nature

Adelheid van Luipen
Maastricht University

In current times, it becomes abundantly clear that humans leave a big mark on the deterioration of our planet, causing many environmental crises, while at the same time our interconnectedness and dependencies on sometimes unpredictable natural phenomena, come to the fore.(1) As some legal theorists argue, one of the most pressing questions philosophers face is to imagine a profound legal response to climate crises and to rethink laws foundations. Critical legal scholarship, building upon new materialist and post-humanist philosophies, have urged for radical imagination in the legal realm to legally accommodate relations in and to a more-than-human world.(2) A legal path with momentum is Rights of Nature theory (RoN).(3) Yet, RoN have also raised numerous questions and critiques, particularly with regards to their emphasis on rights. A question is whether RoN are a mere stop along the way towards a more radical legal approach when they confirm or merely slightly question the dominant Western liberal rights paradigm.

This article aims to bring together different strands in (mainly) legal scholarship on the relation between the liberal legal rights paradigm and RoN. Two questions will be central. Firstly, on the relation between liberal rights and RoN. To what extent do RoN coincide with the liberal rights paradigm? The second part revolves around the question whether, and if so, why this presumed relation is problematic. In particular, this article focuses on the possible problematic character of this relation in light of the relations a more-than-human world would require. This might lead to insights on RoN's potential to legally accommodate relations in a more-than-human world and/or which (other) roads to explore.

(1) IPCC 2022: Climate Change 2022: Impacts, Adaptation and Vulnerability. Contribution of Working Group II to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change.

(2) See for instance A. Grear, 'Towards new legal futures? In search of renewing foundations', in: A. Grear and E. Grant (ed.), Thought, Law, Rights and Action in the Age of Environmental Crisis, Cheltenham: Edward Elgar Publishing 2015, p. 284-289; E. Jones, 'Posthuman international law and the rights of nature', in: A. Grear e.a. (ed.), Posthuman Legalities: New Materialism and Law Beyond the Human (Journal of Human Rights and the Environment 2021 (12), special issue), Cheltenham: Edward Elgar Publishing, p. 93-94.

(3) We follow the policy of various scholars and organizations, including the United Nations Harmony with Nature Programme, which writes Nature in uppercase to establish it also semantically as a subject (See, for instance, footnote 1 of the UN GA Resolution A/75/266). This policy also emphasizes the contrast to a lower-case, anthropocentric object vision of nature.

ID: 148 / Panel 5-2-C: 1
Individual paper
Topics: Ethics (general and applied), Metaethics, Moral Psychology
Keywords: moral progress, moral bioenhancement, genuine morality, instrumental morality

Biomedical enhancement of morality and moral progress

Karolina Kudlek

Technologies have been influencing and facilitating processes of social, cultural and moral change throughout human history – from agrarian and industrial, to sexual revolutions. Therefore, using technologies to bring about morally desirable change should not be anything unusual or controversial in and of itself. Nevertheless, biomedical technologies that specifically aim at directly improving morality have been causing a lot of controversy in ethical discussion.

Biomedical means are often seen as ''shortcuts'' in moral improvement which would bypass relevant processes of moral deliberation and, therefore, produce ''faux'' morality. Some strongly believe that it is essential for morality that moral agents make free, autonomous choices based on reason, evidence and argument, rather than simply enhancing the doing and not the knowing of what is good. This implies that there is an important distinction to be made between ''fake'' and ''genuine'' morality or, more concretely here, between 'instrumental'' and ''genuine'' moral progress, and that biomedically assisted moral change will not amount to genuine moral progress.

In this presentation, I dig into the philosophical discussion on moral progress to examine whether, and under what conditions, would biomedical enhancements of morality lead to progressive moral change. I distinguish more closely between the conception of genuine and instrumental moral progress to show two main things. For one, although it is unlikely to happen in the near or mid-term future, we cannot entirely dismiss the possibility of biomedically assisted moral change leading to genuine moral progress. Two, genuine moral progress may be overrated since instrumental moral progress is doing more work than we originally thought. Surely there are good reasons to think that genuine moral progress is overall better than instrumental one (because it is deliberate, deeper, longer lasting etc.), but that does not render instrumental moral progress bad or irrelevant. I offer three additional reasons: i) genuine moral progress is an ideal construct that many instances of paradigmatic moral progress do not match; ii) instrumental moral progress plays a relevant role in and can contribute to genuine moral progress (in a similar way as instrumental value contributes to intrinsic value); and iii) it seems reasonable to think that sustaining moral progress instrumentally can be justified in certain circumstances, especially in high-risk, urgent situations such as those described in the moral bioenhancement literature.

Therefore, we do not have a sufficiently strong case to render biomedically assisted moral progress implausible, ineffective and unnecessary.

ID: 187 / **Panel 2-2-G: 1**
Individual paper
Topics: Epistemology
Keywords: Epistemology, universities, higher education, co-creation, multidisciplinary research

Co-creating Tools to Foster the Epistemic Responsibilities of Universities

Iris Lechner, Emma Ajdari, René van Woudenberg, Jeroen de Ridder, Lex Bouter, Joeri Tijdink

To answer the question as to what constitutes a good university many turn to a narrow range of measurable indicators (e.g. citations, funding acquisition, academic awards and student numbers). At best, such indicators capture only part of what makes universities good. The question what constitutes a good university, however, can also be approached from a broader perspective. One such way is to allude to their epistemic responsibilities (ERs). In its broadest sense, ERs concern equipping and empowering researchers, educators, and students to attain, produce, exchange, disseminate and safeguard knowledge. In our previous study we identified and described six core ERs: 1) to foster research integrity 2) to stimulate the development of intellectual virtues 3) to address the big questions of life 4) to cultivate the diversity of disciplinary fields 5) to serve and engage with society at large and 6) to cultivate and safeguard academic freedom. Fulfilment of ERs might contribute towards redefining what makes universities good.

In recent work our aim was to translate the fulfilment of ERs from a theoretical to a practical level, in order to develop, or specifically ¬co-create a tool for universities to think about their ERs. Co-creation is a method in which researchers and participants collaborate to solve a particular problem, or design a specific product. It uses democratic knowledge formation, including diverse and opposing perspectives. Co-creation moreover assumes that everyone is creative, and that the people involved or impacted by a specific problem hold the relevant knowledge to solve it. We hosted three rounds of workshops in the fall of 2022 (seven workshops in total). The analyses of each workshop fed into and shaped the next workshops, making use of an iterative process.

We included 25 participants (in six online sessions and one face-to-face session), from different backgrounds, nationalities, disciplines, and career stages. During the workshops participants were asked to do a series of exercises, followed by a dialogue about the findings. In a workshop we made use of (online) materials such as images, drawings, post-it notes, etc. While we initially set out to develop one tool, throughout the course of the study three tools were developed in more detail.

These are: 1) to establish independent red teams at universities to critically reflect on the ERs 2) to develop collaborative spaces to foster ERs and 3) to integrate events and activities in the university to stimulate ERs. These three ideas focus on building communities within the universities, focusing on fostering, stimulating, and cultivating ERs. In the first months of 2023 we expect to finalize the analyses and further develop abovementioned tools.

Our research project is at the intersection of philosophy, qualitative research and policy design. We began with a philosophical concept – i.e. ERs of universities - and developed concrete tools which can ultimately be used by different stakeholders in universities. In this talk I will present the findings of the co-creation study, and reflect on whether co-creation can be a valuable method to translate philosophical concepts to address societal and practical problems.

ID: 134 / **Panel 5-2-B: 1**
Individual paper
Topics: Logic, Metaphysics
Keywords: relativism, fragmentalism, time, standpoints, immersive thought

On standpoints and facts

Martin Lipman

There is a metaphysical picture that emerges in a range of philosophical discussions, according to which the world comprises facts that obtain from different standpoints even though these facts seem to conflict. I take it to be a challenge to make sense of this picture, but one worth facing.

A prominent area in which this picture emerges is the philosophy of time. For instance, Kit Fine, in his 'Tense and Reality' (2005), discusses what he calls 'non-standard realism' about time, which accepts that reality is constituted by tensed facts while denying that a single time is privileged with regards to the metaphysical status of the facts that obtain at it.

The aim of this talk is to sketch the key parts of a new conceptual framework that helps us make sense of this type of view. I start with a brief discussion of the motivation behind it in the case of time. Building on this motivation, I sketch accounts of what it is 'to adopt a standpoint' in our metaphysically serious descriptions of the world, of what it is 'to switch' such standpoints, and of what it is for a fact 'to obtain relative to' something. Key to these accounts will be a distinction between two kinds of logical structure: the logical structure constraining the facts tied to a given standpoint and the logical structure constraining the facts found across standpoints.

If one could accept the sketched conceptual framework, we obtain an attractive way of understanding how a world can comprise the conflicting facts found across standpoints. The framework is generally applicable, and my hope is that it may shed some new light on a variety of philosophical discussions.

Autonomy of the politics and realist normativity

Ivane Lomidze

Recent debates regarding the nature of political normativity between proponents of realist political theory and its opponents are largely concerned with the extent to which politics could be an autonomous domain producing its own sources for normativity without being reliant upon pre-political moral assumptions. The main purpose of the present article is twofold:

First, it suggests dissolving the decades-long question of whether politics has primacy and whether ethics should play any role in political normativity. The debate between realists advocating the idea of politics as an autonomous domain (Geuss, 2005; Williams, 2005; Bellamy, 2010; Newey, 2010; Galston, 2010; Sleat and Rossi, 2014; Jubb and Rossi 2015) and the authors questioning the capacity of the political to produce its own normative sources (Erman and Möller, 2013; Estlund, 2017; Larmore, 2018; Leader Maynard and Worsnip, 2018; Gutmann 2018) rests on a wrong assumption regarding the ways normative reason works. Those approaches assume that there is a robust metaphysical difference between ethical and political reasons, discuss which of them has the primacy and to what extent political judgment could be autonomous from an ethical sphere. In my opinion, the question is not that of primacy or independence of politics from ethics; both spheres have their unique subject matter requiring the unique adjudicative conceptual apparatus, but the reason when judging a case at hand follows the very same patterns in every sphere. Thus, instead of being concerned with the question of primacy, it could be argued that both political, as well as ethical judgments, follow the general patterns of the ways normative reason works – in a "holistic" way as defined by moral particularists, most notably by Jonathan Dancy. This does not annul the autonomy of the politics and does not lead to methodological moralism, but on the contrary: it grants politics a high degree of autonomy and grounds political normativity as a distinct domain; politics remains distinct from and irreducible to morality;

Second, based on the argument of "holism in the theory of reason", the paper reintroduces the idea of principles – defeasible generalizations, as advocated by moral particularists and outlines the ways it could inform realist political normativity. More specifically, the discussions in the theory of ethics during the last two decades showed that ethics could be "particularistic", conceptualizing the idea of principles in a way that on the one hand, avoids generalism and universalism and on the other hand, maintains certain general criteria to make a judgment regarding particular cases. Nevertheless, political realism and the idea of general principles are compatible only when principles are seen as "defeasible generalizations" or "weak principles" as offered by moral particularism. I think discussion among moral particularists – between "strong" (Dancy 1993, 2000, 2004) and "weak" particularists (Little 2000, 2008; McNaughton 1988; McNaughton & Rawling 2000, 2008, 2013) could help us to reframe our understanding of political principles and define a normative core for realist political theory.

ID: 163 / **Panel 5-2-D: 1**
Individual paper
Topics: Philosophy of Science, Aesthetics and philosophy of art, History of philosophy (from ancient to modern)
Keywords: Movement, Übermensch, Eternal Recurrence, De-deification, Topography

The Topography of the Gay Science

Jordan Georgiev Mantov

In his seminal translation of The Gay Science, Walter Kaufmann famously remarked that while the book may seem to be a near random collection of aphorisms, there is in fact an intricate structure which constitutes the book as such. In this essay, I argue for a holistic structural reading of The Gay Science and I approach, what I term, its 'topography'. First, I draw attention to the books basic conceptual structure. Briefly, Book I sets out the problem; namely, the distinctively human craving for other-worldly purpose to justify existence. Book II furthers the problematic and offers a social commentary of contemporary morality and states of affairs. Book III offers a structured critique of traditional moral values in the context of Nietzsche's project of 'de-deification' of nature and the death of God. Book IV begins to set out Nietzsche's morality of 'yes-saying', which culminates in the teaching of the eternal recurrence and the appearance of Zarathustra. Finally, Book V, written in the aftermath of Zarathustra, is a celebration of this new morality professed by Zarathustra. Second, I argue that a structural understanding of The Gay Science is aided significantly by approaching the work as having a topographic aspect. Briefly, I argue that throughout the work Nietzsche employs imagery of terrain and encourages the reader, as it were, to embark and navigate on a journey through shifting terrain. I do so by drawing attention to a number of spatial-physical imagery and motifs which play an important role in understanding many of the core ideas of the work. Treating The Gay Science as possessing a topography, I argue, permits the reader to understand the importance of movement in Nietzsche's philosophy of becoming. Finally, I apply these new insights by examining the topographic imagery in two important ideas in Nietzsche's thought; namely, the death of God and the übermensch. The death of God is described by Nietzsche as a collapse of all solid ground and the loss of cardinal directions; when the implications of this event are thought through, it is akin to awaking in a horizon of 'infinite nothingness'. It is in this context that Zarathustra, the übermensch, emerges as a traveler, explorer, and navigator.

Civil disobedience in semi-authoritarian regimes

Hrayr Manukyan

When is it justified and legitimate to resort to civil disobedience? This question confronts citizens who have strong moral objections to their governments' (in)action. Existing normative theories of civil disobedience in political philosophy are useful but were written for civil disobedience within functioning liberal-democratic states. Being disobedient in such a context means challenging one aspect of the policies and laws of such a society, while the general background constitution of the society accepts fundamental and democratic rights to participation. This is very different for semi-authoritarian regimes. The normative principles that should guide our political thinking about legitimate civil disobedience in such a context must be very different.

In this paper, I suggest a three-level theory of civil disobedience: 1) general theory, 2) regime/context-specific theory, and 3) country-level theory. This approach to civil disobedience is new because existing theories do not distinguish between the general and context-specific features of civil disobedience but work against an assumed background of liberal-democratic societies. This three-level approach paves the way to make the major innovation of this paper: to articulate a comprehensive approach to civil disobedience for semi-authoritarian regimes.

I conceptualise civil disobedience in the semi-authoritarian context in light of the 2018 Armenian protests. This way of theorising about civil disobedience is called the "learning from the streets" approach. I also use theoretical resources of three standard theories of civil disobedience (liberal, democratic, and religious-spiritual) and two non-standard approaches to civil disobedience (radical democratic approaches and constructive disobedience). I introduce civil disobedience as an underappreciated democratisation tool in semi-authoritarian regimes.

The structure of the paper is as follows. In the first part of the paper, I present a general theory of civil disobedience and demarcate civil disobedience from other types of protests and practices. As a general political practice, I define civil disobedience as a kind of 1) non-violent protest, 2) manifested in the form of public lawbreaking, 3) which is predicated on the deep respect for the rule of law. I argue that these three features are the core features of civil disobedience in all contexts (including in liberal-democratic and (semi-)authoritarian contexts). In addition to these three core features, civil disobedience has context-specific (context-dependent) features.

In the second part of the paper, I present liberal-democratic context-specific features of civil disobedience. Particularly, I argue that civil disobedience in the liberal-democratic context aims to change laws or policies or to (radically) reform the existing democratic political system/order. Changes in laws/policies usually concern the basic rights of minorities (as the liberal theory of civil disobedience highlights), democratic institutions (as the democratic theory of civil disobedience highlights), or climate/environmental issues.

In the third (and main) part of the paper, I present semi-authoritarian context-specific features of civil disobedience. Particularly I argue (in light of the 2018 Armenian protests) that civil disobedience in semi-authoritarian regimes aims to change or has the potential to change the existing semi-authoritarian political system/order. In other words, in the semi-authoritarian framework, civil disobedience is a much more radical political practice than in the liberal-democratic context.

ID: 238 / Keynote 2: 1
Individual paper
Topics: Metaphysics, History of philosophy (from ancient to modern)
Keywords: Parmenides, Plato, Aristotle, essentialism, division, hylomorphism

Parmenidean Essentialism

Anna Marmodoro

It is a (philosophical) commonplace to think that Parmenides' philosophical influence on Western thought is all or mainly due to his ban of change, understood as the passage from being to non-being and vice versa. Here I argue that the most influential philosophical idea Parmenides bequeathed to us is a criterion for substancehood, according to which there is no division of any kind between a substance and what makes it what it is (its essence, in Aristotelian terms). This is a type of essentialism which denies 'essential predication' and all types of substance-making-relations. I call it Parmenidean Essentialism. I contend that it is different from today's essentialism (wrongly, I contend, attributed to Aristotle), according to which substances are characterised by essential properties, whose loss they cannot survive. I show that Plato and Aristotle endorsed Parmenides' criterion for substancehood, and in conceiving of their respective substances as kath' auto beings, identified and responded to various 'threats' of division within a substance. Such 'threats' arise from a substance's qualitative complexity, its mereological complexity, its hylomorphic complexity, and the complexity of property-instantiation. I examine the metaphysical 'mechanisms' Plato and Aristotle developed to address such 'threats' of division within a substance, such as the emergence of form, the descendance of form and the holism of the parts of a substance.

ID: 238 / Keynote 2: 1
Individual paper
Topics: Metaphysics, History of philosophy (from ancient to modern)
Keywords: Parmenides, Plato, Aristotle, essentialism, division, hylomorphism

ID: 116 / Poster session 2: 1
Poster
Topics: None of the topics above (please include your topic under 'keywords')
Keywords: Philosophy of Mind, Belief, Perception

Overlap: On the relation between believing and perceiving

Auke Montessori

Imagine seeing a red apple. You then come to believe that the apple is red. The common assumption is that this is a succession of separate states. I argue that, in some cases, there is identity between the experience and the belief. Belief and experience are not identical qua type, but they share some of their tokens. I call this phenomenon overlap.

Perceptual experience, in most cases, provides access to content. When I see a dog in a tree, the content there is a dog in the tree becomes accessible. I can directly infer that something is wrong, for example. This allows experiences to engage in belief-like behaviors. Since, on most views, functioning like a belief is all there is to being a belief, it follows that the experience in question is a belief. The experience retains its access providing role, its content, and its phenomenology, all while being a belief. Of course, not all perceptual experiences become beliefs. We do not always believe our senses.

By perceptual experience, I mean the conscious undergoing of a perceptual state, complete with content and phenomenology. I ignore forms of perception that are either unconscious or sub-personal.

A major obstacle to overlap is that perception and belief are often considered to have different kinds of content. Belief has conceptual content, consisting only of concepts the subject possesses. Perception has non-conceptual content, which contains elements that are not possessed concepts. Some even claim that they have different formats all together. If that is right, then perception and belief cannot overlap.

The solution is to claim that while most beliefs have propositional content, some have non-propositional content. When we perceive, we often gain access to non-propositional perceptual content. If endorsed, we can use this content in action and inference. That, plausibly, makes the experience count as a belief despite being non-propositional. Two problems are considered. The first is the idea that non-propositional content cannot feature in inference or action. I reply that if non-propositional content cannot fulfill these tasks, we should not consider perception to be non-propositional in the first place. The second problem concerns maintaining perceptual beliefs after the experience ends. I claim that when an experience ends, the perceptual belief ends too. It gets replaced by a regular propositional belief. This regular belief gets it content from stored information gained from the ended experience.

Embracing overlap has attractive benefits. There are modest intuitive and architectural reasons to like it. It also helps relieve epistemological problems introduced by Alex Byrne and Kathrin Glüer. They claim that the traditional view that beliefs and experiences are always separate states leads to epistemic trouble. Their solution is to claim that all experiences are beliefs. That solution fails. If I think I am looking at an illusion, I won't believe what I experience. Overlap can also resolve the problems introduced by Byrne and Glüer, without the implausible claim that we believe everything we perceptually experience. This shows that overlap is a useful tool for the philosopher's toolbox.

The challenges of cognitive and psychological language in contemporary AI

Diego Morales

Artificial intelligence (AI) and the cognitive sciences have a rich, but complicated relationship. Although not as close anymore as in the days of the 'good old fashion AI', where both fields shared common assumptions and pursued a consistent set of goals (Boden, 2008), the strong influence that the cognitive sciences have exerted over theories, methodologies, and conceptual frameworks in AI is palpable.

This may be appreciated in the fact that novel processing architectures, designs, and capabilities in AI continue being described using cognitive and psychological language. Whether this is historical aftertaste or a more substantial phenomenon, the fact remains that, as computational and processing power escalates, and AI systems exhibit increasingly sophisticated new techniques and applications, terms such as metacognition, awareness, agency, theory of mind, and consciousness are increasingly being used to explicate and conceptualize features of AIs (e.g., Cuzzolin et al., 2020; Dehaene et al., 2017; Johnson, 2022; Williams et al., 2022).

As convenient (Ringle, 1983) and beneficial (Hassabis et al., 2017) as the employment of cognitive and psychological language in AI may be, it remains, however, a problematic and tricky business. There doesn't seem to be a consensus on how to import and apply terms from psychology or the cognitive sciences, generating an array of situations that range from simply using a term, while ignoring the theoretical baggage and commitments that comes with it, to literal attributions of properties or capacities, and everything in between. Furthermore, the use of psychological and cognitive terms may readily give way to suspicions of idealized mnemonics (McDermott, 1976), misleading anthropomorphism (Salles et al., 2020), and harmful moral conceptualization of AI's capacities (Watson, 2019).

In this context, the aim of this talk is to critically examine the use of terms from psychology and the cognitive sciences in the field of AI, and suggest a plausible explanation for the state of affairs previously described. I will argue that the process of conceptual pollination that continues to occur from the cognitive sciences into AI is heavily influenced by an unresolved identity crisis, namely, whether AI is a field closer to an engineering discipline or one akin to the cognitive sciences. This ambivalence, I suggest, is of greater importance than it has been credited for, as it systematically plays out in three fronts: how these terms should be used; what sort of import, if any, should these terms bring into the new domain; and, what methods should be employed to test the phenomena that they purportedly denote. Throughout the talk, I will strive to show that engineering-inclined attitudes towards this identity crisis yield different ways of facing these fronts than cognitive-oriented attitudes, thus generating the ample array of situations mentioned above.

ID: 209 / **Panel 2-1-H: 1**
Individual paper
Topics: None of the topics above (please include your topic under 'keywords')
Keywords: metaphilosophy, philosophical method, artificial intelligence

Digital philosophy: a methodological proposal

Vincent C. Müller

15th March, 2023

There are unresolved methodological issues with the "conceptual analysis" or "conceptual engineering" that we do in contemporary philosophy in that we know it cannot be just a linguistic analysis as in good old fashioned analytic philosophy and ordinary language philosophy, but neither do not want to leave conceptual issues to empirical research. Furthermore, we have a tradition to look at the conceptual issues of traditional philosophical problems from the case of the human, our own case – though occasionally we invoke special kinds of humans, or animals. When we do not want to do this, we are left with thought experiments (the superspartan, swamp man, angels, superintelli-gent AI, etc.).

But perhaps contemporary philosophy of Artificial Intelligence offers us a new tool for the philosophical toolbox: We can test our conceptual analysis by asking what would be required for an AI system to have the property we are discussing (e.g. moral agency). This would allow abstraction from the hu-man case, plus access to a testable environment of machines we know fairly well. Given the range of cognitive skills we can find, or at least imagine, for AI systems, this new tool allows us to cover a wide range of philosophical issues around the philosophy of mind, language, epistemology, ethics, etc.

I will outline this concept of a "digital philosophy" and illustrate its fruit-ful use through philosophical efforts that have already taken place, and a few that are under development.

Case studies of digital philosophy avant la lettre would include:

- H. Putnam: What is belief? The mental content of a person is underdetermined by their mental state; there is always a "contribution of the environment" (Twin-Earth & machine functionalism) (Putnam 1960, 1981)

- D. Dennett: What is rational decision-making? Rationality cannot mean the consideration of all that is relevant for a decision (Frame Problem) (Dennett 1984; Shanahan 2016)

Case studies of current digital philosophy would include:

- Moral agency – for AI systems and humans (Allen et al. 2000; Müller 2021; Nyholm 2023)

========

Allen, Colin; Varner, Gary and Zinser, Jason (2000), 'Prolegomena to any future artificial moral agent', Journal of Experimental & Theoretical Artificial Intelligence, 12 (3), 251-61.

Dennett, Daniel C (1984), 'Cognitive wheels: The frame problem of AI', in Christopher Hookway (ed.), Minds, machines, and evolution: Philosophical studies (Cambridge: Cambridge University Press), 129-52.

Müller, Vincent C. (2021), 'Is it time for robot rights? Moral status in artificial entities', Ethics & Information Technology, 23 (3), 579–87.

Nyholm, Sven (2023), 'Robotic Animism: The Ethics of Attributing Minds and Personality to Robots with Artificial Intelligence', in Tiddy Smith (ed.), Animism and Philosophy of Religion (Cham: Springer International Publishing), 313-40.

Putnam, Hilary (1960), 'Minds and machines', Mind, Language and Reality. Philosophical Papers, Volume II (Cambridge: Cambridge University Press 1975), 362–85.

— (1981), Reason, truth and history (Cambridge: Cambridge University Press).

Shanahan, Murray (2016), 'The frame problem', in Edward N Zalta (ed.), Stanford Encyclopedia of Philosophy (Spring 2016 edition; Palo Alto: CSLI, Stanford University).

ID: 172 / **Panel 4-2-G: 1**
Individual paper
Topics: Ethics (general and applied), Metaethics, Moral Psychology
Keywords: Collective Action Problem, Act Consequentialism, Climate Change, Causal Inefficacy

Why the problem of inconsequentialism might not (always) be a problem for consequentialism

<u>**Benjamin James Mullins**</u>

'Pure' moral collective action problems -- such as Derek Parfit's 'Drops of Water'/'Harmless Torturers' examples -- are often thought to provide a counterexample to Act Consequentialism. This is because many individual acts can promote significant increases in welfare. Yet, because each act is inconsequential -- i.e. each act itself makes no difference -- then Act Consequentialism does not say that any individual has a reason to perform such acts. Which, on the face of it, seems implausible.

Be that as it may, certain authors -- call these the 'critics of Act Consequentialism' -- then go on to present a whole array of 'real-world' moral collective action cases which seemingly have the same structure as the 'pure cases' given above. Cases are many, but we mainly focus on two prominent examples from the literature: (1) Climate Change and (2) Becoming a vegetarian.

The critic argues that, given these 'real-world' moral collective action problems have the exact same structure as the 'pure' moral collective action problems, then it looks as if they too present counterexamples to Act Consequentialism. And provided such scenarios are ubiquitous in everyday life, this presents an additional blow to Act Consequentialism.

In this article, I want to draw attention to important differences between the 'pure' cases and the 'real-world' cases. I argue that once these differences have been cleared up, then it is no longer obvious that the real-world scenarios provide a problem for Act Consequentialism.

The strategy will be as follows. I will argue that, contrary to the 'pure' cases, in which the agent usually has only a few select options available to them, in 'real-world' cases, the options available to the agent are far greater. And that Act Consequentialism may imply that agents perform other such acts which might make a difference. If this is so, then, I submit, such cases may no longer pose such a problem for Act Consequentialism, and the burden of proof lies with the critic to show that they do so.

This result should be welcomed by the Act Consequentialist. I conclude, however, that there remain pressing issues yet. I end by noting that Act Consequentialism still faces issues concerning overdemandingness and cluelessness.

ID: 140 / **Panel 1-2-B: 1**
Individual paper
Topics: Aesthetics and philosophy of art
Keywords: Fictional characters, characterhood, imagination, internal perspective, cognitive poetics

Fictional Characters and Ourselves: Towards a Definition of Characterhood

Alfonso Muñoz Corcuera

There is much disagreement regarding the metaphysics of fiction. However, authors with very different metaphysical commitments tend to agree on one thing. When engaging with a fictional work, its fictional characters seem to be like us (see e.g. Frow, 2014; Lamarque, 2007; Nussbaum, 1992; Ryan, 2018; Searle, 1979; von Solodkoff, 2019). This claim will be the departing point of my inquiry. And it is important to note that I take it to be metaphysically neutral. For some, the claim that fictional characters are like us means that their authors just pretend to be referring to beings like us (Searle, 1979, pp. 71-72). For others, it means that fictional characters are indeed beings like us who happen to exist in a different possible world (Ryan, 1991). But whatever fictional characters turn out to be from a metaphysical point of view, in some sense they seem to be like us. Using Lamarque and Olsen's terminology, we can say that, regardless of what they are from an external perspective, from the internal perspective they are imagined to be significantly like us (Lamarque & Olsen, 1994, pp. 143-148).

This claim is far from being controversial. Many different authors have endorsed it. However, besides some exceptions (e.g. Frow, 2014; Muñoz-Corcuera, 2016; Ryan, 2018), the claim has usually been set aside as either irrelevant or problematic. After all, fictional characters are not that much like us. And if we take the comparisons between fictional characters and ourselves literally, we could end up misconceiving both fictional characters and ourselves (Lamarque, 2007; Strawson, 2008). On the contrary, even if fictional characters are not that much like us, I think that we can learn something from these comparisons. Specifically, we can find a definition of what a fictional character is from the internal perspective. That is, we can find a class C so that we are Cs and being an individuated fictional C is a necessary and sufficient condition for being a fictional character.

At first, my inquiry can seem irrelevant. However, as I will try to show, reaching a definition of what fictional characters are from the internal perspective will have some interesting consequences. For example, it will offer new insights into the problem of the identity of fictional characters. In this regard, the structure of my paper will be as follows. First, in section 2 I will go through different unsuccessful candidates to be the class C that I am looking for. This inquiry will reveal some insights that, in section 3, will lead me to defend my main claim in this paper: fictional characters are fictional intentional systems. Finally, in section 4 I will explore some interesting consequences of my claim in a programmatic way.

ID: 123 / **Panel 3-3-C: 1**
Individual paper
Topics: Ethics (general and applied), Metaethics, Moral Psychology
Keywords: digital recordings, moral judgment, moral disruption, privacy, surveillance

Digital Recording and the Hazards of Unbounded Moral Judgment

Bart Kamphorst, <u>Elizabeth O'Neill</u>

The widespread adoption of internet-connected, camera-fitted devices such as smartphones, tablets, laptops, home security systems, and "smart" doorbells, combined with the continuously growing availability of cheap, reliable, cloud-based storage technologies, has facilitated a techno-social environment in which it is viable to make and store digital recordings such as photographs, audio fragments, and video streams at an unprecedented scale. We contend that today's digital recording practices threaten to radically alter how we perceive and evaluate ourselves and others, producing an ongoing, socially and morally disruptive shift towards unbounded moral judgment. We argue further that the trend toward unbounded moral judgment poses several hazards, including widespread, difficult-to-restore reputation damage, negatively altered self-perceptions, and even the stifling of morally right behavior.

To establish the existence of a shift towards unbounded moral judgment, we begin by noting four key, interlocking trends. First, the long-term availability of digital recordings makes it possible to retrieve and resubmit recordings to critical scrutiny and moral judgment with a higher frequency than in the past. Second, the ease of distributing digital recordings to others, both globally, but also over time and even to future generations vastly expands the potential number of people who can form judgments about the individuals and actions captured in the recordings. Third, the long-term availability and extended accessibility of digital recordings make it impossible to predict or control when and in what context recordings will reappear. Fourth, and finally, important features of digital files, such as their malleability, duplicability, etc., makes it impossible to predict or control how, by whom, and for what purpose digital recordings will be used and repurposed.

We argue that these trends together are facilitating the aforementioned shift towards unbounded moral judgment, which in turn creates significant ethical hazards. In particular, we hold that today's digital recording practices risk inflicting difficult-to-restore damage to reputations (through moral judgments by others) as well as negatively altering self-perceptions (through moral judgments of oneself by oneself and via the influence of others' moral judgments on one's own self-assessment). Moreover, we contend that these two risks will for some people be severe enough that anticipation of backlash can lead to morally problematic behavior when people are unwilling to speak up or stand out in contexts where (they believe) recordings are being made.

While there are often obvious immediate benefits to making and sharing digital recordings, the potential for digital recording to radically alter moral judgment practices remains underappreciated, as do the hazards associated with unbounded moral judgment. With an eye to how the associated technologies are projected to advance in the future—e.g. growing use of deep fakes and augmented reality—we conclude the paper with recommendations about technical, regulatory, societal, and individual approaches to mitigating these issues.

ID: 151 / Panel 5-1-A: 1
Individual paper
Topics: Social and political philosophy, History of philosophy (from ancient to modern)
Keywords: Quobna Ottobah Cugoano (c.1757–c.1791), slavery and antislavery, freedom, reparations, violence

After abolition: Cugoano on 'lawful servitude' and the injustice of slavery

Johan Olsthoorn

The Black antislavery theorist Quobna Ottobah Cugoano (c.1757–c.1791) is increasingly recognized as a noteworthy figure in the history of philosophy (e.g., Bernasconi 2019; Jorati [forthcoming]). Born in present-day Ghana, Cugoano was enslaved at the age of 13 and trafficked to Grenada, before being taken onwards to England, where the 1772 Somerset court ruling in effect freed him. His Thoughts and Sentiments on the Evil of Slavery [1787/1791] is considered "radical even by abolitionist standards"; it broke new ground by demanding the immediate end of the slave-trade and of slavery itself.

This paper parses Cugoano's vision for a post-abolition future. Existing interpretations focus either on his bold demand for slavery reparations (Best & Hartman 2005; Dahl 2020); or on his role in the failed Sierra Leone repatriation expedition (Peters 2017: 68-70). I will instead analyse his constructive proposals to transform colonial chattel slavery into something he calls "lawful servitude" (34, 98, 124). Though grounded in lived experience and progressive for the time (cf. Sharp 1776: 62-65; Ramsay 1784: 225-251), Cugoano's blueprints for transitional justice seem rather underwhelming today. Every enslaved person should continue to work for plantation owners – unpaid, yet "without torture or oppression" – until they have worked a total of seven years for them (including years prior to the abolition of slavery). Their unpaid labour nominally serves to compensate their erstwhile owners "for the expences attending their education" (98-99).

By reconstructing what post-abolition "lawful servitude" means exactly, I aim to clarify what the injustice of colonial slavery by contrast consists in for Cugoano. What kind of freedom should humans have as a matter of justice? Thoughts and Sentiments mentions at least seven wrong-making features of slavery. Some concern who is enslaved (the innocent; penally enslaving criminals is sometimes justified – 57-58). Others pertain to the nature of colonial slavery itself (e.g., subjection to arbitrary violence; involuntary sale; no Sunday rest).

The British government must "require all slave-holders… to mitigate the labour of their slaves to that of a lawful servitude" forthwith (98). The blueprint for that "lawful servitude", I argue, Cugoano found in the Old Testament. Moses's laws for Canaanite bondage capture the same "lawful servitude" (35-38, 124-127). The "vassalage" of "bond-servants" is a "lawful servitude" – "a state of equity and justice" (35-36). His insistence upon the morally benign nature of Mosaic bondage thus serves a dual role in his antislavery argument. It counters the objection that modern colonial slavery is biblically warranted (38, 127). And, surprisingly, it provides a model for just social relations to come.

For Cugoano, I conclude, the injustice of slavery consists foremost in subjection to arbitrary violence and absence of personal security more generally. Not in absence of freedom of self-direction or imposed inferior social status – as philosophers today have argued (Hare 1979; Roberts-Thomson 2008). We need not conclude that modern accounts of the injustice of slavery are defective. Rather, I suggest, the takeaway is that forms of human bondage come in more forms and degrees of atrocity than commonly recognized.

The (potential) role of critique in Responsible Innovation frameworks: addressing the political deficits

Lisann Penttilä

Frameworks of Responsible Innovation (RI) have been taken up and implemented by various institutions and sectors globally including the European Commission and tech giants such as Meta, just to name a few. While these frameworks have different formulations and emphases, they generally all assume that innovation can provide solutions to pressing problems and that the innovation processes need to be managed in particular ways to ensure they are responsible. RI frameworks commonly attribute a large importance to a multi-stakeholder approach and community or public engagement. However, despite this effort, political deficits have been identified in operative RI frameworks by the critical (academic) literature. Aspects of the political deficits include, but are not limited to, (1) conceptual and normative ambiguity surrounding e.g., participation and democracy, which in turn impact RI's practical uptake and (2) the strategic bypassing of participatory practices. Despite the different aspects of the political deficits identified, the critical literature often shares a concern for the power relations implicit in RI frameworks and uptakes. Yet, what is lacking in the critical literature is an appreciation of the role of social relations and social structures, which is vital to an analysis and criticism of power.

Building on the existing literature, this paper focuses specifically on the political deficits in the European Commission's RI framework, Responsible Research and Innovation (RRI), and explores how the theoretical insights of critical theory and the practice of critique can aid in addressing them. I argue that the RRI framework needs to engage with critique in two ways (1) internally, i.e., on the framework's own terms, examining its normative and conceptual underpinnings and (2) externally, relating to the everyday experiences and critiques of the broader public and local communities (affected both by the problem for which an innovative solution is needed, and the innovation process itself). This paper encourages the RRI framework (and RI frameworks more broadly) to identify and critique the social relations and structures that create, exercise, and perpetuate hegemonic forms of power with the aim of greater social emancipation. To do so, I engage with the work of critical theorist Andrew Feenberg—specifically his work on the role of bias in technological design and his support for the democratization of technology—to argue for the necessity of participatory action and critique in RRI frameworks. It is through participatory action and engaging in critique that the social relations and social structures which taint rationality, and limit the perceived possibilities of action, become apparent and can be addressed.

ID: 241 / Panel 1-3-H: 1
Individual paper
Topics: Philosophy of Science, Social and political philosophy, Ethics (general and applied), Metaethics, Moral Psychology
Keywords: Emotion recognition technology, privacy, relationships, disruption, affects

Emotion recognition technology: revisiting our conception(s) of privacy

Alexandra Prégent

Emotion recognition technology (ERT) imposes a constant scrutinization on people's affective states. Stepping aside from the debate around the scientific foundations of this technology, I argue that this new form of surveillance poses a significant risk to humans' reasoning mechanisms, decision making processes, relationships and well-being - in particular when there is a feedback loop provided to the end users – pressuring us to revise prominent accounts on privacy such as the control and access over control accounts. Emotions are a fundamental component of human experience and a necessity to live a meaningful life. By being recognised both as a part of the inner life and as a necessary means for communication, the status of emotions is a tricky one to place within traditional privacy paradigms.

Addressing the right to privacy through an ERT analysis requires us to recognise that there are elements of our lives that we voluntarily and involuntarily share on a daily basis, which does not take away from their inherently intimate nature, and keeps them as sensitive information.

By discussing ERT within the framework of the social dimensions of privacy (Roesler & Mokrosinska, 2015), this research develops a novel account of privacy and expands the scope of privacy regulations to the (public) emotional life by attributing value and meaning to the apparent futile behaviours, with important repercussions on social control, political manipulation, and preservation of democratic processes.

ID: 203 / **Poster session 2: 2**
Poster
Topics: Ethics (general and applied), Metaethics, Moral Psychology
Keywords: realism, normative truths, ontological commitment, deliberation, justification

Justified sources of ontological commitment towards normative truths

Drishtti Rawat

Normative realists justify their belief in the existence of normative truths on the basis of their features, such as deliberative indispensability (David Enoch 2011) or improving the coherence of theories in the normative domain (Thomas Scanlon 2014). In this paper, I investigate whether realists have good reason to source their ontological commitment in these features.

The objects of normative truths are not observable like trees and tables. Therefore, I assess whether realists have good reason to take these features as their source of ontological commitment based on whether they are justified in believing that these features are indeed such a source. In my negative argument, I argue that the realist is not justified in using these features as a source of their ontological commitment owing to contradictions with their irreducibility and, sometimes, fundamentalist theses.

My positive argument tackles the apparent invincibility of the realist's domain relativity. Owing to the parochial knowledge required to make domain-specific claims, it appears that only realists can determine whether normative truths are indispensable to the deliberative project or improve the coherence of normative theories. Therefore, I move my inquiry to a higher level of description. The realist cannot have an exclusive conception of features, like indispensability or coherence-improving, since all domains still have to do with a single reality.

In my positive argument, I suggest that these features do not belong to normative truths but are, instead, relations that seem to be established by normative deliberation. I do this, first, by arguing that the features of normative truths are contingent on our projects. This argument draws partly from the absence of a deliberative counterpart of an inference to the best explanation (IBE) to infer the primacy of normative truths (as argued by Alex Worsnip 2016). On the other hand, it is motivated by the possibility of not engaging in our deliberative projects well and the fact that not everyone engages in projects of the normative or metanormative domains.

The second part of my positive argument focuses on the malleability of the features described by the realist. What appears to have deliberative indispensability to us may appear to have explanatory indispensability to an alien species. Therefore, one would be more justified in believing that normative deliberation establishes this relation (between the observer, the project, and normative truths) instead of believing that normative truths have features that are the source of ontological commitment. Realists must then find another source of their ontological commitment or concern themselves with reasons for belief apart from justification. For now, without justification, the realist has no good reason to believe in the existence of normative normative truths– even if they indeed exist.

ID: 204 / **Panel 2-3-B: 1**
Individual paper
Topics: Ethics (general and applied), Metaethics, Moral Psychology
Keywords: benefiting from wrongdoings, beneficiaries, moral inconsistency, condemning, compensation

Condemning the wrongdoing you benefit from

Jan Willem Wieland, **Drishtti Rawat**

In December 2022, the Dutch state apologized for its slavery past. But should current generations actually take responsibility for what their ancestors did? Might they even have to do more and compensate for the crimes committed long ago?

According to many philosophers, benefiting from wrongdoing by others (such as our ancestors) can be wrong even when you are innocent of the wrongdoing itself. In this paper, we discuss the "moral consistency argument" (Daniel Butt 2014), which rests on the idea that you cannot genuinely condemn a wrongdoing if you do not compensate the victims. According to this view, beneficiaries must compensate victims because acting inconsistent with one's condemnation is wrong.

Various versions of this argument have been proposed and criticized (see the overview by Avia Pasternak 2017). One question is what the exact inconsistency amounts to: which attitudes are supposed to conflict? Another is what the given inconsistency proves: does it support the "compensation view" (that beneficiaries should compensate the victims), or rather the so-called "disgorgement view" (that beneficiaries should get rid of the tainted benefits)?

We suggest that, if the moral consistency argument is to succeed, it should account for three desiderata. First, the argument should apply to beneficiaries specifically, as opposed to third parties who did not benefit. Second, the argument should generate the "right reasons" for compensating victims, which cannot be only a desire for clean hands or mere consistency. Finally, the argument should generate reasons to remedy benefits received from rights (not only wrongs), following the same reasoning from inconsistency.

In this paper, we defend a version of the moral consistency argument that meets these three desiderata. We propose that the argument applies to beneficiaries specifically, not only because of the benefit they received but, because of a compensatory position defined by their attitudes and the value of those attitudes for the victim. This victim-focus perspective extends to the second desiderata such that the right reasons are those that have to do with concern for the victims and not the self. Finally, the moral consistency argument would be complete if it concluded that beneficiaries who received benefits due to good deeds intended for someone else must also compensate. At the same time, an argument from consistency cannot unintuitively conclude that receivers of good deeds ought to compensate, for example, donors merely because they condemn the act of giving charity.

ID: 206 / **Panel 1-3-D: 1**
Individual paper
Topics: Epistemology, Social and political philosophy, Feminist philosophy
Keywords: gaslighting, politics, self-trust

Collective gaslighting and collective self-trust

Natascha Rietdijk

Accusations of gaslighting occur increasingly often in the political domain. In discussions on, for instance, institutional racism, sexual harassment, or vaccination, it is regularly suggested that groups of people are gaslit. Does it make sense to talk about gaslighting in this collective sense? The term originates in psychology, where it is used to refer to one person manipulating another (often a partner or family member) into questioning their own senses, memory and judgment. I believe there is potential for the concept of gaslighting to be as illuminating in the political domain as in the context of intimate relationships, once we look at how collective gaslighting works to undermine collective self-trust.

So far, various philosophical analyses of specific types of collective gaslighting have been put forward, including accounts of racist gaslighting (Davis and Ernst 2017), misogynistic gaslighting (Stark 2019), and cultural gaslighting (Ruíz 2020). However, what is still missing and what I aim to develop is a more generalized account of collective gaslighting that can accommodate all of these more specific forms and also be applied to new, different forms, like post-truth political gaslighting. An important challenge for such an account is the worry that using the term gaslighting outside its traditional scope and context amounts to conceptual inflation. Philosophers like Abramson (2014) and Sweet (2019) explicitly state that gaslighting is a phenomenon of interpersonal relationships and Case (2019) argues that "overstretching" the concept impoverishes our understanding and is disrespectful to victims. Judging from the sometimes quite loose and polemic use of gaslighting in the political arena, I think this worry is not unfounded. Yet the challenge can be met if we understand gaslighting not in terms of the presence of specific manipulative intentions but in terms of how it affects self-trust.

Drawing on the existing literature on gaslighting and its criticisms, we can then distill three basic criteria that a theory of collective gaslighting should meet. First, it should do justice to the myriad forms gaslighting dynamics can take, without becoming too broad. Second, it should explain collective gaslighting in terms sufficiently similar to one-on-one gaslighting, while clarifying how its dynamics translate to the political domain. Third, it should help us distinguish justified political accusations of gaslighting from unfounded provocations as well as possible. On the basis of these criteria, I develop an account of collective gaslighting which centers on self-trust. Whereas individual gaslighting aims at undermining a target's trust in herself (see, e.g. Spear 2019), I define collective gaslighting as a process by which a collective's epistemic self-trust is illicitly and non-accidentally undermined. Building on El Kassar's (2021) work on collective intellectual self-trust, I will explain how collective gaslighting works to harm it, what this means for victims, and how it relates to and interacts with other forms of political and epistemic oppression. Finally I demonstrate that my account fulfills the three criteria and retains the central value of the term gaslighting: its power to reveal a very harmful dynamic and thereby making it possible to resist it.

ID: 160 / **Panel 3-1-G: 1**
Individual paper
Topics: Epistemology, Philosophy of Science, Social and political philosophy
Keywords: credit economy, social epistemology, values in science, science policy

What is credit in science? A value-based interpretation of the credit maximisation approach to the social philosophy of science

Thijs Hendrik Ringelberg

Questions concerning the social organisation of science are addressed increasingly often (and increasingly successfully) by means of what might be called the Credit Maximisation Approach (CMA). This approach employs computational techniques to model the behaviour of scientific communities on the assumption that scientists act in pursuit of "social credit", and thus are faced by an incentive structure called the "credit economy". But although the functioning of the credit economy is clear in the context of formal models, there is a disconnect between the formal models and the reality which they are meant to represent. The aim of this paper, therefore, is to establish how these models should be interpreted. I argue that the most plausible interpretation 1) casts the credit economy as dependent on and reflective of a pre-existing normative consensus; 2) reveals this normative consensus to be centred on a new type of value, and geared toward the common epistemic good; and 3) restricts the ways in which the CMA can be employed to give policy advice.

I argue that the credit economy functions through an interplay between scientists' attitudes and those scientific institutions I term "praise-institutions" (such as authorship conventions, citations, and academic prizes). Provided that this interplay works well, it is possible to be motivated in one's scientific work by credit incentives: a researcher might strive to make a discovery because she knows that to be the way to further her career. But due to the nature of the attitudes on which these incentives are based, the credit economy must reflect a pre-existing normative consensus, and cannot function autonomously.

The relevant attitudes are attitudes of esteem: scientists' evaluations of each other as scientists. The normative consensus that is a precondition for the functioning of the credit economy therefore focuses on the (shared) concept of a good scientist. I argue that standard analyses employing the CMA show us that this normative consensus is geared toward the common epistemic good: the fact that scientists endorse this picture of a good scientist, and not some other picture, has positive epistemic consequences for the scientific project.

This value-based understanding of the CMA has consequences for the way we can use the CMA to make science policy recommendations. Of central importance in this context are the praise-institutions, which function to communicate esteem-judgements across the scientific community. In order for this communication of esteem to be successful, the praise-institutions need to be perceived as tracking the normative consensus. Policy interventions which make adjustments to the praise-institutions run the risk of tarnishing this perception, thereby endangering the institution's capacity to communicate esteem. To establish whether such tarnishing will take place in a specific case, empirical knowledge is required about the normative consensus in the relevant scientific subdomain. Thus if we wish to base policy on the CMA, sociologists and social psychologists must be enlisted to explore the functioning of the credit economy in specific situations, lest the recommended policy interventions have unexpected effects.

ID: 147 / Panel 1-3-B: 1
Individual paper
Topics: Epistemology, Aesthetics and philosophy of art, Ethics (general and applied), Metaethics, Moral Psychology
Keywords: City, Epistemic Power, Knowledge, Aesthetics, Tourism

How to know a city: Tour guides and epistemic power

Catherine M. Robb, Pilar Lopez Cantero

When we travel to a new city, we gain knowledge about its physical terrain, directions, historical facts and aesthetic features. We can obtain this knowledge in a mediated way (i.e., through someone's testimony) or an unmediated way (that is, through direct acquaintance). When we engage in tourism practices, such as guided walking tours, our experiences are necessarily mediated. This has led experts in tourism studies (Urry 1990) and more recently in philosophy (Kukla 2021, 2022) to question the value of mediated knowledge acquisition. In response to this recent body of literature, we argue that from an epistemic point of view, mediated knowledge of a city is not necessarily disvaluable by nature. However, we suggest that city tour guides are furnished with epistemic power, and that the misuse of this power in their relationship with their audience can be epistemically and ethically problematic.

First, we outline the view according to which tourists are passive and lack automony, and further explain how this stance justifies the epistemic disvalue of city tours. We then show that this view relies on a simplistic view of both the tour guide and the tourist, as well as their collaborative role in generating knowledge of the city. Once the relation is redefined, we move on to define the type of knowledge that is transmitted by the tour guide. We argue that the tour guide has the potential to facilitate knowledge on two levels: testimonial and aesthetic. Although the knowledge gained on each of these levels is partial and mediated, we argue that this limitation does not negate epistemic value, for two reasons. First, testimony itself can be valuable (see for example Kusch 2009). Second, contra Kukla, tour guides do not simply transmit factual knowledge, but also objectual knowledge which does not need direct acquaintance, and would often be hard to obtain without mediation.

To conclude, we identify a feature of the relationship between the tour guide and tourist that is often overlooked in the discussion: the fact that the tour guide has epistemic power. We outline the nature of the epistemic power, and then explain how this power can be violated and produce negative outcomes that are both local (with respect to the tourist's experience of the city) and general (in terms of wider epistemic and ethical implications). We show that the potential disvalue of mediation in urban experiences lies in the misuse of this epistemic power, and not in the nature of the obtained knowledge as it has traditionally been argued.

References

Kukla, Quill R. 2022. Knowing Things and Going Places. European Journal of Philosophy, 1-17. https://doi.org/10.1111/ejop.12790.

Kukla, Quill R. 2021. City Living. Oxford: Oxford University Press.

Kusch, Martin. 2009. Testimony and the Value of Knowledge. In, Pritchard, Haddock & Millar (eds), Epistemic Value, Oxford: Oxford University Press, pp. 60-94.

Urry, J. 1990. The Tourist Gaze, London: Sage.

ID: 199 / Panel 2-2-C: 1
Individual paper
Topics: Philosophy of Science
Keywords: Philosophy of psychiatry, data science methods, personalized medicine, data-driven science

Thick descriptions in data-driven psychiatry

Jan-Willem Romeijn, Hanna van Loo

Owing to new computational methods and the availability of ever larger data sets, psychiatric research is becoming increasingly data-driven and so-called personalized. A good illustration is the I-SHARED project (Kan et al. 2019), in which clinicians use machine learning methods to predict the efficacy of interventions, proceed to share these predictions with their patients, and then involve them in clinical decision making. To arrive at individually tailored predictions, psychiatrists build up a profile that represents the patient, based on a given catalogue of variables.

A common concern about data-driven predictions is that they rely on tacit and unreflective assumptions in the machine learning methods (Beijers et al. 2020). A related concern, and one that has received less attention in the philosophy of science, is that the data do not carry enough information. Our paper directs attention precisely there. The worry is that, while rich in implicit assumptions, the data are nevertheless too poor for bringing psychiatric illness into view. There are many aspects to the doctor-patient interaction that do not get represented in the predictive system, while at the same time they are vital to an adequate understanding of the patient's situation and potential recovery.

To substantiate these claims, we investigate the concrete practice of predictions in the psychiatric clinic. When diagnosing problems and determining therapeutic interventions psychiatrists build up a narrative for their patients. These narratives include variables that can be found back in the psychiatric data structures but they often contain much richer content: they indicate the causal background of the patient's problems, and they provide specific meanings to data items about the patient. In the conversion from clinical practice and the narratives that capture doctor-patient interactions to an SPSS file, these finer semantic and causal details about the patient are at risk of being lost. Furthermore, it is simply impossible to convert the vast number of features and aspects of the clinical reality that may have some relevance to the case – curious behaviours, abnormal speech, disproportionate reactions, etc. – into a manageable data set.

Our primary conclusion is that, if we want to rely on data-driven predictions in the clinic and in research psychiatry, we are in need of what Geertz (1973/2000) calls "thick descriptions", i.e., highly detailed data that includes intentions and meanings, over and above the direct observational facts. Now admittedly some thickening of clinical descriptions may be accomplished relatively easily, by inviting designers of medical data infrastructure to "keep good contact with the clinical context". There are, on the other hand, obvious limits to this, e.g., variation in what the personal specifics mean to a patient is much harder to accommodate in database infrastructure. The paper delves into such considerations, and offers an indication of how far we might get in converting thick descriptions into data sets. The discussion is set against the background of a much older debate on the nature of explanation and understanding in the human sciences, and the more recent debates on the promise of machine learning in medicine.

ID: 239 / Keynote 3: 1
Individual paper
Topics: Critical Theory, Hermeneutics, Phenomenology
Keywords: tba

TBA

Anne Sauvagnargues

Tba

ID: 226 / **Panel 1-3-C: 1**
Symposium paper
Topics: Ethics (general and applied), Metaethics, Moral Psychology
Keywords: psychiatry, mental health, bioethics, philosophy of mind, moral psychology

Developing Moral Resilience: A Matter of Self-Care or a Condition of Moral Agency?

Katharina Bauer

Erasmus University Rotterdam

The term "moral resilience" is "a concept under construction" (Rushton 2016, 112), which is currently mainly discussed in nursing ethics, as the capacity of caregivers to cope with moral distress (Rushton 2018, Lachman 2016, Lützén & Ewalds Kvist 2013). In this context, moral distress is defined as "the anger, frustration, and anxiety of nurses who believed their ability to sustain moral integrity in their work was compromised by institutional pressure and constraints" (Rushton 2018, 24f.). Moral distress has been shown to generate feelings of powerlessness, isolation, and frustration, leading to diminished moral responsiveness, disillusionment, a sense of moral deficiency and shame. It can imply an experience of moral failure, but also refer to the experience of uncertainty and moral concerns in relation to complex moral issues, for example concerning decisions on life-prolonging measurements, or conflicts between medical personal, patients, and their families. In addition, moral distress encompasses the 'moral costs' individual care-workers incur due to problematic structural conditions, such as authority structures, a lack of participation in decision-making procedures, or scarcity of resources. Comparable moral conflicts and institutional challenges occur in society at large (Bauer & Hermann 2022). Moral resilience can thus be regarded as a fundamental condition for sustaining mental health under morally challenging or corrupting conditions.

In my panel contribution, I will redefine moral resilience as the capacity to regain a new form of stability as a competent moral agent after experiences of destabilization. In doing so, I aim to reactivate the connection between moral resilience and resistance against morally corrupting circumstances. I will reflect on the (metaphorical) connections between moral integrity and mental health, and discuss the implications and limits of understanding moral resilience as a matter of self-care or as condition of moral agency.

ID: 224 / **Panel 1-1-C: 1**
Symposium paper
Topics: Ethics (general and applied), Metaethics, Moral Psychology
Keywords: psychiatry, mental health, bioethics, philosophy of mind, moral psychology

The Empathy Mirror

Daphne Brandenburg
University of Groningen

Empathy is understood in a number of different ways. It may refer to emotional contagion, perspective taking, mind-reading, or an affinity with others. How empathy in turn relates to moral agency will depend on one's preferred moral and meta-ethical theory. It has been understood as facilitating moral agency, obstructing moral agency, essential to the development of moral agency, and as essential or even equal to the exercise of moral agency. As a result, the relationship between so called empathy deficits and moral agency is complicated and easily misunderstood. To illustrate, mind-reading deficits have often been confused with a lack of affinity with others and a related impairment of moral agency.

These misunderstandings are especially harmful to persons who have received a diagnosis where empathy-deficits are listed among the symptoms. Most commonly, though not exclusively or necessarily, this concerns persons diagnosed with autism spectrum disorder. Ironically, individuals diagnosed with autism have throughout history been misunderstood because others failed to empathize with them.

Contemporary literature is divided about the impact of empathy failures on moral agency. David Shoemaker maintains that persons diagnosed with autism have limited to no moral regard for others, where moral regard refers to a capacity for genuine affective engagement with someone else's normative perspectives. Jeanette Kennett argues that empathy failures are irrelevant to being moved by a reverence for moral principles, which is arguably the core moral motive and at the very least suffices for genuine moral agency. Victoria McGeer disagrees with this, and contends that moral agency is motivated by at least three different types of affective concern that give rise to different evaluative systems, one of which is impaired in persons who struggle to empathize.

In this talk, I first provide an overview of how autism, empathy failures, moral agency, and the relationship between these three concepts have been (mis)understood in this debate. I then provide an alternative account of these failures and their impact on moral agency. An important upshot of this account is that these difficulties in empathizing are not unique to individuals diagnosed with autism but central to all diverse moral communities, and that they need not obstruct moral agency if we recognize this and accommodate for it.

ID: 125
Full symposium
Topics: Ethics (general and applied), Metaethics, Moral Psychology
Keywords: psychiatry, mental health, bioethics, philosophy of mind, moral psychology

Philosophy of Psychiatry

Chair(s): **Matthé Scholten** (Ruhr University Bochum)

Philosophy of psychiatry is an emerging and quickly expanding field of research investigating a wide array of issues surrounding the foundations of psychiatry as a science, the definition of mental disorder, the explanations of psychopathology, and the ethics of clinical practice. The field is highly interdisciplinary, integrating methods from the humanities, social sciences and health sciences. The aim of this panel is to inform philosophers working in applied ethics and philosophy of mind about current controversies and emerging debates in the philosophy of psychiatry and to urge them to engage with the field by presenting a selection of papers on pressing conceptual and ethical issues.

The first two presentations take issue with the conceptualization and implications of impaired social abilities in people with mental disorders. Daphne Brandenburg focuses on so-called empathy deficits in people with autism spectrum disorders and their implications for moral agency. She argues that failures of empathy are not unique to people with autism and need not undermine moral agency provided that reasonable accommodation is made.

Linde van Schuppen focuses on alleged Theory of Mind (ToM) and intersubjective deficits in people who are diagnosed with disorders on the schizophrenia spectrum. Based on empirical findings indicating that intersubjective and ToM-based accounts of schizophrenia insufficiently recognize the perspective-taking abilities of people with schizophrenia in natural and embodied social interactions, she questions the validity of the current scientific conceptualization of schizophrenia and proposes to rethink our methods for studying mental disorders with an eye to mental health stigma.

Katharina Bauer focuses on the notion of moral resilience and its implications for moral agency. Moral resilience refers to the ability to sustain mental health in the face of moral distress, where moral distress refers to a feeling of powerlessness and uncertainty in relation to complex moral issues or problematic structural conditions. While moral resilience is commonly considered a specific competency of health professionals, Katharina Bauer proposes to broaden the scope and to conceptualize it as the stability of a competent moral agent in the face of destabilizing experiences within a morally complex and structurally unjust society.

Matthé Scholten takes a position in a conceptual and ethical debate over the enforceability of self-binding directives (SBDs). SBDs aim to enable people with bipolar and psychotic disorders to maintain control over their life and treatment by giving them the opportunity to give advance consent to involuntary hospital admission and treatment during future mental health crises. Various authors have argued that SBDs are necessarily invalid and non-enforceable based on an analogy with John Stuart Mill's so-called slavery exception. Matthé Scholten argues that this analogy does not hold true as long as the provision of involuntary treatment based on an SBD remains tied to the requirement that service users lack competence to consent with respect to the treatment decision at hand.

Presentations of the Symposium

The Empathy Mirror

Daphne Brandenburg
University of Groningen

Empathy is understood in a number of different ways. It may refer to emotional contagion, perspective taking, mind-reading, or an affinity with others. How empathy in turn relates to moral agency will depend on one's preferred moral and meta-ethical theory. It has been understood as facilitating moral agency, obstructing moral agency, essential to the development of moral agency, and as essential or even equal to the exercise of moral agency. As a result, the relationship between so called empathy deficits and moral agency is complicated and easily misunderstood. To illustrate, mind-reading deficits have often been confused with a lack of affinity with others and a related impairment of moral agency.

These misunderstandings are especially harmful to persons who have received a diagnosis where empathy-deficits are listed among the symptoms. Most commonly, though not exclusively or necessarily, this

concerns persons diagnosed with autism spectrum disorder. Ironically, individuals diagnosed with autism have throughout history been misunderstood because others failed to empathize with them.

Contemporary literature is divided about the impact of empathy failures on moral agency. David Shoemaker maintains that persons diagnosed with autism have limited to no moral regard for others, where moral regard refers to a capacity for genuine affective engagement with someone else's normative perspectives. Jeanette Kennett argues that empathy failures are irrelevant to being moved by a reverence for moral principles, which is arguably the core moral motive and at the very least suffices for genuine moral agency. Victoria McGeer disagrees with this, and contends that moral agency is motivated by at least three different types of affective concern that give rise to different evaluative systems, one of which is impaired in persons who struggle to empathize.

In this talk, I first provide an overview of how autism, empathy failures, moral agency, and the relationship between these three concepts have been (mis)understood in this debate. I then provide an alternative account of these failures and their impact on moral agency. An important upshot of this account is that these difficulties in empathizing are not unique to individuals diagnosed with autism but central to all diverse moral communities, and that they need not obstruct moral agency if we recognize this and accommodate for it.

Conceptualizing Schizophrenia in Science: The Case of Theory of Mind and Intersubjectivity

Linde van Schuppen
Radboud University Nijmegen

Accounts of schizophrenia in cognitive psychology and phenomenological psychiatry which propose a Theory of Mind (ToM) or intersubjectivity deficit at the center of the disorder (Brüne, 2005; Frith, 2004; Fuchs, 2015; Sass & Pienkos, 2015) have received increasing attention during the past decade. Although these accounts are supported by phenomenological analyses of first person reports and empirical evidence of ToM impairments in schizophrenia, there are reasons to doubt their validity.

Firstly, explaining schizophrenia through one central deficit seems to build on essentialist assumptions that insufficiently acknowledge the heterogeneity between and within people with this diagnosis (van Os, 2016). It is not clear how the interpretation of a limited number of first-person reports generalizes to the broader category of 'schizophrenia'. Secondly, evidence for a ToM deficit is mostly based on experimental ToM tasks that do not involve real-time embodied interaction. Evidence from ecological studies that were based on data from natural, embodied interactions with other people have found participants with a schizophrenia diagnosis to be able to accurately explicate the inner worlds of other people, as well as skillfully navigate perspectives in spoken narrative (McCabe et al., 2004; van Schuppen et al., forthcoming). These findings indicate that intersubjective or ToM-based accounts of schizophrenia, especially those that assume it to be a trait deficiency, insufficiently acknowledge that perspective-taking might be situationally scaffolded and fluctuating over time.

In this talk, I will reflect on whether the current scientific conceptualization of 'schizophrenia' as a mental disorder does justice to the embodied and everyday reality of the people with this diagnosis. In current clinical practice, diagnostic classification is often discarded in favor of a more symptom-oriented approach. Labels like 'sensitivity to psychosis' are preferred to the stigma-carrying diagnosis of 'schizophrenia'. This stands in stark contrast to the scientific literature, which still uses schizophrenia as an explanandum.

Taking the conceptualization of schizophrenia as a ToM or intersubjectivity deficit as an example, I argue that we need to rethink the way in which we study psychiatric phenomena in academia in general, and schizophrenia in particular. This holds especially for conceptualizations that center around social skills, since these can feedback into already existing prejudice about people with schizophrenia as 'unreachable'.

Developing Moral Resilience: A Matter of Self-Care or a Condition of Moral Agency?

Katharina Bauer
Erasmus University Rotterdam

The term "moral resilience" is "a concept under construction" (Rushton 2016, 112), which is currently mainly discussed in nursing ethics, as the capacity of caregivers to cope with moral distress (Rushton 2018, Lachman 2016, Lützén & Ewalds Kvist 2013). In this context, moral distress is defined as "the anger, frustration, and anxiety of nurses who believed their ability to sustain moral integrity in their work was compromised by institutional pressure and constraints" (Rushton 2018, 24f.). Moral distress has been shown to generate feelings of powerlessness, isolation, and frustration, leading to diminished moral responsiveness, disillusionment, a sense of moral deficiency and shame. It can imply an experience of moral

failure, but also refer to the experience of uncertainty and moral concerns in relation to complex moral issues, for example concerning decisions on life-prolonging measurements, or conflicts between medical personal, patients, and their families. In addition, moral distress encompasses the 'moral costs' individual care-workers incur due to problematic structural conditions, such as authority structures, a lack of participation in decision-making procedures, or scarcity of resources. Comparable moral conflicts and institutional challenges occur in society at large (Bauer & Hermann 2022). Moral resilience can thus be regarded as a fundamental condition for sustaining mental health under morally challenging or corrupting conditions.

In my panel contribution, I will redefine moral resilience as the capacity to regain a new form of stability as a competent moral agent after experiences of destabilization. In doing so, I aim to reactivate the connection between moral resilience and resistance against morally corrupting circumstances. I will reflect on the (metaphorical) connections between moral integrity and mental health, and discuss the implications and limits of understanding moral resilience as a matter of self-care or as condition of moral agency.

Self-binding Directives and Mill's Slavery Objection: Why we should allow Mental Health Service Users to give advance Consent to involuntary Treatment

Matthé Scholten

Ruhr University Bochum

Psychiatric advance directives are documents by means of which mental health service users can express their treatment preferences for future mental health crises. Psychiatric advance directives receive strong support from service users and clinicians alike and have been shown to significantly reduce the rates of involuntary hospital admission. Self-binding directives (SBDs) are psychiatric advance directives that include a clause in which service users consent in advance to involuntary hospital admission and treatment under specified conditions. These instruments can be helpful for people with episodic mental disorders involving a high likelihood of treatment refusals during mental health crises, such as bipolar and psychotic disorders. Legal provisions for SBDs exist in the Netherlands and various states in the US, but completion rates for SBDs have remained low. While there has been an intense debate among ethicists and legal scholars about the benefits and risks of SBDs, only recently the views of stakeholders have been investigated systematically.

One prominent issue in debates among academics on SBDs is whether the choice of service users to give advance consent to involuntary treatment is a legitimate exercise of freedom which should be enforced by the state. John Stuart Mill once argued that slavery contracts are void and non-enforceable because in the contract the would-be slave forfeits the very liberty that underlies the validity of the contract. By analogy, academics have argued that SBDs are necessarily void and non-enforceable.

In this presentation, I will argue that SBDs can be valid and enforceable as long as their application (i.e., the provision of involuntary treatment based on the SBD) is subject to the requirement that service users lack competence to consent regarding the treatment decision at hand. Competence to consent refers to the ability of service users to understand the potential consequences of the available treatment options, appreciate their impact on their own situation, evaluate their pros and cons in light of their own values and preferences, and express a treatment choice. Competence to consent can be assessed reliably by means of instruments such as the MacArthur Competence Assessment Tool for Treatment (MacCAT-T). I will base my arguments not only on conceptual analysis but also on the findings from extensive stakeholder research on SBDs which my colleagues and I carried out in Germany, the Netherlands, and the UK.

ID: 227 / **Panel 2-1-C: 1**
Symposium paper
Topics: Ethics (general and applied), Metaethics, Moral Psychology
Keywords: psychiatry, mental health, bioethics, philosophy of mind, moral psychology

Self-binding Directives and Mill's Slavery Objection: Why we should allow Mental Health Service Users to give advance Consent to involuntary Treatment

Matthé Scholten

Ruhr University Bochum

Psychiatric advance directives are documents by means of which mental health service users can express their treatment preferences for future mental health crises. Psychiatric advance directives receive strong support from service users and clinicians alike and have been shown to significantly reduce the rates of involuntary hospital admission. Self-binding directives (SBDs) are psychiatric advance directives that include a clause in which service users consent in advance to involuntary hospital admission and treatment under specified conditions. These instruments can be helpful for people with episodic mental disorders involving a high likelihood of treatment refusals during mental health crises, such as bipolar and psychotic disorders. Legal provisions for SBDs exist in the Netherlands and various states in the US, but completion rates for SBDs have remained low. While there has been an intense debate among ethicists and legal scholars about the benefits and risks of SBDs, only recently the views of stakeholders have been investigated systematically.

One prominent issue in debates among academics on SBDs is whether the choice of service users to give advance consent to involuntary treatment is a legitimate exercise of freedom which should be enforced by the state. John Stuart Mill once argued that slavery contracts are void and non-enforceable because in the contract the would-be slave forfeits the very liberty that underlies the validity of the contract. By analogy, academics have argued that SBDs are necessarily void and non-enforceable.

In this presentation, I will argue that SBDs can be valid and enforceable as long as their application (i.e., the provision of involuntary treatment based on the SBD) is subject to the requirement that service users lack competence to consent regarding the treatment decision at hand. Competence to consent refers to the ability of service users to understand the potential consequences of the available treatment options, appreciate their impact on their own situation, evaluate their pros and cons in light of their own values and preferences, and express a treatment choice. Competence to consent can be assessed reliably by means of instruments such as the MacArthur Competence Assessment Tool for Treatment (MacCAT-T). I will base my arguments not only on conceptual analysis but also on the findings from extensive stakeholder research on SBDs which my colleagues and I carried out in Germany, the Netherlands, and the UK.

ID: 225 / **Panel 1-2-C: 1**
Symposium paper
Topics: Ethics (general and applied), Metaethics, Moral Psychology
Keywords: psychiatry, mental health, bioethics, philosophy of mind, moral psychology

Conceptualizing Schizophrenia in Science: The Case of Theory of Mind and Intersubjectivity

Linde van Schuppen

Radboud University Nijmegen

Accounts of schizophrenia in cognitive psychology and phenomenological psychiatry which propose a Theory of Mind (ToM) or intersubjectivity deficit at the center of the disorder (Brüne, 2005; Frith, 2004; Fuchs, 2015; Sass & Pienkos, 2015) have received increasing attention during the past decade. Although these accounts are supported by phenomenological analyses of first person reports and empirical evidence of ToM impairments in schizophrenia, there are reasons to doubt their validity.

Firstly, explaining schizophrenia through one central deficit seems to build on essentialist assumptions that insufficiently acknowledge the heterogeneity between and within people with this diagnosis (van Os, 2016). It is not clear how the interpretation of a limited number of first-person reports generalizes to the broader category of 'schizophrenia'. Secondly, evidence for a ToM deficit is mostly based on experimental ToM tasks that do not involve real-time embodied interaction. Evidence from ecological studies that were based on data from natural, embodied interactions with other people have found participants with a schizophrenia diagnosis to be able to accurately explicate the inner worlds of other people, as well as skillfully navigate perspectives in spoken narrative (McCabe et al., 2004; van Schuppen et al., forthcoming). These findings indicate that intersubjective or ToM-based accounts of schizophrenia, especially those that assume it to be a trait deficiency, insufficiently acknowledge that perspective-taking might be situationally scaffolded and fluctuating over time.

In this talk, I will reflect on whether the current scientific conceptualization of 'schizophrenia' as a mental disorder does justice to the embodied and everyday reality of the people with this diagnosis. In current clinical practice, diagnostic classification is often discarded in favor of a more symptom-oriented approach. Labels like 'sensitivity to psychosis' are preferred to the stigma-carrying diagnosis of 'schizophrenia'. This stands in stark contrast to the scientific literature, which still uses schizophrenia as an explanandum.

Taking the conceptualization of schizophrenia as a ToM or intersubjectivity deficit as an example, I argue that we need to rethink the way in which we study psychiatric phenomena in academia in general, and schizophrenia in particular. This holds especially for conceptualizations that center around social skills, since these can feedback into already existing prejudice about people with schizophrenia as 'unreachable'.

ID: 114 / **Panel 1-2-E: 1**
Individual paper
Topics: Epistemology
Keywords: abilities, affordances, perception

Perceiving what we can do

Tom Schoonen

In this paper, I will present two related arguments. First of all, I will argue that knowledge of general abilities is epistemically more fundamental than knowledge of specific abilities. (In fact, I will suggest that there is nothing specific to the epistemology of abilities required for knowledge of specific abilities.) Secondly, I will argue that this poses a problem for perception-based epistemologies of abilities. These accounts suggest that we perceive affordances and that this provides us with knowledge of what we can do (e.g., McClelland 2019).

I will show that this concerns knowledge of specific abilities and argue that, in fact, it merely concerns what the environment affords, not what we can do. The problem will be phrased in terms of non-existence anchors. If there are no trees around, how could I know that I can climb trees based on perception?

There are two potential solutions that one might suggest. The first concerns the imagination of affordances and the second concerns bodily perception (i.e., internal perception of one's own body). I will argue that in order for the first solution to work, it has to rely on bodily perception. Thus, I show that knowledge of general abilities is epistemically more fundamental than knowledge of specific abilities and that we can acquire such knowledge via bodily perception.

ID: 202 / Panel 5-3-H: 1
Individual paper
Topics: Critical Theory, Hermeneutics, Phenomenology
Keywords: sophistry, rhetoric, audience; argumentation; practice

Rhetoricity: philosophy & rhetoric after the Badiou-Cassin debate

Blake D. Scott

Despite fundamentally disagreeing over the nature of philosophy and sophistry in their ongoing debate, Alain Badiou and Barbara Cassin both agree that rhetoric is a dead end for contemporary thought. For Badiou, the proud Platonist, rhetoric's concern with audiences and their opinions leaves it entirely divorced from truths. For Cassin, the avowed sophist, rhetoric was never anything more than a philosophical ruse designed to contain the creative power of genuine sophistic discourse. In other words, where Badiou finds rhetoric too sophistical, Cassin finds it too philosophical.

Motivated by the incompatibility of their respective interpretations, this paper asks whether rhetoric might not have more to offer than Badiou and Cassin suggest. Responding to their criticisms, I argue that rhetoric remains an indispensable area of concern for contemporary philosophy. In support of this argument, I draw on the largely neglected "new rhetoric project" of Chaïm Perelman (and Lucie Olbrechts-Tyteca) and the hermeneutic phenomenology of Paul Ricœur. To different degrees, both Perelman and Ricœur manage to suspend philosophy's long-standing prejudice against rhetoric and reconsider the significance of the fact that all discourse is oriented towards an audience.

After clarifying the stakes of Perelman and Ricœur's own debate, I use Ricœur's hermeneutic philosophy to expand Perelman and Olbrechts-Tyteca's technical notion of the "rhetorical audience" – which refers to the specific way in which a discourse constructs its hearer or reader. Rather than continue to oppose philosophy and rhetoric as competing disciplines, this expanded notion of audience allows me to identify what I call the "rhetoricity" of discourse – the rhetorical dimension of human action often obscured by philosophy's deep-rooted mistrust of its ancient rival. What emerges from this investigation, pace Badiou and Cassin, is a picture of rhetoric as (1) a dimension of all discourse and action and (2) an integral capacity of human beings most visible in the reception and production of arguments. Thus conceived, rhetoric plays an essential yet underappreciated role in philosophical and social scientific critique.

The most important consequence of this conception of rhetoric is that it allows us to overcome the traditional philosophical opposition between philosophy and rhetoric. By adopting our view, rhetoric is no longer something "more" added on to philosophy, but a capacity philosophers draw upon in their own practice when they reflect on their audience. By viewing rhetoric as essential to philosophical practice, we avoid the problem posed by Badiou and Cassin's positions: engaging in rhetoric while denying that they are doing so.

ID: 197 / **Panel 5-3-G: 1**
Individual paper
Topics: Ethics (general and applied), Metaethics, Moral Psychology
Keywords: Empirical Ethics, Empirical Philosophy, Standpoint Theory, Justice

To take a stand between the 'goods' and 'bads': a proposal for a prescriptive empirical philosophy

Lotje Elizabeth Siffels

Empirical ethics is an approach that can be loosely understood as "the empirical study of … forms of the good in practice" (Pols 2018; cf. Thévenot 2001). On the one hand, mirroring shifts in bioethics, epistemology, and social studies of science, it seeks to move the field of ethics beyond both a reliance on universal principles and an emphasis on the development of normative criteria for what should count as a 'good' practice (Hedgecoe 2004; Pols 2015). On the other hand, empirical ethics also aims to move the field of ethics away from an emphasis on critique that has long motivated much research in the interpretative social sciences (Jerak-Zuiderent 2015; Latour 2004). Instead, empirical ethics takes as its focus the empirical description of everyday morality. In making this double move, the approach has shifted research questions in promising new directions, resulting in a rich body of scholarship in ethics and philosophy of technology, STS and adjacent fields both in terms of content and methods (Haimes 2002; Lehoux et al. 2012; Mol 2002; Molewijk et al. 2004; Pols 2016; Heeney 2017; Swierstra 2015).

This focus on description comes with an aversion of prescription. Empirical philosophers prefer asking how norms and practices are done in practice rather than prescribing which norms and practices should be upheld. 'The philosopher who enters a field with determination and garnished with criticism, norms or convictions, will notice that the actors have preceded her' (Mol, 2000, 15). Mol describes two movements: 'to bring empiricism into philosophy, and return to philosophy her place in the world' (Mol, 2000, 14) with which it bridges a distance that 'traditional' ethics created: a distance between the philosophical theory and real life. In bridging one distance, however, this rejection of prescription creates a new one: between the empirical philosopher and their subject as they go about describing the practices of "goods" and "bads" as if seeing them from afar, or from above. This paper argues that choosing to take such a distance is a dangerous (Bowker and Star, 1999: 5-6) move, especially when what is at stake is nothing less than societal (in)justice, in which case empirical philosophy still only offers a way of describing different "justices" (Boltanski and Thévenot, 2006), without giving any way of prescribing which justice ought to abide. Methodologically, a purely descriptive approach as this seems too distant, perhaps even bordering on what Haraway (1988) calls a kind of 'God-trick'.

This paper investigates whether standpoint theory may offer a way of bridging this newly formed distance. It proposes that the empirical philosopher can offer a prescriptive reflection on the risks and dangers of a practice, which comes from their positionality, from placing themselves as an actor in a deliberative space of norms and values, evaluating the "goods" and "bads" in practices, without relying on the universal claims of traditional ethics, but by reflection on the different conceptions of justice as relating to their own.

Making and breaking decision boundaries: scope and social policy reasoning under radical uncertainty

Helena Slanickova

The reasoning process underlying social policy decisions is governed by multi-dimensional epistemic standards. Getting clear on these dimensions and what they demand of decision makers will give us a better picture of how we should conduct policy reasoning. The purpose of this paper is to draw attention to a somewhat overlooked dimension of these standards – the scope dimension – which, I argue, is crucial for managing radical uncertainty in policymaking. I do this by contrasting the scope dimension with another highly investigated dimension related to evidence use.

I develop a model of policy reasoning based on two key elements of the process: policy claims (the set of considerations that policymakers accept as true and relevant to the decision over the course of their reasoning) and evidence (the collection of reasons for which policymakers accept policy claims).

I use it to analyse two examples of UK COVID-19 policies and make a distinction between the dimensions of policy reasoning standards.

The first example concerns lockdown and the evidence dimension; policymakers neglected to make sure that the policy claims they relied on were warranted given their evidence. The second example, however, concerns hospital visitations and the scope dimension; policymakers neglected to make sure that they considered an appropriate range, breadth and depth of policy claims. They inadequately determined which considerations were relevant to the reasoning task at hand and unsatisfactorily delineated the boundaries of the decision. The evidence dimension demands that policymakers take care that any policy claims they consider are right, the scope dimension demands that policymakers take care to consider the right policy claims. The philosophical literature about evidence-based policy has led to significant progress on elucidating the evidence dimension. I suggest that there is very little analogous philosophical exploration of the demands of the scope dimension.

I argue that the scope dimension is especially crucial for improving policymaking under radical uncertainty (uncertainty over and above empirical uncertainty). Drawing on resources from sociology and management studies, I identify four characteristics of radically uncertain policymaking scenarios – instability, complexity, novelty and stakes. I discuss the epistemic and practical ways that these characteristics make it more challenging for policymakers to consider the appropriate range, breadth and depth of policy claims and the importance of achieving this for managing radical uncertainty.

Drawing on criticisms of UK COVID-19 policies and work in the social sciences, this paper sheds light on the scope dimension of the standards governing reasoning about social policy decisions and its importance for managing radical uncertainty. It concludes that to improve policy reasoning under radical uncertainty, we first need to explore the demands of the scope dimension and identify the practices, structures and processes that are conducive to meeting it.

ID: 109 / **Panel 1-2-F: 1**
Individual paper
Topics: Ethics (general and applied), Metaethics, Moral Psychology
Keywords: ethics and technology, disclosive computer ethics, empirical ethics, sharing economy, platform economy

Do digital platforms have politics? Making airbnb's ethical and political operations visible

<u>Shaked Spier</u>

Notions of ethical technology design are gaining increasing attention from companies, legislators, researchers, and activists. A central challenge in tackling this issue lies in finding ways to uncover values that are embedded into a technology's design. This involves showing where they are located, understanding their meanings and implications, and determining whether and how they relate to broader institutional and political aspects. As digital platforms pervade many aspects of everyday life, ethical controversies arise. Advocates of the platform and sharing economy emphasize platforms' positive environmental, economic, and social potential, whereas critics point out their working conditions, impact on local communities, bypass of regulations, and neoliberal ideology. Responding to this, platform companies often claim that they are merely intermediaries and therefore not responsible for the platforms' social and political consequences—a response with striking similarity to the neutrality thesis argument.

This paper takes digital platforms' technical design and its intersection with their institutional structures as a starting point. I deploy the disclosive computer ethics approach (DCE) to reconstruct the ethics and politics of one of the mainstream sharing economy's flagships—Airbnb. DCE focuses on identifying and evaluating embedded values, moral and political issues, and normativity in information technologies, applications, and practices; especially when these are morally opaque (Brey, 2000, 2010).

Airbnb is more than just a paradigmatic case study – the platform has a normative impact on sharing economy platforms at large. I investigate Airbnb's technical operations to identify the moral and political values that are embedded in the platform's technology (and the technology's intersection with the platform's institutional structures). Based on the identified values, I analyze what constitutes the platform's ethics and politics from a broader perspective. This perspective enables a generalization of the findings by reconnecting them to the mechanisms, operations, rationales, and ideologies of the overall platform and sharing economy.

This paper adds to the existing literature in three main ways: first, I build on existing DCE approaches to develop an applied research methodology and apply it in practice. Thus, the paper makes a methodological contribution to the literature on the ethics and politics of technology and shows DCE's practical and normative potential. Second, I explore Airbnb as a specific platform technology with a variety of ethical and political dimensions. Thus, the paper contributes to refuting the neutrality thesis and similar claims made by mainstream platforms. Finally, the paper contributes to the growing body of literature on the platform and sharing economy through insights into mainstream platforms' ethics and politics and connecting them to the platforms' technical design.

On two problems for workplace democratization and their solution via a workers' dictatorship

Philipp Stehr

The workplace democracy debate has articulated the idea that workplaces are arbitrary exercises of power by employers that have unjustifiably wide discretion over their employees. The focus of the debate so far has been the normative basis for rejecting such structures and the development of alternatives. Far less attention has been paid to the feasibility of alternatives. With this paper I contribute to filling this gap by building on Marx's idea of the dictatorship of the proletariat. In doing so I shed light on an underappreciated kind of problem of internal stability for democratizing workplaces and I show that transitional dictatorial forms of corporate government can, like the dictatorship of the proletariat, be justified from a democratic point of view.

Existing approaches to workplace democracy only offer static understandings of the establishment of institutional alternatives. They envision reforms in corporate law for example that are supposed to bring about more democratic institutions. But they lack serious consideration of the social conditions under which such change will take place. These social conditions can however threaten the transition to more democratic institutions. In the first part of my paper I offer two examples of such threats. First, democratizing workplaces means giving power to those not used to having it. When the formerly disenfranchised are given power they lack experience in wielding it. This might lead to detrimental outcomes. Second, democratizing workplaces means taking power from those who are used to having it. Former bosses might be unwilling to cooperate with their former underlings, also hampering the transition.

In the second part of my paper I suggest that we therefore need to think more carefully about how to institutionalize democracy in the workplace. If workplace democracy is bound to fail if instituted directly, we need to develop alternatives that are not only more democratic but also resilient to threats so that they enable the transition to a fully democratic alternative. Marx's dictatorship of the proletariat is an example of such a resilient transitional structure. Marx foresaw problems similar to the ones mentioned above for the overthrow of a capitalist society and suggested the dictatorship of the proletariat to deal with them. Lea Ypi has recently reconstructed how Marx's dictatorship of the proletariat can be democratically justified. A corporate dictatorship of the formerly disenfranchised can be justified in the same manner, appealing to its transitional status, its sharply defined goals, and the fact that it is working towards its own superfluousness. Such a dictatorship would employ the powers already available to the directors and managers of a corporation to work for the democratic transition.

ID: 161 / Panel 2-1-E: 1
Individual paper
Topics: Phenomenology
Keywords: technique, deconstruction, ecology, Nancy, Derrida, Stiegler

On articulation: technique before the pharmakon

Donovan Alexander Stewart

A major insight of twentieth-century phenomenology and post-Kantian philosophy was the reformulation of human existence as being fundamentally technical, which helped overcome traditional accounts of an originally pure human life that was subsequently contaminated by technics and history. This thinking of human originary technicity, as recently summarised by Lindberg (2023), Crowley (2022), James (2019), and Hörl (2017), was offered by Derrida in his early (1954, 1962, 1967) readings of Husserl (1929, 1939), and later developed by Stiegler (1994) and Wills (1995), and was elsewhere articulated by Foucault (1976), Deleuze and Guattari (1980), Haraway (1985), Preciado (2000) and Agamben (2014). The consensus: human existence is essentially composed by finite, historically contingent, prostheses, pharmaka with toxic and curative potentials, beginning with the languages within which it becomes itself, has a world, and experiences the earth. In this text, I wish to push this thinking of technique further, emphasising not the question of technical objects, but constitutive relation––the finite manner of each singular composition. I suggest that technique, in its most determined sense in English, describes a way, the manner, of not necessarily a being or thing, but an event, that is, a finite relation or contact. I turn to Nancy's (1992, 1996) concept of "ecotechnics" which offers, at times, an understanding of such a sense of technique as the articulation of existence itself––its historical being-put-into-play that is irreducible to human activity and which opens onto a thinking of more-than-human technicity. With Derrida (1967, 2003) and a certain Heidegger (1935), Nancy introduces another moment in the deconstruction of the inherited Western metaphysical understandings of the relation of technique and life: technique as the finite articulation of being, the rising of its limits; offering a renewed sense to originary technicity, the ramifications of which remain to be thought.

ID: 186 / Panel 3-2-B: 1
Individual paper
Topics: Social and political philosophy, Ethics (general and applied), Metaethics, Moral Psychology
Keywords: Instrumental rationality, intrapersonal inefficacy, interpersonal inefficacy

Instrumental rationality in temporally and collectively extended activities

Tessa Elise Supèr

Consider an agent who enjoys eating unhealthy, sugary treats. She knows that eating one treat at most trivially impacts her health, while bringing her great pleasure. However, if she would eat many treats daily, the cumulative effect will negatively impact her health. Examples like these (which I will refer to as "intrapersonal inefficacy cases"), where a series of actions causes a negative outcome, but no action by itself affects this outcome, are ubiquitous in our daily lives, and have raised interesting questions about instrumental rationality over time, versus rationality at each moment during that time. Sergio Tenenbaum (2020), for example, defends the idea that an agent who eats many treats daily can be rational each time she does so, and yet still be irrational over time, as she finds her health declining. Chrisoula Andreou (2014; 2023), on the other hand, while agreeing with Tenenbaum that we should take seriously the notion of temporally extended activity, as distinct from momentary actions, is skeptical about his suggestion regarding rationality at a time versus over time. Instead, she argues that while an activity (e.g., "ruining her health") is temporally extended beyond a particular moment in time, it is still the case that this activity is going on "at" that moment. Therefore, insofar as "ruining her health" is considered irrational, it can be said that the agent, at any moment when she enjoys a treat, is also ruining her health, and therefore can be considered irrational at that moment.

As has been observed by other philosophers, Andreou acknowledges that "intrapersonal inefficacy cases" share structural similarities with "interpersonal inefficacy cases", in which many people's acts combined lead to a certain outcome, but none of the acts individually influences this outcome. Given these similarities, Andreou extends her proposal from intra- to interpersonal cases. For example, activities such as "damaging the environment", even though extended beyond any individual's action, can be going on while one person, say, "drives her gas-guzzling SUV".

However, there seem to be various (metaphysical) disanalogies between these types of cases. In this paper, I will examine to what extent these differences threaten the application of Andreou's account to interpersonal cases. For example, the agent who is "ruining her health" is the same agent who "enjoys a treat", while "damaging the environment" involves many other agents. It is thus unclear if the latter activity should be ascribed to the single agent who "drives her SUV", or to a collective of agents who act in the same way. I will argue that if she accepts the latter, Andreou's arguments cannot be united with her skepticism about Tenenbaum's claims, while if she accepts the former, this would commit her to a view about time-sliced agency that she explicitly rejects. Finally, I will utilize elements from the work of Michael Bratman on instrumental rationality and self-governance (e.g., 2018), in order to enrich Andreou's account. This will allow Andreou's insights to be applied to both intra- and interpersonal inefficacy cases, without running into the dilemma posed above.

ID: 171 / Panel 1-1-A: 1
Individual paper
Topics: Epistemology, Social and political philosophy
Keywords: social epistemology, political epistemology, physical space, epistemic diversity, epistemic justice

A Political Epistemology of Space: How we can Foster Deliberation by Altering the Physical Environment

Merel Melika Talbi

According to some, we are living in uniquely polarized times. This leads Talisse (2020) to despair: given our great political differences, deliberation on political questions is made impossible, due to our mutual distrust and distaste for each other's political beliefs. This is unfortunate from an epistemic perspective since someone like Landemore (2017) insist that deliberation is fundamental to knowledge production, especially when those who contribute to that deliberation have different ideas, opinions or identities.

The idea that deliberation and argumentation, or the giving and asking for reasons, will lead to better and more robust knowledge, is thus thwarted somewhat by situations of adversariality or conflict, where winning an argument, rather than learning, is the main focus of the communicative exchange (Dutilh Novaes 2020). However, this negative effect may in some cases be mitigated or even prevented by changing the setting or context of the deliberative exchange, such as when we employ certain communicative interventions when dealing with exclusion in political discussions (Young 2002).

One avenue that has not yet been explored in much political or social epistemology, is how physical or material space may affect the course of deliberation, and how it might foster the exchange of reasons and arguments, even in situations of polarization. We might wonder how the setting of a room or the design of an urban locale, may either hinder or ameliorate social processes – and through it, the exchange of reasons and production of knowledge. The re-appreciation of matter is embraced in Science Studies, by authors such as Latour (2007) and Barad (2003). It has also been explored in Honig's work on the importance of public things, where Honig argues that public things are crucial for the proper functioning of a democracy (Honig 2017). Kukla (2022), finally, has recently explored how the physical spaces of cityscapes allow for particular interpersonal interactions. I argue that this, in turn, will also lead to different kinds of epistemic results.

In this talk, I explore how spaces of deliberation may work to foster the exchange of arguments and information in a constructive and productive way. Assuming, from the work of Latour and Barad, that our interactions are at least in part shaped by our material surroundings, I will draw on Kukla's analysis of cityscapes to argue that particular kinds of physical spaces allow for certain kinds of knowledge to be produced, while the production of different kinds of knowledge might be obstructed. Additionally, I will join Honig in addressing how we might ensure these spaces are truly public spaces, thus fostering particularly democratic processes. I will illustrate this using a case-study from Amsterdam and New York based neighborhood courts, a practice where judges adjudicate small cases in very local settings. These processes show how the particular knowledge that is produced and exchanged in these contexts is fundamentally different from what occurs in regular courtrooms, and how this practice might be understood as fostering democratic deliberative practices through the constructive use of space.

ID: 175 / **Panel 2-3-C: 1**
Individual paper
Topics: Ethics (general and applied), Metaethics, Moral Psychology
Keywords: Neuroethics, Mental integrity, Extended mind, Neurotechnologies

The right to mental integrity and the extended mind

Vera Tesink

Neurotechnologies allegedly form a unique threat to our minds. Some scholars have proposed a right to 'mental integrity' to protect against mental interferences by neurotechnologies. Considering the vague boundaries of what constitutes the 'mental,' an ensuing question is what would fall within the protective scope of such a right. Generally, the proposals for a right to mental integrity assume a brain-based conception of mental states, and accordingly, propose a right that protects what exists within the skull. However, an alternative conception of mental states as things that can extend beyond our skull may have very different implications for its protective scope. In this article, we will contrast a more traditional, brain-based view of the mind with an extended view of the mind and consider its implications for (the scope of) the right to mental integrity. We find that an extended view of the mind would imply a (perhaps implausibly) large protective scope of the right to mental integrity—not only protecting our internal mental states, but also external artifacts that may be constitutive of mental states, such as brain implants, or even phones or notebooks. Because such external artifacts would also be liable to mental integrity infringements, neurotechnologies would no longer constitute a unique threat to our mental integrity. Moreover, if our minds can extend beyond our internal states, this might suggest that neurotechnologies could become part of our minds, and therefore the subject of mental integrity infringements themselves. We conclude that a right to mental integrity would only sufficiently protect against neurotechnologies on a brain-based account of the mind. On an extended account, the right would either be too broad, counting neurotechnologies amongst the many interferences that could infringe mental integrity, or too narrow, by preventing neurotechnologies from infringing mental integrity altogether by becoming part of the mind themselves.

ID: 138 / Panel 3-1-E: 1
Individual paper
Topics: Philosophy of Language
Keywords: experimental semantics, causal theory of reference, descriptivism

Testing the reference of not just proper names

Jeske Toorman, Jussi Haukioja

Experimental work in the philosophy of language has been taken to show that there exists cross-cultural, intra-cultural, and intra-individual variation in whether test subjects take proper names and natural kind terms to refer in accordance with classical descriptivism or the causal historical account. The most common setup used in these studies is one that is modelled on Kripke's Gödel case. The vignettes and probe questions used in these setups have been taken to contain multiple problems by virtue of which the experimental data gained by means of these setups might fail to track a semantically interesting phenomenon. Nevertheless, even when these problems are accounted for, the original studies seem to replicate, leaving the impression that variation in reference assignments of a sort that is of relevance to theories of reference is real.

We suspected that these setups are subject to another hitherto unaccounted for problem, which for us was the motivation for a new follow-up experimental study. In our study, we created analogs of many of the vignettes and probe questions previously used to test theories of reference. But where the original studies used proper names and natural kind terms, we have used terms for which we have independent reason to believe the causal-historical account is not true of them. In all cases, we were able to replicate the analog of the original findings. Thus, in all cases, test subjects gave the same proportion of supposed causal-historical responses. Our results put pressure on the idea that data gained by means of these setups tracks anything of relevance to theories of reference. But what they show, in any case, is that whatever the data gained by means of these setups tracks, it is nothing that is specific to proper names or natural kind terms.

ID: 153 / **Panel 5-3-D: 1**
Individual paper
Topics: Epistemology, Ethics (general and applied), Metaethics, Moral Psychology, History of philosophy (from ancient to modern)
Keywords: Morality, nihilism, dogmatism, experiment, Nietzsche

Beyond nihilism: Nietzsche on experimental moral truth

Luca Gerard Pompeo Tripaldelli

This paper aims to investigate Nietzsche's stance on moral truth. In popular reception, Nietzsche is often understood as 'the wrecker of morality' due to his critique of Christian values. In line with his general criticism of dogmatic truth, one can wonder if Nietzsche leaves any room for morality at all. Consequently, many people have characterized his philosophy as nihilistic. However, within Die Fröhliche Wissenschaft, Nietzsche himself also warns us of the potential 'darkness' that Europe will find itself in when Christian values are destroyed. Thus, he also refers to the danger of nihilism, which implies that Nietzsche distances himself from the idea of a value-free society.

Instead, by reading some crucial passages within Die Fröhliche Wissenschaft, this paper shows that the denouncement of Christian values is not the end of his story. This paper does so by first analyzing how Nietzsche's critique of truth refers to dogmatism. Instead of criticizing truth in its entirety, Nietzsche aims to replace dogmatic truth with 'the experiment.' Hence, for Nietzsche, we must keep on taking an experimental stance to reshape what we accept as truth. Subsequently, this paper shows that 'the experiment' can also be applied to moral truth. As such, it will become clear that Nietzsche's critique concerns the dogmatism behind 'traditional' morality. Instead of upholding a static moral system, he argues that we must take the experimental stance to inquire into new possible values.

Ultimately, this paper shows that this critique of old values does not mean that we must completely break with these old values. Next to creating new values, we must test the acceptability of the old ones. Hence, through Nietzsche, we can learn how we can break with dogmatism without letting go of moral truth itself. Put differently, his philosophy shows the value of morals without accepting their fixed status. Hence, Nietzsche's morality explains why secular societies still accept that murder is bad and at the same time have changed their stance on abortus or the rights of the LBTQ+ community.

Moreover, this paper shows how Nietzsche's philosophy is a better alternative to classical morality as 'the experiment' keeps on aiming to 'improve' itself by keeping an open stance towards existence. In short, Nietzsche is read as an experimental deontologist. As opposed to traditional deontology, which seeks to deduce an absolute code of conduct, Nietzsche shows that we must keep up the aim to ameliorate the rules we live by.

ID: 113 / **Panel 1-1-E: 1**
Individual paper
Topics: Epistemology, None of the topics above (please include your topic under 'keywords')
Keywords: 4E cognition, scaling-up problem, Ecological Dynamics, neural reuse, reasoning

Reasoning by Doing: using LEGO to think

<u>Menno Michiel van Calcar</u>

The dominant theory in cognitive science, cognitivism, conceives of cognition as the neuronally realized internal processing of representations. However, several decades of research have paved the way for conceptualizing cognition as extended, embedded, embodied and enactive (4E). 4E is praised for its explanations of so-called lower order cognitive activities, such as catching a ball. At the same time, many are not convinced of 4E's ability to scale up to explaining higher order cognitive activities, such as reasoning. The paradigm of cognition, quietly sitting and thinking, without a noticeable role for perception and action, seems to be thoroughly isolated, disembodied and passive. The special status of higher order thought, as opposed to the ontological continuity of other kinds of activity with the rest of nature, is awkward, to say the least.

I propose that we can conceive of higher order cognitive activities as co-constituted by embodiment and action. To that purpose, I will argue that Kiverstein and Rietveld's ecological-dynamic proposal, that we should conceive of linguistic thought as skillful engagement with enlanguaged affordances, points us in the right direction (Kiverstein & Rietveld, 2021). I add that we need agent-specific adaptable elements in the cognitive loop to explain the phenomenon of individual skill-enhancement. These elements are neural structures, which get adapted or exapted to the smallest extent needed (since nature doesn't invent when it can adapt). I argue that such a conception affords explaining higher order thought as typically based on more basic cognitive skills, and that the characteristics of those more basic skills often influence the less basic skills as cortical biases (Anderson, 2010). This influence forms traces in our cognitive activities. Lakoff and Johnson's famous conceptual metaphors, for instance, can be seen as signifying traces of the basic materials used (and thus not as examples of conceptual mapping, Lakoff & Johnson, 2003). That means that our basic bodily engagement with the world co-constitutes higher order thought.

I will focus on one form of linguistic thought, namely deductive reasoning, and will argue that deductive reasoning is built upon our more basic motoric skills in building. I argue that our early skill at stacking, balancing, etc., is reused when we develop the skill to evaluate validity. I have designed a test to pry out this basis, in novices in deductive reasoning, high school pupils. The results of a pilot test I ran suggest that we should favor the ecological-dynamic view over the cognitivist view of deductive reasoning. Pupils who motorically engaged with premises and conclusions (using LEGO) outperformed their control group. This points us toward further exploration of the enactive roots of higher order cognition.

Anderson, M.L. (2010). Neural reuse: A fundamental organizational principle of the brain.

Behavioural and Brain Sciences, 33, 245-66.

Kiverstein, J,. and Rietveld, E. (2021). Scaling-up skilled intentionality to linguistic thought.

Synthese 198 (suppl.), S175-94.

Lakoff, G, and Johnson, M. (2003). Metaphors We Live By. The University of Chicago Press.

ID: 194 / **Panel 4-2-B: 1**
Individual paper
Topics: Aesthetics and philosophy of art
Keywords: fiction, imagination, intention, silly questions, interpretation, digital games, videogames, glitches

Unintentional fiction in videogames: a serious take on 'silly questions'

Nele Van de Mosselaer

Within aesthetics, the dominant opinion is that something is fictionally true in a work if the audience of this work is intended to imagine it (cf. Currie 1990). Yet, works of fiction often seem to represent events that their creators did not intend to be imagined. In Pirates of the Caribbean, for example, a stray crew member, wearing sunglasses and a cowboy hat, can be seen amidst the pirates on the Black Pearl. Why is he on the ship?

Philosophers of fiction have called such questions 'silly questions'. These inquiries call attention to inconsistencies within represented content, but should not be taken seriously when interpreting this content (Walton 1990, 176). The problems they address can simply be avoided by applying the charity principle: we should ignore inconsistencies such as the stray crew member based on our recognition that the author did not intend them to be part of the fictional content (Walton 1990, 183; Matravers 2014, 131).

Instead of disregarding silly questions, I will stress their value for thinking about processes of fiction interpretation. My presentation focuses on silly questions that have not yet received much attention: ones that are raised by computer-generated, interactive works of fiction like videogames. By seriously discussing three silly questions, I demonstrate how videogames challenge the idea that fictional truth is based on authorial intent.

1) Why is baby Ocelotte invisible?

In Dark Souls III, the mad character Oceiros can be seen brutally killing his son Ocelotte. Ocelotte, however, is invisible to players and seemingly does not even exist. By accessing unpublished, earlier versions of Dark Souls III, players discovered that Ocelotte's 3D-model was originally part of the game, but was removed last-minute, likely for reasons of censorship. This information, discovered by approaching the game in unintended ways, raises the question: Should players imagine Ocelotte as a visible baby, an invisible baby, or a mere hallucination of Oceiros?

2) What are these beastly cowboys?

In an early version of Red Dead Redemption, so-called manimals (human characters who behaved like birds or cougars) appeared in the Wild West gameworld. This was caused by an unintentional mistake in the game code. Yet how could players disregard the fictional presence of manimals, while these beasts attacked other (canonically fictional) characters, including the avatar?

3) How did Mario jump that high?

When DOTA_Teabag was playing Super Mario 64, Mario suddenly jumped ten times higher than usual. This event was caused by stray cosmic radiation causing a 'bit-flip' (turning one of Mario's 'height' values from 0 into 1, see Veritasium 2021). Mario's extraordinary jump was not intended to be possible by the game's designers and was not even caused by how the game was coded. And yet Mario did it. Is this not a clear case of unintentional fiction?

References

Currie, G. 1990. The Nature of Fiction. Cambridge University Press.

Matravers, D. 2014. Fiction and Narrative. Oxford University Press.

Veritasium. 2021. 'The Universe is Hostile to Computers.' Youtube video. https://www.youtube.com/watch?v=AaZ_RSt0KP8&t=194s

Walton, K. 1990. Mimesis as Make-Believe. Harvard University Press.

ID: 159 / Panel 4-3-E: 1
Individual paper
Topics: Epistemology, Philosophy of Language
Keywords: representation, pictures, content, color perception

Can you fail to recognize what a picture shows?

Brandt van der Gaast

This paper develops an argument for the conclusion there is a sense in which one cannot fail to recognize what a picture shows.

To argue for this claim, I first develop a way of talking about seeing pictures. First, I distinguish seeing a picture from recognizing what a picture is of. Seeing a picture requires having a good view of what the picture shows, i.e. its content. Recognizing what a picture is of is more involved. It requires knowing, on the basis of what the picture shows, which object or individual is depicted in the picture. This can be easy or difficult, depending on the way something is depicted and depending on which objects and individuals the viewer is acquainted with.

Given this fairly common-sensical way of talking about pictures, it turns out that there is a difference between seeing what a picture shows and recognizing what a picture is of. There is a sense, I argue, in which seeing a picture guarantees knowing what it shows, but not knowing what it is a picture of.

In order for this claim to be interesting, it should be more than just a result from a proposed way of talking about pictures. So, to add more substance, I involve the literature on analog representation, since pictures can be said to represent in an analog manner (Goodman 1968, Haugeland 1998, Kulvicki 2004, Lee/Myers/Rabin 2022). Authors on analog representation claim that the distinctive character of analog representation lies in the fact that it involves a one-to-one correspondence between representational vehicle and content. Pictures and maps and some types of signs and diagrams represent in an analog fashion.

Due to this one-to-one correspondence, analog representation has certain special features. Two of these features are: a lack of ambiguity and a lack of synonymy. There is no ambiguity, unlike in language, because there are no one-to-many relations between representational vehicle and content. There is no synonymy, unlike in language, because there are no many-to-one relations between representational vehicle and content.

The lack of synonymy is responsible for the phenomenon we're interested in: that one cannot fail to recognize what a picture shows. This means that pictorial Frege-style puzzles, where one and the same content (e.g. Venus) is captured by two representational vehicles (e.g. 'Hesperus' and 'Phosphorus'), are not possible. I conclude this discussion of pictorial representation by considering potential counterexamples. Can two colors on a surface not represent one and the same color (e.g. once in full light and once in the shadow)? To answer these objections, we have to decide what exactly we count as part of the content of a picture.

ID: 145 / **Panel 4-2-F: 1**
Individual paper
Topics: None of the topics above (please include your topic under 'keywords'), Ethics (general and applied), Metaethics, Moral Psychology
Keywords: Philosophy of law, temporality, legal practice, digital transformation, technologies, administration of law, adjudication

Slowness of the administration of justice, value or vice? The normative significance of temporality against the backdrop of emerging technologies

Eva Van der Graaf

Digital and automated information technologies are becoming increasingly important in the administration of justice and have the potential to change the everyday practices of legal professionals, in both the judiciary and public administration. The turn to digitalisation and automatisation in the legal domain has often been coupled with presumed efficiency gains: The implementation of digital technologies is promised to reduce costs and the acceleration of legal procedures would be beneficial to increase access to justice.

However, while speed increases in legal procedures have been analysed from the viewpoint of efficiency and access to justice, the normative effects of temporal reconfigurations resulting from increasingly digitally mediated legal practices have been less considered. To what extent does acceleration impact legal practices themselves? For example, how does a turn to the digital affect practices of deliberation and contestation, which are part and parcel to the legal procedure, and, as Latour argued in La fabrique du droit (2002), by nature take time? This paper foregrounds evolving temporal reconfigurations in legal practice and analyses these from the perspective of theories of adjudication and philosophies of law that emphasise the normative and legal value of slowness of the law and of legal decision-making.

A growing body of academic literature points to normatively relevant ramifications of the proliferation of information technologies in legal decision-making, e.g. displacements of discretion and changing interpretation mechanisms due to what emerging technologies afford their users. However, attention to temporal reconfigurations in these analyses has been scarce, and, if such configurations have been noticed (e.g. by Laurence Diver), they are predominantly evaluated in the context of (fully) automated legal procedures and with less attention to the contexts and practices that are vital to the quality of legal judgments.

As new technologies have the potential to radically change and speed up practices in the administration of justice, this paper argues that the substantive normative effects of temporal reconfigurations may signal an undervalued dimension of legal decision-making, which thus far has not sufficiently been taken on board in current debates on the potential benefits and harms of emerging technologies in legal practice.

ID: 152 / Panel 1-3-E: 1
Individual paper
Topics: Epistemology
Keywords: Motivated reasoning, irrelevant influences, belief polarization, bias, Bayesianism, rationality

The skeptical import of motivated reasoning: A closer look at the evidence

Maarten Van Doorn

Empirical work on motivated reasoning suggests that our judgments are influenced to a non-trivial extent by preferences, prior beliefs and identities. For example, a famous study by Lord, Ross and Lepper (1979) gave supporters and opponents of the death penalty the same set of studies for and against capital punishment, and found that their assessments of the strength of these studies correlated with their existing views about the rights and wrongs of capital punishment and that the exposure to the mixed evidence caused the initial disagreements to polarize. Many scholars have concluded that this polarization shows that people process information in a biased manner, so as to support their preexisting views (Baron, 2008; Munro & Ditto, 1997; Ross & Anderson, 1982). These explanations typically emphasize the role of motivated reasoning (Klaczynski, 2000; Kunda, 1990) and suggest that the polarization results from people interpreting information in a biased manner to favor conclusions that they would like to be true.

Indeed, central to much of the recent philosophical interest in motivated reasoning is the idea that it runs afoul of epistemic normativity. For example, Carter and McKenna (2020) argue that motivated reasoning thwarts our assessment of evidence, leading many of us to false beliefs. And Avnur and Scott-Kakures (2015) argue that directional influences constitute evidence of wishful thinking, undermining the beliefs that have been subject to their influence. Here I provide a deeper examination of the empirical evidence underlying these critiques. I argue that this evidence does not convincingly establish that motivated reasoning is epistemically non-normative.

I first make an argument that belief polarization is not a consequence of motivated reasoning, but results from rational metacognitive management of ambiguous evidence (dedicating more investigative resources to scrutinizing purported disconfirming evidence (than to confirming evidence), thereby increasing the chances of finding reasons to dismiss it). Second, contra Avnur and Scott-Kakures (2015), I claim it's highly uncertain whether the use of differential criteria for confirming vs. disconfirming evidence is subjective evidence for desire-based directionally motivated reasoning in the way they suggest. Such asymmetric acceptance thresholds can also be the result of prior beliefs doing their rational job as a Bayesian anchor. I also point out that the wishful thinking effect is weaker than often assumed, rendering the epistemically pernicious impact of directionally influential desires unclear. Third, I note that Carter and McKenna (2020) are too quick to dismiss patterns of information assessment described as biased assimilation. Such patterns are not evidence for motivated reasoning because they derive from study designs that produce results that are not inconsistent with Bayesian principles and do not permit causal inferences about the effect of modes of reasoning on posterior beliefs.

A picture emerges on which it's rational for your initial view and inclinations affecting more than others the view you end up with. Perhaps then we should reconsider the way in which neutrality could be an epistemic virtue.

Keywords: Motivated reasoning, irrelevant influences, belief polarization, bias, Bayesianism, rationality

ID: 144 / Poster session 1: 1
Poster
Topics: Epistemology, Philosophy of Science
Keywords: LLMs, Agency, ChatGPT, GPT3, Intelligence

LLMs and Agency without Intelligence

Maud van Lier

In a very recently published paper, Floridi (2023) claims that with the emergence of Large Language Models (LLMs) like GPT3 and ChatGPT, a new kind of agent has entered the scene: the intelligence-free agent. These systems can learn and act 'correctly' without any intelligence required. In my talk, I argue that calling such systems 'intelligence-free agents' can be misleading for two reasons.

First, calling LLMs 'intelligence-free' agents suggests that such systems can and are functioning in a vacuum. However, the output generated by a LLM is the product of a collaboration between a human user and the system itself. The systems can only function by receiving input in the sense of a 'task' from another agent/system. This input has a necessary and shaping role for the eventual output. And this eventual output of the system might then again shape the new kind of input (when the output is not exactly what the user means or when the output leads to new insights). When agency is understood as the ability to create output, then the agency of this LLMs is thus always collaborative (see also Nyholm (2018) for a definition of collaborative agency). If we choose to attribute agency in this context, then we should attribute it to the whole practice: to the user together with the system co-producing the output. A further consequence of this collaborative agency attribution is that the characterization of 'intelligence-free' is no longer correct. After all, a part of this collaborative agent – the human user – is intelligent.

A second reason why I think that calling LLMs 'intelligence-free agents' can be misleading, is that this attribution suggests that the LLMs act (semi-)autonomously. Acting is often differentiated from 'mere behavior' in that the acting is 'up to the agent' (Steward, 2012) But would we say that what an LLM does is really up to it? Yes it has learned to predict the most likely output, but this prediction is based on probability calculations and statistical analysis of huge amounts of data. Where it does make these predictions by itself, is this the kind of autonomous behavior that we are looking for in an agent? Would calling this behavior automated not be more realistic? In referring to LLMs as intelligence-free 'agents', I argue that the interesting question of what kind of autonomous behavior we expect from an agent is neglected and overlooked.

Even though I thus do not agree with Floridi that LLMs can or should be called intelligence-free agents, I do agree with him that the emergence of these systems is "good news for philosophers looking for work" (Floridi, 2023, p. 10). This new philosophical work, however, lies in (1) exploring how humans together with LLMs can co-produce knowledge (the groundwork for such analyses has already been prepared by Russo (2022)) and (2)

reconsidering the distinction we make between automation and autonomy, especially in light of these LLMs and the attribution of agency.

ID: 205 / Panel 4-3-G: 1
Individual paper
Topics: Metaphysics, Ethics (general and applied), Metaethics, Moral Psychology
Keywords: collective action, causation, omission, requiring reasons, normative asymmetry

Causation, omission, and the problem of collective action

Rutger van Oeveren

Sometimes, a number of actions together make a difference to some morally relevant outcome, but individually do not (and, let's assume, cannot). The question is what reasons we have to perform any individual action in such cases. I argue for a particular understanding of the principle to cause no harm, one that covers partial and omissive causes. This principle tells us one is required not to perform φ when φ causes harm, where φ may be an action or an omission, and may be a merely partial cause of the harm. The principle allows us to make progress on the collective action debate in three respects.

First, it allows us to distinguish more clearly several types of collective action cases: what I call harm cases, harm by omission cases, and mere failure to benefit cases. I discuss several ways of understanding what it takes to harm someone (or something), allowing us to distinguish harming from merely failing to benefit. Second, in these different types of cases, this principle does a better job than some of the existing views in accounting for a normative asymmetry in what one might call the deontic force of our reasons. In particular, it gives us a plausible account of when we have (and don't have) requiring reasons to perform an individual action as opposed to merely favoring or justifying reasons.

In the second part of the paper, I discuss some potential upshots for accounts of omission. In particular, I argue that this picture puts pressure on normativist accounts of omission, such as those of Thomson (2003) and McGrath (2005). Normativist accounts of omission explain which absences of behavior count as an omission by appeal to norms that are violated by those absences. For example, my not watering your plant is a cause of your plant dying if I promised to water your plant (such that my not watering the plant is a failure to keep that promise). But a total stranger's not watering your plant is no such cause—according to normativist accounts, because no norm was violated. Though this seems to get things right in the case just given, as I argue, normativist accounts of omission face a challenge. The challenge for normativist accounts of omission is to find a plausible norm that can play the same explanatory role in an important class of collective action cases. These are cases in which many people not doing something causes a bad outcome, and in which it is therefore plausible that your not doing something was a partial cause of the bad outcome (and in which it is also independently plausible that you had a requiring reason not to fail to perform that act).

ID: 129 / **Panel 4-3-D: 1**
Individual paper
Topics: Ethics (general and applied), Metaethics, Moral Psychology
Keywords: Kant, Kantianism, Constitutivism, Relativism

Why we cannot be 'non-Kantian' agents

Leon van Rijsbergen

According to Kant, the supreme principle of morality, the Categorical Imperative (CI), is a moral law for all agents. All agents, that is, are morally obligated to constrain their actions in accordance with the CI's demands. I call this Kantian moral universalism. An indispensable component of Kantian moral universalism is the connection between an agent's free, autonomous will, and the universal moral authority of the CI. However, interpretations differ when it comes to explaining this connection.

According to one interpretation, autonomy is the constitutive standard of action, and an agent is required to act in accordance with the CI in order to live up to this standard. On this interpretation, Kant thus maintains that the categorical norms for action are explained by the very nature of action. I call this the action constitutivist interpretation. In recent work, David Velleman, a proponent of this interpretation, has suggested that Kant's theory of action fails to pertain to all agents, because of which it is incapable of generating a moral principle to which all agents are constitutively committed. However, instead of entirely rejecting Kant's theory of action, Velleman maintains that it can be subsumed under his own relativistic action constitutivism.

On this account, the constitutive standard of action is 'agency under some particular conception thereof'. Accordingly, an action is good or bad relative to the extent to which it renders an agent attuned with her own particular commitment to agency. 'Kantian agency', according to Velleman, is just one contingent variant among others. Because not every agent makes action their own in the way that a 'Kantian' agent does, and moral norms are derived from one's particular commitment to agency, the CI is a contingent- rather than a necessary and universally valid principle of action. Or so Velleman argues.

In this paper, I argue that an agent must implicitly conceive of herself as free and autonomous in order to be able to act under any specific self-conception at all. In other words: an agent must implicitly conceive of herself as a 'Kantian' agent (viz., the sort of agent that reflects Kant's actual conception agency, not the action- constitutivist interpretation) before she is capable of acting under any specific self-conception. Therefore, Velleman fails to show that the CI cannot be a moral principle that applies to all agents.

ID: 218 / **Panel 2-3-F: 1**
Symposium paper
Topics: Social and political philosophy
Keywords: taxation, inheritance, tax avoidance, passive income

The fairness of tax avoidance

Bruno Verbeek

Centre for Political Philosophy, Leiden University

Tax avoidance, the legally permitted reduction of fiscal burdens, is a hot topic in circles of activists (e.g., the Tax Justice Network) and policy makers (e.g., the BEPS initiative of the OECD). The discussion is typically limited to multinational companies and their aggressive tax planning structures which often border on what is legally permissible. However, in this context it is crucial to distinguish tax avoidance from tax evasion on the one hand, and tax mitigation on the other. In this talk, I will try to give a characterization of tax avoidance, evasion, and mitigation, and try to argue when and why avoidance becomes problematic from a moral point of view. I will conclude that taxpayers, whether natural or legal persons, have fiscal duties that go beyond what the law requires.

ID: 107 / **Panel 4-1-C: 1**
Individual paper
Topics: None of the topics above (please include your topic under 'keywords')
Keywords: psychoanalysis, beyond the pleasure principle, death drive

Death drive as aggression Judith Butler's selective interpretation of the Beyond the Pleasure Principle

<u>Yuting Wang</u>

In the 2020 book The Force of Nonviolence, an attempt was made by Judith Butler to build a new political-philosophical framework that is structured on the Freudian dualistic theory of "Thanatos (death drive)" and "Eros (life drive)." (Butler, 2020, p. 107) The Freudian dualistic metaphysics was adopted by Butler to describe the fundamental "love and hate" dynamic in political scenarios: "Eros, the force that creates ever more complex human bonds," and "Thanatos, the force that breaks them down." (Butler, 2020, p. 107) From this assumption of human nature, Judith Butler tries to falsify the old binary opposition of aggressive violence and passive non-violence. "It is not only that we must sometimes aggressively defend our lives in order to preserve life (the aim of Eros); we also have to commit to living with those toward whom we maintain intense feelings of hostility and murderous impulse," declared Butler. (Butler, 2020, p. 123) And she argues why a blueprint for an aggressive nonviolent society is attainable when human nature vacillates between Thanatos and Eros. What is aggressive nonviolence? According to Butler, it is a way to utilize the destructive death drive constructively and carry out "the desire for the other's desire" to live passionately (Butler, 2020, p. 139). However, Freud's interpretation of death drive and life drive, especially in Beyond the Pleasure Principle (BBP), are less polarized and more complex.

ID: 176 / **Panel 4-3-H: 1**
Individual paper
Topics: Ethics (general and applied), Metaethics, Moral Psychology
Keywords: Life extension, African ethics, ageing, longevity, enhancement

Superlongevity and African ethics

Christopher Simon Wareham

The topic of enhancement – roughly, making use of biotechnologies to improve above the 'normal' – is high on the agenda in western bioethics. By contrast, with a few notable exceptions, African theorists have had had less to say about this subject. Nonetheless, recent African contributions have valuably pointed to problematic relational features of enhancement technologies that tend to be overlooked in more prominent western discussions.

I apply African theory to the topic of 'substantial', or 'considerable' life extension – or 'superlongevity.' While this particular form of enhancement is yet to be directly tackled by African theorists, I make the case that African theories give rise to two specific sorts of moral concern that are distinct from similar objections in western literature. First, significantly longer lives could be an impediment to an African perfectionist account of 'personhood', thereby undermining a key teleological goal for individuals. Second, life extension may be socially divisive, thereby undermining key social tenets of sharing a way of life and communing with others.

Although these distinctive concerns are significant, I suggest that their strength as objections to superlongevity depends heavily on the distribution of life extension technologies. Moreover, since African theories typically embrace the idea that moral excellence correlates with increasing age, they provide a prudential incentive to live longer to achieve 'greater personhood.' The upshot of these responses is that, instead of creating a case against life extension, African thought generates obligations to substantially extend our own lifespans and those of others.

ID: 142 / Panel 4-2-C: 1
Individual paper
Topics: Social and political philosophy, Critical Theory, Philosophy of Religion
Keywords: divine violence, Walter Benjamin, violence, law, emancipation, general strike

Divine violence as nonviolent-violence: Judith Butler's interpretation of divine violence

Hayden Simon Weaver

Abstract

The question of violence and how society can emancipate oneself from it has occupied many philosophers. Walter Benjamin attempted to answer this question in 1920 through the notion of divine violence. This notion has recently been resurrected by philosophers such as Jacques Derrida, Jürgen Habermas, Slavoj Žižek, and Judith Butler. Divine violence is turned to as a means of emancipating society from systemic oppression and coercive law. However, it is a notion that has been met by major critiques. Most notably is Jacques Derrida's critique given in Force of Law: The Mystical Foundations of Authority. This article examines Judith Butler's turn to divine violence in opposition to the critique of divine violence given by Derrida. Butler attempts to merge divine violence and nonviolence to create a means of nonviolent revolution capable of emancipating society from oppression and coercive law. However, in order to make this argument, Butler needs to overcome Derrida's objection that suggests that divine violence is a dangerous notion with the potential to justify horrendous forms of political violence. Does Butler successfully create a nonviolent divine violence capable of achieving this desired emancipation? Or does divine violence continue to be a notion with a dangerously destructive potential as Derrida suggests? These are the questions that this article attempts to answer through a detailed examination of both Butler and Derrida's work on divine violence. Ultimately, it is established that divine violence should be jettisoned into the realm of the divine, rather than harnessed for political ends.

Effective political philosophy: how to advance justice in nonideal conditions

Thomas Wells

Political philosophy is concerned with how political institutions ought to be arranged. An implication of this normative stance is that the recommended institutions should be possible. Yet the more concerned with feasibility – especially 'political' feasibility - the more constrained and compromised the normative analysis. Hence the concern with identifying the correct degree of 'realism' that animates the nonideal theory debate (Valentini 2012).

In this paper I first draw from Amartya Sen's work to make the case that one of the things that political philosophy ought to be concerned with is contributing to making the world a better place in the sense of being characterised by more justice (less injustice) than it otherwise would be. This branch of political philosophy is characterised by its concern with practical achievements, and this implies far more than usual commitment to accommodating the demands of the actually existing world, including contingent local political conditions.

I next draw from economists' reflections on nonideal theory to develop the concept of 3rd best theorising. A key finding of the General Theory of Second Best (Lipsey and Lancaster 1956) is that "if one of the Paretian optimum conditions cannot be fulfilled a second best optimum situation is achieved only by departing from all other optimum conditions". Hence, successful policymaking requires practical judgement about what is economically possible rather than the mere application of ideal theory. In addition many applied economists (such as Dani Rodrik) note the significance of particular – perhaps regrettable - political and economic institutions and engage in what I call 'third best' theorising: in which they ask the question 'Given how things are here, what is the most ambitious but politically feasible policy(s) for making the people here do better?'.

Finally I demonstrate how 3rd best theorising in political philosophy would look by examining the case for liberal democracies adopting UAE style temporary guest worker programmes. The basic idea of this is that tens of millions of people from poor countries would be allowed to take up jobs in rich countries like the Netherlands for up to 5 years on condition that they accept a distinctly 2nd class status (civil rights but not political or social rights).

This proposal is justified by two distinct arguments.

Firstly, from a consequentialist perspective there are enormous gains to be had from enabling more economic migration, especially but not only for poorer countries (Lant Pritchett 2006). Thus it is a practical way to improve the world.

Secondly, lowering rich countries' moral standards for the treatment of the global poor will enable far more of them to help themselves to a better life than anything else rich countries might actually do – which makes it the most ambitious but still politically feasible policy to attempt. This is because there is – regrettably - no political consensus in liberal democracies that would support a morally superior solution to the problem of global poverty.

ID: 180 / **Panel 5-1-C: 1**
Individual paper
Topics: Ethics (general and applied), Metaethics, Moral Psychology
Keywords: COVID-19, public health. justification, crisis, harm principle

Public health policy justification – not just a matter of individual harm

<u>Lucie White</u>

The COVID pandemic was an exceptional public health situation – which brought with it unprecedented restrictions across the global populace. But what was it about this pandemic which caused us to implement such drastic restrictions on liberty? Much of the ethical debate on lockdowns and vaccination measures during the pandemic focused on the potential harm caused to other individuals by the risk of infection, and, following a tradition in pre-COVID debates on mandatory vaccination, utilised J.S. Mill's "harm principle" to argue that posing another person such risk of harm provided sufficient grounds to limit liberty. But this line of argument misses a distinctive sort of harm posed by the COVID pandemic, which could arise in future public health crises, and which may provide a more solid justification for extreme measures: the potential failure of healthcare systems.

Although the impending collapse of healthcare systems was commonly invoked by policy-makers when justifying the most extreme measures taken to slow the spread of the virus, the individual-harm-focused debate in philosophical ethics missed the ethical significance of this line of justification. I will draw out three ways in which harm-principle-based justifications of public health measures fail to capture the full scope of harm wrought by the collapse of healthcare systems. First, they can't adequately capture the cumulative and "looping effects" of the harm caused by strained healthcare systems. Second, they fail to capture the widespread ripple effects the failure of a central societal institution can have on other institutions. And third – much more difficult to quantify, but important nonetheless – the failure of a healthcare system can impose what Gerald Dworkin and Joel Feinberg refer to as "psychic costs", affecting the moral character of all members of society, reducing trust in institutions, and potentially posing an existential threat to the fabric of society. After laying out the ways in which the harm-principle-based approach is unable to capture these concerns, I will briefly consider what a shift to focusing on the societal harm of institutional failure might mean for the ethical justification of public health measures, both in the recent past, and in the future.

ID: 128 / **Panel 5-2-H: 1**
Individual paper
Topics: Social and political philosophy, Critical Theory, History of philosophy (from ancient to modern)
Keywords: Imagination, Politics, Spinoza, Superstition, Social Conflict

Imagining harmony: Imagination, harmonious cohabitation, and conflict in Spinoza's Tractatus Theologico-Politicus

<u>Joost Wijffels</u>

That we lack personal and collective imagination might be one way to diagnose the predicament of our times. Indeed, there are increased rates of depression, somberness, and pessimism about the future among individuals. Simultaneously, we seem to be politically disenchanted and disconnected, as suggested by the fallen trust in politicians, escalating polarization, and the rise of conspiratorial ideologies. These feelings are understandable, but also unviable in a time where crises are consistently demanding our attention and collective action. What role can the imagination play in enabling this collective action? And what could be the political effects of imagination gone missing or gone wrong? It seems that recently, more scholars, perhaps disillusioned by dominant Enlightenment thinkers, have realized that Spinoza's writings might contain the answer to questions like these. Interestingly, historians of philosophy have regained interest in Spinoza's Tractatus Theologico-Politicus, but few have translated his work to our current political context. Many critical scholars, on the other hand, ground their politics in the holistic ontology outlined in Spinoza's ethics, but largely overlook his political writings. This is a missed opportunity. Spinoza's realism provides a fresh alternative to the idealist theories that remain dominant today and his work, written about the imagination in times of conflict, might prove particularly relevant given the political challenges we face today.

I thus aim to respond to some of the socio-political crises of our time and contribute to the emerging tradition of critical Spinozism by examining what Spinoza can teach us about the importance of imagination for living harmoniously together. To answer this question, I spell out how Spinoza understood the imagination and its role in political life. Spinoza saw the imagination as a highly affective, relational, and contextual form of knowledge through which we make sense of our relation to the world. The cultivation of this imagination is crucial for inspiring willing obedience to one's political community and the state and thus is indispensable for living harmoniously together. Subsequently, I apply Spinoza's work to our current context, revealing Spinoza's hesitations about the imagination and highlighting how these hesitations manifest today. This reveals that the issue we face today is not so much a lack of imagination – which would be a near possibility – but rather the competition of various conflicting imaginative schemas that too easily revert into dangerous forms of superstition. This sets up certain groups as scapegoats, prevents the true and structural alleviation of harm, and can result in an escalating and vicious cycle of conflict. Lastly, I consider how we might cultivate the positive facets of imagination whilst protecting ourselves from its negative sides. In doing so, I argue that the imagination is a powerful force that we can and must foster to act responsibly and contribute meaningfully to our world as long as it revolves around collectiveness and mutual understanding, is limited in its violent expressions, and it backed up by material intervention.

ID: 190 / Panel 4-1-B: 1
Individual paper
Topics: Aesthetics and philosophy of art
Keywords: immersion, interactivity, artworks, imagination, K. Walton

When is a work immersive?

Nathan Wildman

The immersive experiences literature has primarily focused on the phenomenon of immersion, particularly its cognitive nature. However, little has been directly said about what it means for a work – a novel, videogame, VR application – itself to be immersive. The aim of this paper is to shed some light on this neglected notion of work immersiveness. Specifically, building off a broadly Waltonian conception of works as props in games of make-believe (Walton 1990), I here develop two conceptions of work immersiveness. Prop immersiveness is a relational, contextual property, such that a work is prop immersive for a user at a time to the degree that, as a prop, it facilitates the user's make-believing prescribed content; consequently, prop immersiveness comes in degrees – a work can be more or less immersive – and is user relative – a work can be highly prop immersive for one user but not very immersive for another. Meanwhile, interactive immersiveness is characterized in terms of how rich a choice range a fiction provides players (alternatively, how forced-choice incomplete the work is; see e.g. Wildman & Woodward 2018). These two complement each other, making sense of different ranges of cases.

To do so, I begin by quickly detailing the Waltonian background. I then turn to motivating and describing the notion of prop immersiveness. Having sketched the concept, I proceed to apply it to a range of examples from literature, videogames, and VR, demonstrating how these works can be understood to be (non-)immersive in specific contexts. I also use the concept to explain why many generally find VR more immersive than videogames, despite the two media being similar in many ways. Next, I sketch the alternative, albeit related, concept of interactive immersiveness. I then employ this alternative to explains some outlier cases of works that are said to be immersive due to how well they allow players to 'take a role in the story'. Finally, I conclude by showing how my two conceptions of work immersiveness neatly fit with four existing accounts of imaginative immersion; specifically, those of Schellenberg (2013), Liao (unpublished), Langland-Hassan (2020), and Chasid. The upshot is that these two notions provide a better understanding of what it means for a work to be immersive, and hence a better overall picture of immersive experiences.

References

Chasid, Alon (2021). Imaginative immersion, regulation, and doxastic mediation. Synthese online first. DOI: https://doi.org/10.1007/s11229-021-03055-1

Langland-Hassan, Peter (2020). Explaining Imagination. Oxford: Oxford UP.

Liao, Shen-yi (unpublished ms.). "Immersion is Attention/Becoming Immersed." Available at: https://philarchive.org/rec/LIAIIA

Schellenberg, Susanna (2013). "Belief and Desire in Imagination and Immersion." Journal of Philosophy 110 (9): 497-517.

Wildman, N. & Woodward, R. (2018) "Interactivity, Fictionality, and Incompleteness" in G. Tavinor & J. Robson (eds), The Aesthetics of Videogames, Routledge: New York.

Walton, Kendall. L. (1990). Mimesis as Make-Believe. Cambridge MA: Harvard University Press.

ID: 237 / **Keynote 1: 1**
Individual paper
Topics: Social and political philosophy, Ethics (general and applied), Metaethics, Moral Psychology, History of philosophy (from ancient to modern)
Keywords: Confucius, leadership, ethics, politics

Zhuangzi on Not Following the Leader

David Wong

I begin with identifying Confucian metaphors of leadership for the way the mind (or the heart-mind) should lead the whole person. I then discuss how the Daoist text Zhuangzi criticizes this conception of the mind's leadership as too fixed and rigid--unresponsive to the fluidity and unpredictability of the world. The text suggests as an alternative a way that the whole embodied person can fluidly respond to the world. This alternative ties into some contemporary work, scientific and philosophical, of how the whole person and not just the deliberating mind processes information from the world. I end by discussing how the critique of the fixed and rigid mind can suggest alternative models of political governance that distribute and integrate guidance throughout the body politic.

How should representatives change their minds in deliberations?

Erica Celine Yu

Deliberative democracy is a theory of legitimacy which traces the authority of laws and policies to the participation in authentic deliberation by all those subject to them (Dryzek 2001). In large democracies, however, the inclusion of all affected in deliberations is impossible. This makes deliberative representation—the making present of individuals not present literally or in fact in deliberations—essential. A member of a representative committee represents an individual in deliberations to the extent that she can be expected to 'make present' the values, views, and interests of that individual in the committee's deliberations, and be sensitive to and convinced by the same arguments that would convince the individual were she present in these deliberations (Chamberlin and Courant 1983).

When individuals select representative committees, the only information that they have is backward-looking. In other words, they would not know how their representatives would transform their judgments during deliberations themselves. On the one hand, it would be in the interest of the individuals represented for the representatives to never change their minds during deliberations and keep representing their values, views, and interests. On the other hand, one of the goals of deliberation is precisely to facilitate such changing of minds. The question I would then like to address in this paper is then the following: When does the transformation of judgments of representatives in deliberations threaten the representation relationship?

Take the following example: Anna is Betta's highest-ranked committee member on the issue of gun control, as both Anna and Betta believe that strict gun control should be legislated on the national level. Say that in the course of deliberations, Anna undergoes judgment transformation such that she now believes that she believes that no gun control should be implemented at all. Intuitively, this would mean that Anna no longer represents Betta on this issue as they are now fundamentally opposed. Alternatively, say that Anna transforms her belief to be that gun control should be legislated on the state level. This is something that Betta could see herself supporting were national implementation deemed unattainable. This belief transformation of Anna thus does not threaten the representation relationship.

To determine what kinds of preference and belief transformation threaten the representation relationship and which do not, I turn to Urbinati's (2000) theory of political representation as advocacy. Representatives should both have a passionate link to the elector's cause as well as exercise their autonomous judgment in deliberation. Using this as a general guideline for how representatives should act in deliberations, I investigate how plausible different measures of representativeness are at capturing this double role of representatives. More specifically, I propose desiderata for such measures in order to capture that representatives can change their minds in ways that are either justified or unjustified. I then look at the measure of representativeness proposed by Brenneis et al. (2020) as the distance between weighted argumentation graphs of individuals, and see in how far it satisfies the desiderata. I then propose an alternative that I argue satisfies the desiderata better.

ID: 143 / **Panel 5-1-D: 1**
Individual paper
Topics: Social and political philosophy, History of philosophy (from ancient to modern)
Keywords: Hegel, habit, madness, resistance

Habit, madness and the laziness of philosophers: Hegel on resistance to philosophical thinking

Bart Zantvoort

Philosophical thinking is often conceived of as a purposive activity, which is supposed to develop in a certain direction – whether it is an autonomous development in the direction of truth or, as social criticism, a development in the direction of a more just society. Historically, it has often been maintained that the free development of philosophical thinking is constrained by factors exogenous to philosophy; for example by a lack of empirical data, by the interference of the body or emotions or desire, or by social conditions like censorship, ideology or the distorting effects of power. Using Hegel's account of the role of habit in philosophical thinking, I will argue that there are forms of resistance to philosophical thinking internal to philosophical thinking itself.

On the one hand, Hegel seems to argue that thinking is irresistible: it exposes false limitations imposed on the subject matter of thought and continuously forces us to transgress boundaries, leading to an 'unhalting forward motion' towards the truth. On the other hand, Hegel also explores the various obstacles that get in the way of thinking. These include bad philosophical habits such as an overemphasis on particular things (empiricism), an overemphasis on general categories ('formalism' or rationalism) as well as the lazy desire to avoid the hard 'labour of the concept' and shoot straight for conclusions.

I will briefly outline Hegel's account of madness and habit in the 'Anthropology' section of the Encyclopaedia of Philosophical Sciences to explore the forms of resistance to philosophical thinking internal to philosophy, and show that philosophy, on Hegel's account, is an embodied 'ethos' that may be able to find the right mean between bad philosophical habits, but that can never detach itself from habit altogether.