

*Book of abstracts*

*2022 Utrecht Philosophy Graduate Conference*



12th of May (Thursday)

**Maja Griem** (Ruhr University Bochum): “Animal Play Signals as a Form of Intentional Communication in a Weak Gricean Sense”

Intentional communication has long been thought of as an ability unique to humans, as the classical Gricean (1957; 1989) communication requires third-order intentionality, with no evidence suggesting anything like that in nonhuman animals. However, it has been argued that the strong Gricean prerequisite is too sophisticated as it excludes children from communicative intents (Andrews 2015). Therefore, we need an account for intentional communication that is appropriate for describing interesting cases of communication in humans and allows for comparability between human adults, children and other animals.

I suggest taking a weak Gricean stance on intentional communication to offer the possibility for interspecies comparison and developmental investigations. One possibility includes ostensive accounts of communication. Here, communication can be split into two parts: (1) a message and (2) a signal that the message is intended (Andrews, 2015). If the message, for example the gesture of pointing or the gaze towards something interesting, is combined with an ostensive cue (e.g., eye contact), which shows the intention of directing the other’s attention, this can be seen as a communicative act. Tomasello (2008) goes even further in arguing that a signal is intentional – and, thus, communicative –, if it is used in a flexible way, includes a learning component, and the user is aware of the attentional state of the receiver.

Having this set, I provide available evidence for such a form of communication in nonhuman animals using the example of two common play soliciting signals: the play bow and the play face or relaxed open mouth display (e.g., Bekoff & Allen 1998). Since these two signals are relatively well studied and common among mammals, they are good candidates to look at the intentional communicative aspect of their use across species. Therefore, I will take a closer look at the learning component, flexibility of use with respect to context and receiver and the sender’s sensitivity for the receiver’s attentional state during signal production in these two play-soliciting patterns.

The available data shows that more and clearer play signs are used in situations, in which the risk of escalation is high. These include species-specific circumstances – e.g., strict hierarchies (Nolfo et al. 2021) – as well as ambiguous contexts and unknown partners – e.g., integration of newcomers (Stahler et al. 2002). Signalers are sensitive of the receiver’s attentional state and show attention-getting

behaviors before play signal performance wherever necessary (Horowitz 2009; VanderLaan et al. 2012; Nolfo et al. 2021). This indicates that play signals are used flexibly, taking into account the receiver's attentional state and probably possess a learning component, rendering them intentionally communicative signals according to a weaker Gricean view.

Given the complexity of play and involved pretense-like structures, play signals should be an object of further investigation. Since a wide range of species exhibits diverse forms of social play, further examination could provide a broad comparative perspective of play, communication, and cooperation across species

---

**Ludmila Reimer** (Ruhr-University Bochum): “Gesture as Feature Highlighting by Exemplification”

Spoken language and gestures are interconnected, yet it is still puzzling whether and how gestures contribute meaning to a spoken sentence. Do they have meanings of their own, interacting with spoken sentence meanings? Or do they merely modify, facilitate, or enhance speech? Is their function dependent on what type they are? And can we transfer findings about one type of gesture to other types?

To tackle these questions, I propose to redefine the core mechanism of gestures to be feature highlighting of spoken about objects or actions by exemplification borrowed from Goodman's theory of symbols. At first, this will be applied to iconic co-speech gestures, i.e. gestures produced simultaneously with speech and contrasted with the possible functions of pro-speech gestures, i.e. gestures produced in place of a spoken utterance, usually a word or short phrase. Both kinds are assumed to carry meaning on their own and are researched in this fashion, leading to theories that often treat them functionally similar to adjectives or appositives. Other approaches follow McNeill's imagistic view and his growth point theory, according to which gestures are windows into our thoughts and are inseparable from language. This gives little insight into why we gesture and what the advantages of one gesture over a spoken alternative would be. Here, some findings that gestures enhance speech – both for the speaker in better word retrieval, and for the listeners in information – by probably providing redundant information, set in and open a door for feature highlighting by exemplification.

To see how this can work, let's turn to the following example: Mary says, "My new tent is round", and makes a full circular movement with her hand or index finger while uttering "round". This seems redundant at first but she actually is disambiguating "round" by highlighting the most distinguishing feature of the tent, its circular base and probably circus tent shape. Compare this to a different gesture, namely a half-circle motion with a (cupped) hand. Now the tent's shape is to be more likely be a dome tent, which does not have necessarily a round base but needs a top resembling a hemisphere to be categorised as such. So even though "round" already picks up a feature of the tent that is separating it from the prototypical wedge or A-shaped tent, it leaves open other possibilities. It is the seemingly redundant gesture that is highlighting the importance of that feature and that it sets it apart by exemplarily tracing it.

Overall, this re-definition allows for more flexible interpretations of iconic co-speech gestures and might even be extendable to other kinds of gestures; e.g. in case of beat gestures, which put emphasis on a word or a syllable, the highlighted feature could be simply the intonation of the simultaneously uttered word. Additionally, it might solve problems arising in the debate around at-issueness of gestures, since they rely on some-what fixed gesture meanings. However, this re-definition cannot handle the potentially compositional nature of gestures.

---

**Michaela Fikejzová** (Metropolitan University Prague): "How Does a „Normal“ Speaker Refer?"

The conference paper deals with the problem of reference in the contextualist notion of mental files proposed by François Récanati (2003, 2012, 2016, 2018). He develops the conceptual foundations of the theory of mental files and focuses on the issues of singular reference and thought concerning the so-called modes of presentation. Central to Récanati's notion of mental files is the concept of so-called "normal" speaker, that is fully functional concerning classical problems in analytic philosophy of language, such as coreference of the names Cicero and Tully, or Hesperus, Phosphorus, and Venus, that are the base for Récanati's theory. In my paper, I focus on the problematic nature of the concept of the "normal" speaker when confronted with "less standard" situations, which, however, occur abundantly.

In the center of my critique of Récanati's "normal" speaker stand situations in which misgendering happens. My definition of misgendering is purposefully broad: misgendering is when a speaker refers

to a person with gendered language that does not align with the gender with which given person identifies. Récanati presupposes the automatic connection between and coreference of (gendered) pronouns and (gendered) proper names, as well as that only one proper name can refer to one person and gender is generally presupposed to be an inherent property of an object, and even more important that this “fact” is easily accessible to any “normal” speaker. Building on social constructivist theory of gender (esp. Butler 1990, 1996; Halberstam 1996, 2018) I am debunking this “automatic” connection, just as the supposed “clearness” of reference employing definite descriptions (such as “a woman with blonde hair”) when the identity of the object we are referring to is unclear.

My aim is to ameliorate Récanati’s framework in such a way that it can grasp these “non-standard” situations. In this pursuit, I am following the steps of Bettcher (2013, 2017), Saul (2012), and Diaz-Leon (2016), who tried to find a more suitable, emancipated, and, most importantly, trans\*-inclusive definition of the term “woman” and expanding their logic to a wider range of terms. Through reconceptualizing Récanati’s notion of mental files, I am queering the concept of “normal” speaker itself, as trans\* people and people referring to them with their preferred pronouns would be considered “irrational” in the current state of the framework.

Even though this is the main aim of this paper, I argue that the analysis of misgendering is useful for Récanati’s notion of proper names, definite descriptions and indexicals in general. Firstly, Récanati focuses purely on English, therefore considers gender to be a marginal problem. When applying his framework to more gendered languages, e.g., French, German, Czech, there are crucial connections missing in the schema. Secondly, Récanati is omitting the phenomenon of name-giving and newly arisen nicknames. When misgendering in the form of deadnaming happens, it can be extremely harmful linguistic behavior. Understanding the mechanics of such a practice leads us to better understanding how name-giving and nicknaming occurs and is accepted as a proper reference.

---

**Sofia Meléndez Gutiérrez** (Cambridge University): “Fictionalism and Empty Names”

Fictive sentences are those which purport to refer to ficta, or quantify over them. In everyday life, we often assert sentences of this kind, and assent to them. Consider:

1. Sherlock Holmes is a detective.

(1) is intuitively true; and, if it is, then the name 'Sherlock Holmes' must refer. It appears platitudinous, though, that fictional entities do not exist; and, if they do not, then (1) is untrue.

Fictionalists believe that our intuitions that (1) is true can be honoured even if we deny the existence of ficta. On their view, (1) is untrue because Sherlock Holmes does not refer; but, ordinarily, we use it as elliptical for a true statement:

2. According to the story, Sherlock Holmes is a detective.

The fiction operator governing (2) is circumstance-shifting. The truth of (2), then, does not require that 'Sherlock Holmes' refers to a detective. According to fictionalists, (2) is true just in case Doyle's story says that Sherlock Holmes is a detective. It follows, they conclude, that (2) is true. Moreover, given that (1) is normally used to convey (2), ordinary assertions of (1) are truthful, as our intuitions suggest. Thus these intuitions have been honoured.

To my mind, fictionalists made a mistake concerning circumstance-shifting operators. They claimed, we have seen, that these operators may take untrue sentences and deliver true statements; but this claim requires a crucial precisification: circumstance-shifting operators may deliver true statements when they take sentences that are untrue and complete in sense of not comprising any empty terms. If the name 'Sherlock Holmes' is empty, then the sentence 'Sherlock Holmes is a detective' is incomplete, and embedding it within the scope of the fiction operator cannot change that. The fiction operator, in other words, cannot do anything to fill the semantic gap that empty names produce. If this is right, then the proposition expressed by (2) is not truth-apt:

3. According to the stories, \_\_\_\_ is a detective.

If my arguments are sound, then fictionalists have failed in establishing that (1) is typically used to convey truthful information. My paper consists of three sections. In §1, I show that fictionalists of all stripes (both content theorists like Brock [2002 and 2016] and force theorists like Crimins [1998], Everett [2005 and 2013], and Recanati [2018]) have grossly overestimated the power of circumstance-shifting operators in a systematic way. I will contend, accordingly, that virtually all fictive sentences are as meaningless as (3) on their fictionalist analysis.

Brock (2016, pp. 400-401) thinks that this meaninglessness problem may be comfortably solved by the adoption of either descriptivism, or a Fregean semantics. In §2 and 3, I argue that he is mistaken: these theories can only be adopted by fictionalists at a very high theoretical cost. I conclude that the meaninglessness problem is much more threatening than it has been recognised thus far.

---

**Denis Kazankov** (Central European University): “Meta-Semantic Normative Encroachment”

In this paper, I defend what I call Meta-Semantic Normative Encroachment (MNE). MNE says that the meaning of a linguistic concept can be at least partially grounded in the normative facts about its value for the linguistic community of its users. I claim that MNE has an explanatory utility in meta-semantics because it presents a promising alternative to an influential meta-semantic theory called Temporal Externalism (TE).

According to TE, a concept’s meaning is sometimes determined by the future linguistic practice of its users. Hence, TE denies that what a concept means at a time necessarily supervenes upon the totality of the events that have taken place up to that time. What makes TE a compelling meta-semantic thesis is that it purports to provide the best explanation of the following two phenomena:

(Determinacy) There are cases in which the events that have taken place up to a certain time seem to underdetermine a concept’s meaning, but we are still reluctant to describe the concept as semantically indeterminate at that time.

(Deference) There is a widely accepted thesis in meta-semantics called social externalism, according to which we defer to experts in our linguistic community for the meaning of our concepts (Putnam, 1975, p.186; Burge, 1979). The main evidence for such deference is that we are disposed to correct our understanding (and use) of a concept when discovering that experts understand (and use) it differently. This evidence seems equally compelling even if experts are situated in the future.

I defend MNE on the grounds that it can account for Determinacy and Deference better than TE.

The body of the paper is structured as follows: In §2, I lay out how TE is supposed to explain Determinacy. I introduce Jackman’s case (1999) of a putative linguistic community that possesses the concept ‘ave’ which seems semantically underdetermined by the totality of the events that have taken place until a certain time,  $t$ , but whose meaning doesn’t seem to change at  $t$  and thus doesn’t seem indeterminate before  $t$ . I explain why TE *prima facie* looks like the strongest explanation of this and similar cases.

In §3, I lay out how TE is supposed to explain Deference. I present some reasons why TE is thought to be a special kind of social externalism which pertains to the cases when experts are situated in the future.

In §4, I show that the way in which TE is supposed to explain the two phenomena in §2 and §3 is problematic because the members of the putative linguistic community in §2 aren't disposed to defer to their future selves on the meaning of 'ave'. I argue that MNE avoids similar problems and thus offers a better explanation of Determinacy and Deference which still allows for a weaker version of TE. According to my proposed account, normative facts about what makes a concept all-things-considered valuable for its users, which may also depend upon the facts about the future linguistic practice of its users, sometimes directly determine its meaning as long as other meta-semantic factors such as its dominant use or its central function underdetermine it.

---

**Matteo Nizzardo** (University of St Andrews & University of Stirling): “Singular Reference & Indiscernible Entities”

Can we refer to indiscernible entities singularly? The relevant literature seems uniformly committed to a negative answer: singular reference to indiscernibles is impossible. In contrast with this trend, I argue that statements like “Let  $a$  be one of the indiscernibles” can indeed single out indiscernibles as subjects of singular reference, and I provide a formal framework in which singular reference to indiscernible individuals can be achieved.

Arbitrary Reference (ar for short) is the thesis that in some contexts we can refer to individuals with some degrees of arbitrariness. Although the literature on ar is fairly developed, ar has so far been applied only in the context of mathematical proofs, and never to indiscernible objects. Thus, my proposal is the first of its kind.

In this paper, I draw heavily from Wood's (2014) account of ar, in which names are modeled as choice functions. Within this framework, the sentence “Let  $n$  be a natural number” is formalized with the help of Hilbert's  $\epsilon$  operator, as ' $n = \epsilon x.x \in \mathbb{N}$ ', and the meaning of the term ' $n$ ' is fixed by choice functions  $f: \wp(D) \rightarrow D$ . I argue that Wood's framework can be adopted to obtain singular reference to indiscernible entities.

Finally I consider one possible objection, according to which indiscernibles cannot be modeled using Zermelo-Fraenkel's Set Theory with Choice (zfc), for no zfc set can contain indiscernible elements — and so any talk of choice function is misguided, when applied to indiscernible individuals. Whilst I grant that zfc doesn't allow for the existence of sets containing indiscernible elements, I argue that our intuitions on choice functions can be preserved under different set theories. zfc's obedience to a

version of the Identity of Indiscernibles is a result of the interplay between the Axiom of Extensionality and the Axiom of Well-foundedness. On the one hand, by dropping the latter, we obtain Non well-founded set theories. (See Aczel 1988). One of these theories, known as Boffa's Set Theory, allows for both (1) indiscernible sets and (2) choice functions on any non-empty collection. On the other hand, we can decide to weaken the power of the Axiom of Extensionality by adding to the ontology of our theory objects that are not sets. This is the case of theories like Zermelo-Fraenkel's Set Theory with Urelements (zfu), where urelements and sets appear side by side. Since the Axiom of Extensionality is restricted to sets, any two urelements are indiscernible from the point of view of zfu. This makes them good candidates to model indiscernible individuals in a mathematical framework. Further, zfu can be augmented with a version of the Axiom of Choice. The resulting theory, zfcu, is an example of a theory of sets that allows choice functions to single out indiscernible elements.

To conclude: the fact that we can refer to indiscernible entities singularly is important in many respects. First, it sheds light on some of our linguistic practices, in both philosophy and physics, where it is customary to assign names to indiscernible particles. Second, it helps us understanding how indiscernibles are individuated. Finally, it shows that skeptical arguments against the possibility of indiscernibles based on the impossibility of their reference are fundamentally misguided.

---

**Cécile Rosat** (University of Neuchâtel): "Actional Theory of Adam Smith's Sympathy"

What is to be understood when someone says: "Put yourself in her shoes!" Without a doubt, it is empathy or sympathy. But may this kind of phenomenon be an action? Intuitively, it seems that sympathizing could be something done by an agent. To argue in favor of such an ontological move, one line of argument consists in showing that it requires an effort (Massin, 2017). Whether it is projective empathy or Smithian sympathy, this effort is either implied or explicitly acknowledged (Bailey, 2018; Kaupinnen, 2017; Fleishacker, 2012; Coplan, 2011). Although researchers stress the importance of effort, this aspect has received little attention. Overall, existing studies do not define the concept fully nor address the possibility of actional sympathy.

In this talk, I nuance Adam Smith's (1982) understanding of sympathy, currently known as projective empathy (Darwall, 1998), and propose an actional analysis of sympathy. It shows that firstly an effort requires an action. Then, it argues that a specific sort of sympathy implies an effort. Lastly, it provides a substantial account of sympathetic effort, defined in terms of forces (Massin, 2017).

Actional sympathy consists of two main characteristics: A projective and an other-oriented perspective-taking process. When Julie sympathizes with Tom, she puts herself in his shoes and adopts his point of view. She does not try to imagine what she would feel if she was in his situation but what she would feel if she was Tom. In this way, actional sympathy requires not merely a change of place but of perspective. To adopt the other's point of view is not an easy thing because we are continually conscious that we are not the one who actually suffers. For this reason, to sympathize implies a sort of resistance.

To sum up, if there is an effort, there is an action. Since sympathizing involves effort, then sympathy - in a Neosmithian account - can claim the status of action.

---

**Robin Timothée Bianchi** (University of Neuchâtel): “The disappearing agent problem revisited”

The main idea of what we may call the Hobbesian picture of action is that actions are (bodily) movements caused by efficacious desires. Roughly speaking, when we act, we end up doing what our strongest motivational state aims at. Agency is here understood mechanistically in terms of (motivational) forces pushing in different directions, the strongest determining the behaviour of the agent, hence its name as the ‘hydraulic model’ of agency (Wallace 2006; Schroeter 2004). A metaphor is often used to describe this model of agency: agents are the arena in which different psychological forces compete for the guidance of the agent’s behaviour. The Standard story of action echoes this hydraulic conception of behaviour (see Smith 2004 & 2010; Schlosser 2011). In this picture of agency, actions are events caused by pairs of desires and beliefs or sui generis intentions. That means that desires or intentions “are given a substantive explanatory role in the aetiology of intentional action. They determine which action we perform by causing the bodily movements that we make in acting.” (Wallace 2006, 173). Although it is rarely made explicit, the Standard Story has its conceptual roots in the Hobbesian picture of action (see Mayr 2011, Ch.3). As a result, some have worried that human agents play no role in bringing about or refraining from bringing about their actions, since all the causal work is done by causally efficacious intentions or desires. This worry is found encapsulated in the so-called disappearing agent problem. Several diagnosis of the problem can be found in the literature and different culprits have been identified: determinism (Steward 2012); the humean picture of causation (Mayr 2011; Lowe 2008); the lack of interest for powers and their exercise; the failure of the standard story to account for agency par excellence (Velleman 2015; Bratman 2001; Schroeter 2004; Mele 2003) and so on.

In this paper, I would like to offer a different diagnosis of the problem. In my view, the disappearing agent problem arises because of a failure to recognize that human agency, as other forms of animal agency, is a form of collective agency rather than individual agency. As I argue once we start appreciating that human agency is really the agency of complex substances with functionally differentiated parts (Hyman 2015, Ch.2), the disappearing agent problem breaks down into two separate issues. On the one hand, we found the question of the attribution of agency to wholes – complex substances with parts such as human beings – to which I suggest the key answer is the notion of functional integration (Hyman 2015, Ch.2). On the other hand, lies the question of what makes human agency, and more specifically reflective and deliberate agency, so special – to which, I suggest the answer is to be found in the notion of ‘conscious self’ as a central thought-processing system involved in selecting options in accordance with the different plans the agent possesses. The upshot is that talk of efficacious desires, volitions or intentions are less important than has been thought to understand human agency and its distinctive features contra the standard story.

---

**Giulia Schirripa** (University of St Andrews and the University of Stirling): “Classical Extensional Mereology, Social Groups and the Coextensionality Objection”

Social groups are everywhere. Hiking groups, political action committees, African American women, scientific teams, the Supreme Court, and Palestinians, just to mention a few, shape and permeate our existence in innumerable ways. Although their impact on us is often elusive, their importance could be hardly overestimated. Many among us would be willing to recognise them as entities, but what their ontological status is, is a matter of debate.

If we look at the literature of social ontology, we will find out that there are several options available, but that the classical mereological view is not among them. The classical mereological account of social groups holds, roughly, that the relation holding between a social group and its members is the wholepart relation. This view used to be attractive in the Sixties and Seventies but has fallen more and more out of philosophical favour due to a number of objections, among which that from coextensionality has been one of the most detrimental. The aim of this paper is to defend the classical mereological account of social groups by presenting an alternative approach to the standard one - which identifies membership with parthood. According to my view, rather than being just mereological

wholes, social groups are mereological wholes with a specified understanding of the membership relation.

I intend to structure the presentation as follows. After an introductory contextualisation of the issue, I will present the standard axiomatisation of Classical Extensional Mereology (CEM) and, subsequently, the account of social groups that used to be derived from it and that, following Strohmaier (2018), I shall label ‘naïve CEMSG’ (nCEMSG), i.e., naïve classical mereological account of social groups. I will then focus on some problematic features of social groups emerging from this characterisation and, among then, I will discuss the extensionality principles, which prevent two different social groups to be composed by the exact same members. For instance, according to nCEMSG, if A, B, C, and D compose a reading group, then they cannot also compose, say, a running group, because fusions are unique. Given this scenario, the coextensionality objection stems from the intuitiveness of the idea – and real cases awareness – that there can well be different social groups composed by the exact same members that are nonetheless distinct. From that, I will move on by presenting my diagnosis of the problem: nCEMSG is unable to account for group membership because it treats the socially constructed membership relation as if it was just parthood. Otherwise put, the alleged failure of CEM to provide us with a satisfactory account of social groups is misguided because nCEMSG is not the only possible way to argue for a mereological account of social groups based on CEM. My proposal is to regard membership as parthood plus further conditions. By ‘parthood’ I mean the fundamental relation CEM aims to grasp and by ‘further conditions’ I mean any possible condition that might characterize social groups but that is chosen from outside CEM. In the last part of the presentation, I will show that the proposed account is desirable and that it succeeds in rejecting the objection from coextensionality.

---

**Roos Geerse** (Utrecht University): “A neurophilosophical account of thinking for oneself”

Thinking for oneself can be difficult. Even if you have good arguments, it may be hard to stick to your own opinion if the people around you or a person with authority for instance are of a different opinion, especially if you have a tendency to conform. As thinking is often equated with conscious reasoning, this may seem like a cognitive problem. However, what I will argue in this paper, using insights from various disciplines, is that reasoning is a largely unconscious process that involves affect as well as

cognition. In itself this may not be a very original view, but what I submit is new is its neurophilosophical backing.

I first outline a neurophilosophical model, based on the following assumptions. People have generative models of the world in the form of representational networks that they continuously update. Both on the mental and on the neural plane we need to distinguish between modal (non-semantic) and amodal (semantic) concepts. Moreover, within the category of the amodal concepts, we need to distinguish between the referential and the relational ones. Lastly, representations in one subsystem can be what I call reiterated in the other subsystems, i.e. the representations in for instance the semantic system can activate the corresponding representations in for instance the affective system.

Taken together these assumptions suggest how we can draw conclusions about a situation outside of awareness. Firstly, representations can be activated through spreading of activation. Secondly, relations between referential concepts can be encoded because relational concepts like agent and action can be nodes in the representational network as well as referential concepts can. Thirdly, because conclusions and inferential rules can be nodes as well, these too will be activated and combined into new ones, when applicable. As a result many inferences can be made swiftly, in parallel and outside of awareness.

Because relative to the other inferences the more relevant ones will get their activation from more elements that in addition will be activated more themselves, as a rule they will also be activated stronger and for a longer period of time. Additionally, a highly activated conclusion can be activated extra when it is reiterated in another mental system, thus leading to focus. This may be especially true for the evaluative conclusions that are drawn when an element in the situation is related to a so called concern, i.e. a state of affairs that people desire either for themselves or others. These will be reiterated in the affective system, thus leading to an emotion that will then draw the individual's attention to the element that started this process.

This explains why in some cases it is easier to think for oneself than in others. The concern to understand a situation can get competition from the concern to feel part of the group, and the arguments we come up with ourselves can get competition from the idea that our supervisor is the expert. We get and keep the feeling that we are right when our own ideas are activated above a certain threshold, so they can enter awareness, and there is not a competing representation that gets activated more. But obviously this cannot always be the case.

**Caroline Stankoz** (Ruhr University Bochum): “Gut Microbes and Human Cognition”

Did you know that your gut microbes affect your food intake (Agustí et al. 2018, p. 11)? Choosing a meal from a menu seems to be a cognitive process. Does it span gut microbes? There is a growing field of research on the impact our gut microbiota – the trillions of microbes living in our gut – has on our cognition (Cryan et al. 2019, p. 1878). But what is the nature of this impact? Is the gut microbiota a constitutive part of our cognitive system? Or is it contributing to our cognition in a less direct way? Does the gut microbiota function as an organ? Our natural body parts are self-produced. Does it matter that we have not produced the microbiota ourselves?

Federico Boem, Gabriele Ferretti, and Silvano Zipoli Caiani (2021) argue that our cognition extends internally into our gut microbiota. This phrasing might not be necessary, as they treat the gut microbiota as an organ, a part of our body. Such a relation might be suitably explained in terms of embodied cognition. This family of theories argues that “the body’s interactions with the environment constitute or contribute to cognition in ways that require a new framework for its investigation. Mental processes are not, or not only, computational processes.” (Shapiro & Spaulding 2021) If one understands the gut microbiota as an organ, referring to the body’s role in cognitive processes might be enough to emphasise and explain the gut microbiota’s contribution to cognition.

If one does not understand the gut microbiota as an organ of the human body, one needs more than embodied terminology. In the following, I propose another approach towards internally extended cognition: If we understand the gut wall as something akin to our skin, the contents of our gut can be understood as an internal environment. As such, gut microbes would remain distinct individuals, strictly speaking outside of our body. This picture would provide room for internally extended cognition, since extended cognition is generally meant to span processes outside the human body (Clark & Chalmers 1998) – not only outside the skull (as Boem et al. 2021 use the term). A big question is: To whom does human cognition extend? To individual bacteria, to a group (quorum) of bacteria, or to the whole gut microbiota?

This talk will (i) sketch Boem et al.’s (2021) view on internally extended cognition. Since they view our gut microbiota as an organ, (ii) their account seems to be explainable in terms of embodied cognition. I will (iii) highlight the difference between gut microbiota and former microbes that actually became organs in the course of evolution, such as mitochondria. Next, I will offer (iv) another explanation of the influence of gut microbiota on our cognition in terms of internally extended cognition. I will end

with (v) a brief overview of empirical data, indicating to whom cognition extends – to the whole gut microbiota or to individual microbes.

---

**Saga Norrby** (Utrecht University): “Meditating on Fichte’s I”

“I am, because I am,” said Johann Gottlieb Fichte, and went so far as to say that everything that exists, exists because of this highest act of the intellect – the I’s positing of itself, within itself: “This act is all that ever was, is, or will be; as soon as anything exists, this act exists.”

Fichte’s *Wissenschaftslehre* has since been dismissed by many, including his successors Hegel and Schelling, as a “one-sided, ‘subjective’ idealism and a prime instance of the ‘philosophy of reflection’”. In other words: as an intellectual self-absorption that is out of touch with reality, or how things really are.

This presentation explores a radically different reading, by drawing out the similarities between Fichte’s inward turn – whereby the intellect thinks “simply ‘I am’, without at the same time thinking of something which is not itself” – with the practice of meditation. On this reading, which describes meditation and Fichte’s inward turn as similar acts of remembrance, Fichte’s philosophy comes out as potentially quite practical and worldly, rather than overly intellectual and self-absorbed.

The starting point of my investigation is the question of what it would practically mean for the thinking subject to make itself the object of its own thought, i.e., how we could go about the task of turning thought onto itself. My answer is inferred from Fichte’s vivid description of what happens when a subject thinks about an object which is not itself, and sketches some preliminary similarities to meditation. These preliminary similarities have to do with sensory stimuli, and with seeking to minimize these. For, to see – or otherwise sense – an object invites thinking about this object, and in thinking about an object, Fichte points out, “one disappears into the object; one thinks about the object, but one does not think about oneself as the subject who is doing this thinking”. Meditation, meanwhile, is plausibly described as a practice of remembrance, remembrance of the very fact Fichte says we forget when we turn thought onto objects other than itself – when I meditate, I remember that I am the thinker of my thoughts, not the thoughts themselves.

In the subsequent discussion, important notions are attention and mantra. I compare these, respectively, with Fichte’s notion of complete abstraction, and his ‘I’. At the chore of the analysis,

however, lies the breath. We will see that whether meditation is best understood as an analogy to Fichte's inward turn, or as literally what Fichte's inward turn amounts to, depends on whether the breathing body falls within or without Fichte's 'I'.

---

**Taku Uchiyama** (Georg-August-University Göttingen): "The Capacity View of First-Person Reference"

First-person reference —self-reference by means of 'I'— has some remarkable features. First of all, we never fail to refer to ourselves when we first-personally refer. Further, we thereby never fail to know that we self-refer. These features seem to be well clarified in terms of the reflexive character of first-person reference, namely the character that 'I' refers to the very person who produces a token of this expression. Thus, it is natural to appeal to the token-reflexive rule of 'I' to account for first-person reference. However, such an approach proves problematic. This is because the merely reflexive reference does not capture the self-reflexive aspect of first-person reference, namely the aspect that we refer to ourselves as ourselves when we first-personally refer. Further, any attempt to supplement the referential rule of 'I' to capture this aspect becomes circular. Such an attempt inevitably re-introduces the subject's capacity for the first-person reference, which is the original explanandum. Facing this problem, many scholars argue for a special class of primitive self-awareness. Some identify it with the essential feature of bodily awareness, while others characterize it as agential awareness of actions. Despite this difference, they all try to offer a non-circular account of first-person reference by appealing to a specific type of self-awareness that does not require a capacity for first-person reference but provides a basis for this capacity. One main purpose in this paper is to show that such an approach fails to succeed because the alleged kind of awareness cannot be qualified as self-awareness. Such awareness is, though self-related, essentially selfless. As such, it can be at best construed as merely reflexive awareness, but such awareness falls short of offering a sufficient ground for self-awareness.

This paper also proposes my own first-person reference account, the capacity view. Drawing on an analogous discussion on knowledge-how, this view rejects the idea that there must be a basic class of self-awareness prior to and independent of first-person reference. In my view, self-awareness is not such a pre-condition for the first-person reference. Instead, the former is something that is, ultimately, only manifested by the latter. Accordingly, the view takes self-awareness and first-person reference to be mutually interdependent. Nevertheless, its account is not circular because the view is not committed

to clarifying the one in terms of the other. Rather, it accounts for self-awareness and first-person reference as two aspects of one and the same capacity, namely the capacity of applying 'I' correctly. This capacity is acquired and justified through communication with community members to which one belongs. As such, the capacity in question is not dependent on any privileged perceptual state of individuals or an agential character of actions. Notably, the capacity view does not deny such fundamental features. They surely ground the so-called immunity to error through misidentification. The view only denies that those features serve as a genuine basis for the self-ascription at issue. It is rather the capacity the view puts forward that is crucial for such self-ascription.

---

**Sam Hafkenscheid** (Utrecht University): "Natural Selection: Cause, Feature or Dead End."

The nature of selection has long been a controversial topic in the philosophy of evolutionary biology. Walsh, Lewins and Ariew (2002) argue that natural selection should be understood as a statistical theory analogous to that of the kinetic theory of gases as opposed to a dynamical theory analogous to that of Newtonian mechanics. Authors like Stephens (2004), Brandon and Ramsey (2007), Shapiro and Sober (2007), Northcott (2010) and Hitchcock and Velasco (2014) responded by defending the idea that evolutionary theory should be understood as a theory of forces as argued in Sober's seminal work *The nature of selection* (1993).

The core contention is whether natural selection is a cause of evolution or a feature of evolution. Does natural selection cause evolution or does evolution exhibit natural selection? The dynamical position takes natural selection to be a force driving evolution at the sub-population level while the statistical position takes natural selection to be an emergent process at the population level itself.

The problem with both approaches, however, is that they fail to establish the empirical foundations for their terminology. Both sides fail to adequately distinguish empirical fact and Darwinian interpretation: selection coefficients in evolutionary models are not and cannot settle philosophical disagreement about the nature of selection if the very fact that these model variables are called selection coefficients is in dispute. Only by systematically showing how abstract terms like natural selection are physically instantiated is it possible to make progress: too often authors conflate metaphorical use of natural selection with literal use, resulting in the misguided attempt to identify some particular aspect of physical reality as natural selection.

This article first introduces a purely empirical framework in terms of trait spaces through which to analyze biological evolution in its most general form. It then shows how Darwinian terms like natural selection, fitness and evolutionary forces find natural and well defined expressions in this trait space representation. Unlike previous approaches which tried to analyze the nature of selection in terms of its relation to other abstract concepts like fitness and adaptation, this approach analyzes the nature of selection in terms of the different possible evolutionary trajectories and their causes through trait space, which itself is a direct representation of the physical reality underlying biological evolution.

It is concluded that even though there are many classes of evolutionary forces which have selective effects, none of these classes are inherently selective. Based on this analysis the evolutionary forces of, for example, mutation and migration are empirically identifiable, but the evolutionary forces of natural selection and drift are not. This conclusion has important consequences for the unit of selection debate, our interpretation of the price equation and our understanding of fitness.

---

**Claudio Davini** (University of Pisa): “Explanation, Normativity, and Biological Functions. The Rivalry Between the Etiological Theory and the Causal Role Theory.”

Biological functions are central to many debates – e.g., the rivalry between the consortium called ENCODE and the Human Genome Project–, yet philosophers and scientists disagree about what a theory of biological function is supposed to be or to do. Many of them believe that a theory of biological function should account for the existence of traits to which we attribute a given function, holding that the only way to make sense of the presence of a biological trait is to refer to the effect of the trait in virtue of which natural selection favoured it; many others stress that functional attributions are often made when trying to understand how a biological system works in the present, claiming that the function of a trait is simply its contribution to the functioning of the overall system and that we should shift our attention from why-is-it-there questions towards how-does-it-work questions. My purpose is to argue that the latter approach ultimately implies the former.

First of all, I lay out the most significant desiderata for a theory of biological function that have been developed in the literature over the last decades: a) the distinction between functions and accidents; b) the explanatory dimension of functions, that is, the reason why a given trait is there; c) and their

normative dimension, meaning that a theory of biological function should be able to elucidate that it is possible for a trait to have a function that it cannot currently perform.

Then, I introduce the etiological theory – which states that the function of a trait is whatever it was selected for by natural selection – and, following, the causal role theory – which holds that the function of a trait consists in its current contribution to some system capacity –, and I consider in what way these two theories are linked to the desiderata for a theory of biological function. In this regard, I show that both theories can satisfy the first desideratum, but I add that the causal role theory, in order to effectively make sense of the remaining desiderata, cannot but make an appeal to natural selection. For example, when we identify the causal role of sweat glands in our thermoregulatory system, we are implicitly assuming that this system is meant to achieve some goal, that is, keeping body temperature constant. But the assumption is only justified if the system has evolved by natural selection to achieve that very goal, meaning that we can identify the function of sweat glands with their causal role in the thermoregulatory system only if it is the result of natural selection. The same goes for the third desideratum: in fact, we can affirm that sweat glands are not functioning only if they are not currently performing the function in virtue of which they were selected. So I claim that for these reasons the etiological theory accounts for the above mentioned desiderata better than the causal role theory.

---

**Lara Maassen** (University of Oxford; TU Delft): “Negotiating sex: the role of negotiation in sexual ethics”

This talk focuses on the role of negotiation models in feminist sexual ethics. Feminist theorists have long argued that we need to move our focus beyond consent if we want to advance our sexual ethics. One strategy has been to foreground 1 other communication models, such as negotiation, in the place of consent. Negotiation is usually presented as an approach that can remedy some important issues with consent, such as a heteronormative and flattened focus on invitation-and-reaction.

The aim of my talk is to demonstrate that negotiation models face significant limits that they are ultimately unable to overcome. Consequentially, I suggest that those who are committed to advancing our sexual ethics should think differently about sexual communication: both about the form that it takes and the role that it plays in sexual ethics.

First, I will distinguish two different (but overlapping and often confused) roles that communication rituals play in our sexual ethics. The first is a kind of third-person-metric: one that is used to describe what goes ethically right or wrong within sexual encounters. We do this, for example, when we suggest that sex is ethically wrong because no consent was given. The second is a first-person action-guiding role: the one we use when we promote sexual communication as a ‘pedagogy’ on how to ensure that sex is ethical. For example, we may teach teenagers that they should always verbally ensure that sex is wanted. Both of these roles are, I suggest, taken up by negotiation.

Next, I argue that negotiation faces important limits on both of these levels. First, I suggest that negotiation promotes the wrong kind of communication insofar as it remains reliant on contractual language. Such a notion of sexual communication makes a phenomenological mistake: it fails to capture the simple fact that sexual communication, especially when it goes well, does not operate like a contract: it does not present modes of formality, self-interest, and preformed goals. Hence, ‘negotiation’ does not seem to offer the third-person-metric that we want it to have, nor does it offer an action-guiding idea that someone might follow.

The second objection concerns the role that negotiation plays here. By focusing on language alone, I suggest that theorists overlook other important factors, such as intention, social construction, and power. Casting them in this way risks reinforcing the potential for rote rituals to be played out without the players paying any attention to the background conditions needed to make such rituals valuable. As a result, they fail to account for the fact that negotiation is perfectly compatible with unethical sex. This means, again, that negotiation is both an insufficient metric as well as a wrong action-guiding norm for ethical sex.

The upshot of this is twofold. Firstly, those who are committed to sexual ethics should think more carefully about the content of communication they promote: they should move beyond contractual language. Secondly, they should also think more carefully about the role that sexual communication plays. Rather than casting these rituals as inherently valuable, I argue that we should see them as embedded in a larger web of values and virtues — such as reciprocity, care, and attunement— that are not inherent to communication, but that may nonetheless be enhanced by it.

**Myriam Bellamine** (Utrecht University): “Beyond consent: taking into account the specificities of the intimate sphere in sexual ethics”

Consent is at the center of many public discourses about sexual ethics. While the idea that ‘no means no’ has become uncontroversial, the most interesting ethical questions arise when it comes to determining when ‘yes means yes’. It is accepted that one of the background defects that can make consent invalid is coercion, but the exact definition of coercion is debated.

The philosopher Alan Wertheimer (1996) defines coercion as “when A threatens to make B worse off than she has a right to” (p. 333). In this paper, I argue that this liberal definition fails to present a satisfying account of sexual coercion. To show this, I engage with an example given by Wertheimer, in which a man A threatens to break up with his female partner B if she refuses to have sex with him. I claim, against Wertheimer, that this should be considered coercion even though partner B “has no right that A continues his relationship with B on her preferred terms”

I argue that the language of right is inappropriate because we never have a ‘right’ to a romantic relationship with someone. Rights language is not sufficiently sensitive to allow us to draw the line between acceptable and unacceptable behavior in relationships. Considering the empirical evidence that most cases of sexual coercion for both men and women take place in intimate relationships, I argue that such a case should be considered a paradigmatic case of sexual coercion rather than an edge case, and that it is urgent to create ethical tools that are more adapted to the specific structure of the intimate sphere.

What characterizes the intimate sphere? I argue that key feature of affective bonds is not choice. Though we can decide in romantic relationships who we engage with, once a long-term bond is created, it pre-exists and supersedes the specific content of interactions. This means that the desire for the bond to continue most of the time will supersede other considerations in relationship negotiations. This creates a mutual vulnerability and a power over the other partner that each member of a relationship has the responsibility not to abuse. This vulnerability is however deepened for female partners in heterosexual relationships because of the economic, social and practical dependency on men created by the unequal distribution of resources and power in society, which reinforces the need for heterosexual men to cultivate a sense of responsibility and reflexivity about power.

Moreover, vulnerability is not a defect but has a desirable role to play in the intimate sphere. Sexual relations, more than other activities, should not be up for bargaining because sex comes with a

heightened vulnerability, and giving in to unwanted sex degrades the sense of safety and connection to one's body that is a pre-condition for sexual flourishing.

I conclude that our reflections about sexual ethics and sex education must thus go beyond consent and include reflexivity about power relations, vulnerability, responsibility and creating conditions of safety in which enriching exploration can flourish.

---

**Joost Wijffels** (Utrecht University): “Selective Sexual Preferences and the Limitation of Sexual Autonomy”

Although the practice of choosing sexual partners on the basis of certain selective criteria, also called *Selective Sexual Preferences* (SSP) is deeply ingrained in our society, it is far from innocent. A large body of research has highlighted that SSP mirror racist, sexist, or other oppressive structures. It is mostly minority groups that are excluded or fetishized on the basis of SSP, and members of these groups report feeling objectified and dehumanised by SSP. Several activists and philosophers have consequently deemed the practice of SSP morally objectionable and call for their dismantlement through critical reflection.

This paper sides with these critical voices and argues for the moral objectionability of SSP. To support this argument, this paper draws on Amia Srinivasan's reinvigoration of feminist debates regarding the politics of sexual desire. This analysis shows how the supposed legitimacy of SSP is grounded in the liberal conception of sexual autonomy that aims to liberate individuals' sexual practices from moral scrutiny and poses consent as the sole determining factor for the permissibility of one's sexual practices. The analysis also shows how this conception of sexual autonomy is at odds with a moral responsibility to critically reflect on sexual desires. This responsibility stems from the fact that desires do not pertain solely to isolated individuals, but that they are rooted and, in turn, perpetuate objectionable social structures. A tension between sexual autonomy and a moral duty to critical reflection emerges.

This paper suggests that this tension may be a fabrication. To show this, it reconceptualises sexual autonomy as an integrated sexual perspective that encompasses the biological, individual, and social significance of one's sexuality and desirability. By drawing on research about effeminophobia in the gay community, this reconceptualization illustrates that SSP are a symptom and perpetuation of limited

sexual autonomy. One's SSP are likely to be the result of one's own disintegrated sexual perspective for which one wishes to compensate. The imposition of SSP on others, moreover, is a reification of social norms that can (and in reality, often do) cause the disintegration of others' sexual perspective. Although it may not be sufficient, critical reflection regarding our sexual preferences may be a first step in mitigating the harmful effects of SSP on our sexual autonomy. This paper maintains, therefore, that SSP are morally objectionable because they signify a limitation of our sexual autonomy. It also argues that, rather than being in tension, sexual autonomy and the moral duty to reflect on one's sexual preferences are mutually supportive.

---

**Femke Reniers** (Utrecht University): "Pandora's Lab: Queering science and story through Classical myth"

Based on what feminist philosophy has coined the dualistic epistemology of Western culture, the sciences and the humanities are assumed to be an exclusive dichotomy, with attributes such as objectivity and neutrality associated with the sciences, while the humanities are connected to subjectivity and partiality. This dualism is a part of a greater dualistic epistemology which assigns negative value to some categories while attributing positive value to their "opposites." Examples of these categories are people of colour, emotion, and women, versus the white colonizer, rationality, and men. These dualisms are seen to be interwoven, with the qualities on either side of the divide being associated with one another. This means that the opposition between the sciences and the humanities also plays a role in for example the subjugation of Indigenous knowledge as "less than" Western Euro-American knowledge. The maintaining of any binary in this scheme also supports the existence of other binaries, such as the man-woman gender binary, which is known to be both unrealistic and damaging.

Instead of accepting that the sciences and humanities are indeed inherently and incommensurably different, this thesis posits that the opposition of the "objective" sciences and the "subjective" humanities is actually already incomplete and therefore not a realistic or useful construct. Instead, I present the use of narrative as a valuable method for knowledge-making in science.

Making use of "queer defamiliarisation" as discussed by Helen Palmer, in the first part of this thesis I compare scientific practice to Greek mythology to identify and recognize the already existing but unacknowledged narrative elements in science. While science is stereotypically seen as an objective

source of information, in many instances the scientific method is not sufficient to answer certain questions. Rather, science relies on subjective narrative elements in thought experiments, experiment construction, and knowledge-communication. A queer perspective on myths is used, recognizing a myth as a story with one identity but many different possible forms.

In the second part, using theory from queer, feminist, and Indigenous work on knowledge, method and translation, I argue that the explicit use of narrative story-telling in science can lead to greater accountability, knowledge, and accessibility in the sciences. Rather than attempt to reach unobtainable objectivity in science, I posit that by acknowledging narratives as a valid method of knowledge-making, we are made aware of the locality of our perspectives, even in science.

Translation between different forms of knowledge, similarly to how myths can be told differently by different story-tellers, allows us to glean new patterns and questions that we wouldn't have seen from within our own context. Lastly, the acknowledgement of narrative as a valuable method could create more space within Western culture for recognition of other cultures' knowledge-practices that have long been deserving of equal consideration, such as storytelling-based knowledge in Indigenous cultures.

By allowing for the acknowledgment and further incorporation of narrative into the sciences, this thesis shows how the concept of knowledge can be queered to be less dependent on strict boundaries between disciplines, thereby making space for creative development of new ideas and perspectives both within and beyond scientific practice.

---

**Roxane Pret Théodore** (University College Dublin): “Rethinking Political Organization from a Feminist Standpoint: Politicizing an Ethics of Care and Vulnerability”

Sexism in the political field comes from deep-rooted patriarchal structural causes. As a result, political institutions from so-called “western” democracies continue to exclude women and other minoritized groups, hence reproducing and maintaining this bias. However, were these political institutions subverted, they could become tools of resistance and change, which would contribute to minoritized groups' political empowerment. Thus, there is an urgency in responding to what is referred to as a “crisis of democracy” (Manin, 1995), of which political sexism is a part, by rethinking the way we organize ourselves politically. Democracy in modernity has taken the shape of a representative system. As such, there has always been a gap between electors and elected. The “crisis of democracy” described

in the theoretical literature, however, specifically refers to the exponential widening of this gap, which leaves a growing number of electors feeling underrepresented (Manin, 1995). Therefore, the central questions that arise are the following: What instruments can be used to remedy this situation of persistent structural injustice? How would a political reassessment of feminist concepts help shape a more democratic organization of our political institutions?

In this paper, I aim at developing an alternative philosophical account demonstrating how feminist concepts, and more precisely an ethics of care and vulnerability, would help improve our political policies and practices, when applied to the political field. I suggest that reviewing our political institutions of representation from a feminist standpoint allows for a rethinking of these structures and practices in a more democratic framework, opening the field to all minoritized groups that have been and still are excluded from it. My objective is to examine what an ethics of care and vulnerability would entail conceptually when applied to the political field. I rely on a non-binary definition of the figure of the “caregiver” (Tronto, 2013) and on a positive perspective on vulnerability (Butler, 2004), understood as a relational dynamic of interdependency and solidarity rather than a demonstration of weakness, or in other words, as a “general openness towards the other” (Petherbridge, 2016). I address the ambiguity vulnerability carries in the literature by asking a follow-up question to those previously stated: if caring relations, as a response to the expression of our inherent vulnerability, allowed us to rethink the way we organize ourselves politically, how can we overcome the ambiguity of vulnerability itself? How is such a positive reframing of the concept of vulnerability possible?

To this end, in the first section, I look more closely at Adriana Cavarero’s work (Cavarero, 2011) and contrast her approach to vulnerability with that of Judith Butler (Butler, 2004). In the second section, I suggest that in Cavarero’s specific treatment of violence, there is an underlying radical approach to vulnerability and care that would allow for the emancipation of vulnerability from its dichotomous ties to violence. This could have political ramifications for addressing a revisited account of vulnerability that informs alternative political institutions and practices, which I explore in the final section.

---

**Joseph Conrad** (University of Edinburgh): “The emotions-responsive corporation: why corporations are appropriate targets of moral blame”

In this paper, I defend the attribution of moral responsibility to corporations. Particularly, I argue that corporations are appropriate targets of moral blame. While we routinely blame corporations, it is not obvious how they *can* be appropriate targets of blame, and why they *should* be. The discussion of corporate blame links to the broader and important question: To what extent do corporations form part of our moral community?

My argument proceeds in three steps. First, I use the examples of big fossil fuels corporations, such as ExxonMobil and Shell, to illustrate egregious corporate moral wrongdoing and the emotive reactions that such corporate wrongdoing can elicit in us. My point is to illustrate the common *intuition* to blame corporations for their wrongdoing.

Second, I argue that corporations *can* be blamed wherever it is appropriate. I respond to what I call the *moral emotions objection*, stating that corporations lack the necessary agential capacities for blameworthiness. According to this objection, blame involves the expression of moral emotions, but corporations lack a capacity for affect, particularly the capacity to feel guilt. Hence, it is argued that there is little sense in blaming corporations because they cannot feel guilt (Sepinwall 2017).

My argument against this objection will be that corporations, though not strictly speaking capable of experiencing a phenomenology of moral emotions, can nevertheless recognize and respond to our expressions of moral emotions so that our moral communication with them still makes sense. I contend that if we conceive of corporations as *reasons-responsive agents*, which we should, we are at the same time committed to the claim that they are also *emotions-responsive agents*. By this I mean that corporations can be adequately sensitive to the moral emotions that bear on their actions.

I briefly address the broader claim that corporations, being ‘affectless creatures’, are ill-suited as members of our moral community. In response to this, I contend that rather than excluding corporations from the moral realm, we need to grapple with their participation (or intrusion) in our moral life.

Third, I argue that corporations *should* be blamed wherever it is appropriate. I begin by responding to what I call the *collective punishment objection*, stating that holding corporations morally responsible results in illegitimate practices of collective punishment (Hasnas 2017). Here, I show that corporations can, at least theoretically, be held collectively responsible without thereby negatively affecting any individual human person. Further, even if individual human persons are negatively affected, this can often be justified on the grounds of individual role responsibility, voluntary risk-taking, or benefitting from a wrongdoing.

I conclude by offering a positive argument for why corporations should be blamed wherever it is appropriate. Building on Manuel Vargas' (2013) instrumentalist theory of moral responsibility with 'Strawsonian adjustments', I argue that we can justify the practice of holding corporations morally responsible on broadly consequentialist grounds. The basic idea is that blaming (and praising) corporations can foster a better form of corporate agency that is more sensitive to moral concerns.

---

**Eline Gerritsen** (University of St Andrews & University of Groningen): "Conventional Norms: The blind spot of metaethics"

Metaethics is no longer concerned only with morality: attention is increasingly paid to other types of norms as well. Metanormative theorists have turned to questions about the normativity of the law, rationality, aesthetic values, political principles and other types of norms and standards. While it is widely assumed that morality has a serious normative force, there are vigorous debates on the normative status of other types of norms. Yet, one important class of norms receives hardly any attention in this context: conventional norms. By conventional norms I mean informal, non-codified and often implicit norms that emerge through interactions within a society (Bicchieri 2006). It includes norms about how to greet others, how to form a queue and what to wear to a funeral. Metaethicists have assumed, without serious discussion, that conventional norms are not normative in a significant sense. It is telling that the normative force of morality is often highlighted by contrasting it to etiquette (e.g. Mantel 2019, 217; Cuneo 2016, 77; Plunkett 2019, 113; Werner 2018, 617; Olson 2014, 120). Yet, in reality, conventional norms take an important place in our lives: they regulate many details of our behaviour and shape how we act in general. We feel bound by conventional norms and overwhelmingly comply with them, sanctioning others who do not. Since real agents act within a social context, ignoring or dismissing conventional norms constitutes a blind spot that makes metaethics less connected to real, everyday experiences of normativity.

In this talk, I will explain the problem, diagnose a cause of it, and show how we can do better. I will highlight that research disciplines both within and outside of philosophy – including psychology, evolutionary anthropology, and social ontology – take conventional norms seriously in a way that is not currently reflected in metaethics. The blind spot therefore extends both to our lived experiences and to academic treatments of the topic in other fields. I will argue that metaethics's dismissal of conventional norms is explained by a popular but unnecessarily restrictive understanding of

normativity. There is an increasingly recognised distinction between formal normativity, which all norms have, and authoritative normativity, which is more significant. The question of interest is whether conventional norms can be authoritatively normative. The standard view of authoritative normativity stands in the way of an open discussion of this: in recent literature, a common but unwarranted assumption can be found that authoritative normativity must be irreducible to natural facts (e.g. Olson 2014; Parfit 2011). It is plausible that such a form of normativity could not apply to norms that are simply based in natural facts about conventions and others' expectations. I will suggest an alternative understanding of authoritative normativity, based on Humean constructivism (Street 2008), which does allow us to ascribe a significant normative force to conventional norms. This ability to validate the bindingness of conventional norms that we experience in our everyday lives, I will argue, is an important advantage of Humean accounts of normativity.

---

**Charlie Blunden** (Utrecht University): “Are Our Moral Progress Judgements WEIRD, and Is It a Problem?”

The philosophy of moral progress has recently seen a renewed interest across several disciplines (e.g., Anderson, 2014; Anomaly, 2017; Buchanan, 2020; Buchanan & Powell, 2018; Buddeberg, 2019; Eriksen, 2019; Hermann, 2019; Hopster, 2020; Huemer, 2016; Kitcher, 2021; Pinker, 2018; Schinkel & de Ruyter, 2017; Tam, 2020). When talking about moral progress, one of the first things authors must explain is what kinds of developments they count as progress. Authors usually do not first stake out their comprehensive theory and then derive their examples of moral progress from this theory. They think this is unnecessary, and I agree with them. Instead, authors typically rely on canonical examples of moral progress that “are very widely held and stable under favorable social-epistemic conditions” (Buchanan and Powell 2019, 288). Canonical examples include the legal abolition of slavery (Pleasant 2010; Anderson 2014; Kumar and Campbell 2016; Jamieson 2017; Hermann 2019), the ongoing struggle for equal rights and emancipation for women (Eriksen 2019; Kitcher 2011; Lowe 2019), the ongoing movement for LGBT+ emancipation and equality (Kumar, Kodipady, and Young 2022; Kitcher 2021), and, for an example from further back in history, several authors have noted how, over time, different ancient and medieval legal codes changed to place greater emphasis on *intentionality* in establishing responsibility for crimes and establishing who should be punished (Kitcher 2011; Buchanan and Powell 2017).

However, thus far, insufficient attention has been paid to how widely shared these judgements are: in particular, whether the canonical examples offered in the literature might be skewed towards examples of progress that appeal to WEIRD people (people who are Western, Educated, and from Industrialised, Rich, and Democratic countries). A growing body of evidence shows that WEIRD psychology is often very atypical and unrepresentative of human psychology in general (Henrich, Heine, and Norenzayan 2010; Barrett 2020; Henrich 2020), and it has been noted that this research might have philosophical implications (Henrich et al. 2022). Furthermore, we have evidence that at least some of the canonical examples of moral progress are supported most strongly in WEIRD countries: most notably the progressiveness of LGBT+ emancipation (Poushter and Kent 2020; Welzel 2013; Inglehart 2018; Mendos et al. 2020; Kumar, Kodipady, and Young 2022, 34–36) and the importance of intentionality in ascribing moral responsibility and assigning punishments (Curtin et al. 2020; McNamara et al. 2019).

In my paper, I will first summarise this evidence to establish the extent to which the canonical moral progress judgements in the literature are WEIRD. I will then explore whether this WEIRD-ness is a problem for the philosophy of moral progress. In particular, I will survey various philosophical strategies one could adopt in order to (a) admit that certain moral judgements are more strongly supported in WEIRD societies than in others, but (b) maintain that this fact doesn't undermine judgements of progress based on these moral judgements. I will aim to assess how plausible these strategies are, and, thus, how much of a problem WEIRD-ness is for the philosophy of moral progress.

---

**Jenny Zhang** (University of Edinburgh): “A new direction for moral psychology? A case for moral experiences qua experiences”

One prominent strand of discussion in moral psychology is sparked by Jonathan Haidt's (2001) seminal paper, in which he proposes the Social Intuitionist Model (SIM) of moral judgement. SIM is proposed as an alternative to the rationalist approaches, and it differs from them in two important ways: a) it emphasises that social and cultural influences play a significant role in shaping moral judgements; b) it proposes that moral intuitions, as opposed to moral reasons, usually directly cause moral judgements. The debate that ensues Haidt's proposal primarily revolves around identifying or establishing a theory of moral judgement that best accommodates the growing empirical evidence on reasoning and moral

behaviour, one that wins the “contest between models of the partnership between reasoning and intuition” (Haidt, 2010, p. 182-183).

I argue that the abovementioned debate is empirically informed in a skewed way, and headed towards a wrong direction. It is skewed because many of the studies that it appeals to are conducted in laboratory settings, devoid of the nuances and complexity present in everyday moral experiences. The discussion is headed towards a wrong direction because of the assumption that empirical evidence is useful insofar as it gets us closer to building an overarching model, a well-organised, all-encompassing outlook of moral psychology. Under this picture, our moral experiences are of no value other than that they are instantiations of some mechanisms or processes underpinning morality.

I believe that our moral experiences are significant qua experiences. They are rich, messy and particular to us, and we experience them as such. Importantly, we often take them to be central to our moral life. This, I think, points to the possibility that any reductionist or foundationalist accounts of moral judgement inevitably overlook, mischaracterise or oversimplify (at least some) morally relevant experiences. To capture the complexity, richness and sometimes difficulty of moral life, it is necessary to recognise the irreducible nature and immense value of moral experiences. This requires that we shift and expand the locus of discussion that has preoccupied debates in moral psychology for the past few decades: instead of confining the key issues to the extent to which reasoning or intuition plays a role in moral judgement formation, what neurobiological mechanisms or processes are involved in moral development, or the extent to which social or cultural conditioning exerts influence on moral behaviour, we need to see moral functioning as necessarily multi-faceted, interactive, and importantly, experiential.

In my paper, I first outline the discussion following Haidt’s (2001) proposal. I then draw on and adapt Cora Diamond’s (2008) insights to show that there is something fundamentally amiss in the discourse. Specifically, I argue that the debate neglects, sidelines or detracts from what moral experiences comprise, and what it is like for someone to undergo these experiences. I suggest that to rectify the problem, we need to look beyond studies conducted in laboratory, and reflect upon the status, characteristics and constituents of moral experiences qua experiences.

---

**Migdalia Arcila-Valenzuela** (Cornell University): “The Recalcitrant Body: Chronic Pain and Moral Agency”

Chronic Pain Syndrome (CPS) is the leading cause of disability around the world, affecting 30% of the population (Cohen et al. 2021). However, philosophical theorizing about moral agency have paid little attention to this profound debilitating condition. In this paper, I argue that this neglect is a mistake: Chronic Pain Syndrome is a multipronged attack to moral agency that has been overlooked in the literature.

In the first section, I explain, from a neurobiological perspective, what CPS is and what its major effects are. Then, based on patient’s reports, and my own experience as a CPS sufferer, I paint a more detailed picture of the effects that the syndrome has on people’s lives. In section II I present three relevant features of moral agency in light of which we can understand the devastating effects of CPS. In the moral psychology literature, moral agents are frequently characterized as planning agents (Bratman 2000): we are diachronically extended and future-oriented, and we want our agency to be effective over time. Prominent features of planning agency identified in the literature are (i) a sense of one’s agency or moral self (Prinz & Nichols, 2016), since we want our plans to be executed in the future by us; (ii) collaboration (Doris 2015; Bratman 2010), sociality facilitates the exercise of individual agency; and (iii) hope as a unifying and grounding force of agency (McGeer, 2004; Calhoun, 2018). I argue that (i) (ii) and (iii) are significantly impaired by CPS.

To conclude, I’ll consider a fourth feature of agency less commonly identified in the moral psychology literature: attention. A distinctive feature of pain, i.e. intrusiveness (Martinez 2015), prevents people from attending to their relationships and projects in the way required for a planning agency. Therefore, CPS, by creating a pervasive “practical distraction” (Cheyne et al. 2009), undermines the attentional capacities required for the exercise of one’s agency.

---

**Lily Tappe** (University of Marburg): “Are there Moral Reasons to Suspend Judgement on God's Existence”

Imagine a scenario of a person trying to find information and form a belief about a particular fact X, e.g. Tim is trying to find out whether Mary is dating Joe. Conducting excessive research on the topic and gathering every information available, the person finds their evidence inconclusive. However, they

opt for believing that X independently of their basis of evidence. Maybe an equal amount of Mary's and Joe's friends deny and confirm they are an item. Since Tim finds them equally credible and doesn't have access to other sources, his epistemic reasons for and against believing X cancel each other out. Nevertheless, he is a devoted advocate of true love and soon convinces himself that the romantic involvement of Mary and Joe is material for the latest talk of the town.

Beyond the notion that one shouldn't stick their nose in other peoples' love affairs nor engage in gossiping and speculating about such delicate matters in public, I hold that this scenario appeals to two central notions of normative epistemic interest. First, it is wrong to maintain a belief despite the awareness that it is not supported by evidence. Second, if the evidence is inconclusive or otherwise insufficient to support a decision, one should not make a judgement. Both of these sentiments are captured in philosophical thought, the former as a way in which norms should guide epistemic practice in the "ethics of belief". The latter appeal to suspension of judgement is a form of Agnosticism. The scenario presents a case of how a theory from the ethics of belief, namely evidentialism, can overlap with and support Agnosticism. I want to investigate whether this overlap holds for the locus classicus of Agnosticism: the existence of (a) god(s). More specifically, I want to investigate whether moral evidentialism can provide arguments for being agnostic about God's existence.

Merging these lines of thought is not a new idea. Among Victorian secularists such as William Kingdon Clifford and Thomas Henry Huxley, the claim that norms should guide our belief formation was strongly intertwined and discussed within debates about religious belief, the existence of God and the authority religion should have. However, their positions on Agnosticism and evidentialism are not explicitly combined or used to support each other. Further, they call for immense clarification, deradicalisation and modernisation. Therefore, I aim to develop a moderate and modern version of moral evidentialism and identify the argumentative support it can offer for existing agnostic theories.

I will roughly present my argument in the following way while the focus of my presentation is to defend

1) against multiple charges such as the problem of assuming doxastic voluntarism.

1) If there are moral reasons to form beliefs in accordance with one's evidence, then there are moral reasons to suspend judgement when the evidence is inconclusive.

2) The moral reasons apply to forming belief about God's (non-)existence.

3) The evidence for God's (non-)existence is inconclusive.

C) There are moral reasons to suspend judgement about God's (non-)existence.

## 13th of May (Friday)

**Fengyuan Liu** (Utrecht University): “The Concept of Individual in Spinoza: A Dynamic System?”

The concept of the individual has always been a central topic for philosophical discussions. It is a symbol of the connection between single human existence and their capacity of carrying out generality. Among all the individual theories in history, Spinoza’s theory of individual should arise our attention, not only because its existence has been largely overlooked by the Spinozian scholarship, but also because of the particular definition used by Spinoza. The Hegelian interpretation of Spinoza has dominated scholars’ understanding of Spinoza’s philosophy for a considerable period. (Renz, 2018; Melamed, 2010) According to Hegel, there is no individual left in Spinoza’s system. Everything has fallen into his concept of the only substance.

However, recent studies have provided a quite different picture. The more systematic investigations of Spinoza’s early works and the Ethics suggest that Spinoza is not only considering the substance itself but also leaning a lot of points on the existence of the finite and individual as special kinds of modes. (Di Poppa, 2010; Renz, 2018; Schmaltz, 2021) According to the recent studies, the dynamical elements in Spinoza’s philosophy are not only adequate causation but also the model of production in his discussion of the finites. (Sangiaco and Nachtomy, 2018) These are all strong evidence to prove that Spinoza’s concept of individual needs to be considered much more seriously.

Even though much work has been done to clear the view that the individual does play a role in Spinoza’s philosophy, the nature of the individual in Spinoza’s system has remained largely unexplored. In his paper, Melamed claimed that Spinoza might deploy a functional use of the concept of the individual. (Melamed, 2010) Different from this approach, this paper is trying to show that Spinoza takes the “individual” as a system composed of simple bodies based on the studies of Di Poppa and Sangiacomo.

There are two points discussed in this paper: Firstly, based upon the previous studies, how does Spinoza’s concept of the individual is different from the Hegelian interpretation and even from the “weak individual” defense offered by Melamed. We are going to prove that the individual is a model as a system for Spinoza. Secondly, since the individual as a system is composed according to the rule of velocity, the individual is not a purely static substantial existence. Referring to Deleuzian interpretation, an individual could be understood as an expression of the substance. It is not the product of the expression. The individual is the expressive process itself. We are going to prove that dynamism is an essential aspect of the existence of the individual.

---

**Harmen Grootenhuis** (Groningen University): “All for one and one for all; how a theory of grounding explains Spinoza's monist metaphysics”

Spinoza is a controversial philosopher. His view on God made him a target for his seventeenth century contemporaries and still divides the modern-day readers of his works. In short, Spinoza is a monist metaphysician that puts the unity of reality at the forefront and only allows for a derived diversity of things in the world. That unity is the unity of God or substance – two terms used interchangeably throughout Spinoza’s works. The perhaps most famous line of his magnum opus – the *Ethics* – is: “Whatever is, is in God, and nothing can be or be conceived without God.”

In this paper, I will trim the thorny issue of what it means to ‘be in God’. I argue that an intricate system of *grounding relations* is at work in Spinozist metaphysics. Drawing on Kit Fine’s theory of grounding, it becomes possible to dissect Spinoza’s early modern monist metaphysics using an analytical tool from modern day metaphysics – so I claim. The benefits of this procedure are twofold: not only does Spinoza scholarship profit from a coherent interpretation of substance inherence, but scholars working on metaphysical monism in general can also draw on the case study of Spinoza as an example of how ontological diversity can be integrated in a (historical or contemporary) monist framework.

The paper assesses two attempts at harmonizing the world’s diversity with Spinoza’s claim that the cosmos is ultimately united in God – or substance – before introducing my adaptation of Fine’s theory of grounding to the context of Spinoza. The first harmonizing attempt is Aristotelian in nature and interprets the diversity of things as a diversity of *properties* belonging to the one substance. The second attempt utilizes Spinoza’s notion of causation to interpret the diversity of things as a diversity of *effects* following from one common cause i.e. the one substance. I reject both attempts. The first attempt is incomplete as it only explains how the *essences* of individual things are in substance, without explaining how *actual* things are united in one substance. The second attempt fails due to the fact that Spinoza explicitly denies the possibility of one substance causing a multitude of things. I then propose that the actual plurality of things is dependent on substance via grounding relations, and that multiple types of grounding are at play. One type explains how the essences of the individual and diverse things are united in substance, and a second type of grounding explains how these things inhere in substance insofar as they are actual. The availability of multiple types of grounding provides multiple dependence

relations and for that reason improves on the more austere Aristotelian suggestion. It also improves on the second attempt by not casting the dependence relations between substance and the actual plurality of the cosmos as causal relations.

---

**Alberto Cavallarin** (Utrecht University): “Mystical Ontology: Can Mystical Experiences Ground Metaphysical Claims?”

A so-called mystical experience is the experience of losing one’s own sense of subjectivity, one’s sense of being an individual opposed to an external world, and of becoming one with “the universe”, or “God”, or “the universal consciousness” (depending on one’s preferred terminology and background). The fact that mystical states are a class of experiences is beyond any doubt. Said states can be induced fairly easily through different practices, such as the use of psychedelic substances, breathing techniques, meditation, prayers and other. What is far more complicated to assess is their metaphysical status: are mystical experiences just hallucinations – “comforting delusions” – or can they tell us something about the nature of reality? Most people who have had these experiences themselves report a certain sense of authority, or even of absolute certainty; however, the mainstream scientific consensus seems to be, so far, on the hallucination hypothesis. The slow decriminalization of psychedelic substances worldwide, accompanied by the spread in popularity of psychedelic-assisted therapy, is making mysticism and its implications relevant again; and while a good amount of ink has been spilled on the matter in the past, more work is needed. This paper aims, first of all, at summarizing the academic debate on the metaphysical relevance of mysticism. After assessing all the different major arguments in favor and against grounding metaphysics in mystical experiences, I will conclude that none of them is particularly convincing (both in an academic and non-academic context) and I will propose an argument of my own – I will propose a pragmatic argument for grounding metaphysical claims in mystical states. To put it sharply, this paper ultimately aims at showing that those who have had these experiences, and that came out of them with a metaphysical insight, have good reasons – based on pragmatic considerations – to believe that they were not just hallucinating. The next question is: what is this metaphysical insight? I will argue that, given the complexity of assessing what a mystical experience precisely is, a mystical ontology should be minimal. I will suggest that, if one agrees with the characterization of mysticism I propose in this paper, the only metaphysical claim we can put forward is: at a fundamental level there exist no subjects and no objects; ultimate reality must be some

sort of unity. A more detailed characterization of “unity” would be, I believe, speculative. While my claims are not particularly revolutionary or unheard of, I believe this paper can work as a basis for reviving this fundamental and extremely relevant debate.

---

**Samuel Ferns** (University College Dublin): “Epistemic Recognition or Epistemic Freedom? Two Paths to Social Epistemology in Honneth's Work”

Work within the field of social epistemology has sought to explain knowledge as a normative and power-embedded phenomenon through such concepts as “epistemic injustice” (e.g., Fricker 2007), “ignorance” (e.g., Sullivan and Tuana 2007; Peels and Blauuw, 2016), and “ideology” (e.g., Mills 2017). The tradition of Critical Theory, stemming from the work of the Frankfurt School of the 1930s and intended to constitute a theory of society with emancipatory intent, has from its beginning taken up the issues of knowledge and non-knowledge as matters of normativity and power, and for this reason has much to offer contemporary research. Notably, recent studies (for example, Congdon 2017, 2018; Giladi 2018, 2020) have sought to establish the epistemic import of Axel Honneth’s recognition-theoretic iteration of critical theory. These studies have been concerned (a) to utilise Honneth’s recognition theory to establish the normative nature of the concept “knower”, such that wrongs suffered by knowers can be made sense of as failures of recognition; and (b) to make clear precisely where matters of epistemic significance fit within the architecture of Honneth’s recognitive theory: Does epistemic recognition pertain to each of the three spheres of recognition delineated by Honneth (affective relationships, legal rights, and social esteem), or is an alternative, specifically epistemic category of recognition required?

While these studies make strong cases for the epistemic significance of Honneth’s theory of recognition, they are nonetheless limited in two ways. First, thus far such work has concerned itself only with Honneth’s early theory as elaborated in *The Struggle for Recognition* (Honneth 1995). However, Honneth has since developed an institutional account of social freedom on the model of Hegel’s mature political philosophy in *Freedom’s Right* (Honneth 2014). The former provides the conceptual means to establish an anthropologically grounded theory of *epistemic recognition*. Meanwhile, by way of the latter’s claim that personal relationships, the market economy, and the democratic public sphere and constitutional state embody so many “institutionalized spheres of the promise of freedom” (Honneth 2014, vii), one can ask whether these spheres make and fulfil a promise of *epistemic freedom*.

In this paper, my first aim is thus to ask and answer the following questions: (i) As they pertain to Honneth's work, what are the differences between "epistemic recognition" and "epistemic freedom"? (ii) What are the relative strengths and weaknesses of an anthropological account of epistemic recognition and an institutional account of epistemic freedom? Second, while work in this area has so far engaged with Honneth's recognition theory in its fully articulated form - as presented in *The Struggle for Recognition* - it does not attempt to reconsider from the standpoint of Critical Theory's potential contributions to social epistemology the concerns and arguments along the theoretical path that brought Honneth to formulate his critical theory as one of recognition. My second aim in this paper is therefore to evaluate the paradigm's conceptual foundations to establish its possible contributions to current problems in social epistemology. Ought any future critical theory of knowledge be grounded in the paradigm of recognition?

---

**Giorgia Foti** (University of Bologna): "Monitoring goes social: a new account of epistemic vigilance in the epistemology of testimony."

The debate between reductionists and antireductionists in the epistemology of testimony has hitherto largely spanned around the question whether the recipient of a piece of testimony should engage in some sort of monitoring for trustworthiness in order to get justified testimony-based beliefs. It has been argued that a kind of *passive* monitoring is actually compatible with antireductionism, in that it can contribute to the overall reliability of testimony as an epistemic source, without implying a reduction of testimonial justification to non-testimonial evidence (Henderson and Goldberg, 2006). In this respect, I will focus on the distinction between moderate and strong versions of social anti-reductionism. They both give a social explanation of *default prima facie* justification provided by testimony. According to the latter, a recipient of testimony is *prima facie* justified to believe the speaker on mere say-so, in absence of undefeated defeaters (Simion, 2020). The former, on the contrary, adds some further requirements in order to warrant a rational testimonial uptake, such as a counterfactual sensitivity to defeaters on behalf of the recipient (Goldberg, 2014) or a monitoring for honesty *filter* (Graham, 2010 and 2020).

This paper has two aims: one negative, one positive. The first two sections will be devoted to reject the moderate-antireductionist claim that some kind of epistemic vigilance toward the speaker in the context of a testimonial uptake can make an effective contribution to the overall reliability of testimony

(Michelian, 2010). I will motivate this rejection on both empirical and theoretical grounds. The literature on deception detection clearly shows that we are not able to catch a liar or an incompetent speaker when we face one (e.g. Ekman, 1996). Furthermore, on the theoretical side, even if some kind of epistemic vigilance embedded in the mechanism of comprehension did indeed contribute to the reliability of testimonial exchanges, I think it could be aptly captured by the no defeater condition alone, as strong anti-reductionists maintain—especially if we included also *normative* defeaters.

Does this mean that we can do without the notion of epistemic vigilance when providing an explanation of the reliability of testimony? The *pars construens* of my paper will be devoted to show that we cannot, if we reshape the notion of epistemic vigilance in non-individualistic terms. I will argue that we are not able to catch a liar *on the spot*, but we are much better in exercising *ex-post* monitoring or after-the-fact deception detection. A speaker is motivated to comply with the norm of assertion when giving testimony because of the *reputational costs* she will have to pay if it will be discovered that she didn't (Wiener, *forthcoming*). Information about the past performance of a speaker is then 'stored' in the form of reputational clues (Origgi, 2012) which are *available* to receivers as members of a given social group and play the role of normative defeaters in future testimonial exchanges.

---

**Yasha Sapir** (University of Southern California): "Deliberative inquiry"

Inquiry is an activity where an agent tries to answer a question. A popular view is that the aim of inquiry is at least knowledge or better (Williamson 2000, Schaffer 2005, Kelp 2018, Sapir and van Elswyk 2021, a.o.) If so, then it's natural to think that an agent may not settle their inquiry until at least they know the answer to their question

An objection: if knowledge is sufficiently hard to achieve then an agent may settle their inquiry before they know an answer (cf. Avery 2021, Falbo *forthcoming*). But in such cases, though you may stop your inquiry, you may not settle your inquiry. Consider chess. Chess players can stop playing before there's been a win or draw. But settling play under such circumstances can't be done. Nevertheless I defend that agents may settle their inquiries even under conditions of extreme uncertainty.

I start by considering a view that's distinct but related:

The exhaustivity requirement: If you only know an in-exhaustive answer to your question then you may continue inquiring into the question.

This view is influential in the literature on inquiry (cf. Friedman 2017). But a consideration from the literature on the semantics and pragmatics of questions demonstrates that it's false.

Consider the case of mention-some questions (cf. van Rooy (2003)). If someone asks “where can I get coffee” an in-exhaustive answer will suffice. This lesson from the semantics and pragmatics of questions applies to the case of inquiry. Once you learn of a few good places to get coffee, you should usually settle your inquiry.

I extend a claim by van Rooy (2003) to inquiry and posit that one reason knowledge of incomplete answers can render continued inquiry impermissible is that sometimes we inquire in order to figure out what to do. In such cases we look for answers that will help us make decisions – not necessarily answers which exhaustively answer the question guiding our inquiry. I call this kind of inquiry deliberative inquiry.

If decision problems regulate things like informativeness, then we should think they also regulate when deliberative inquirers may settle their inquiries. The thought is this. First, decision problems quite plausibly regulate when deliberative inquirers may stop inquiring. And if so, then we should think that deliberative inquiry may be permissibly stopped when the expected cost of acquiring additional information is not exceeded by the expected cost. But this can obtain even under conditions of extreme uncertainty.

This result extends to when agents are permitted to settle their deliberative inquiries. It's permissible to settle a process just in case it's permissible to stop the process and the internal purpose of the process has been attained. The internal purpose of deliberative inquiry is to make a decision. Thus, when an agent permissibly stops their deliberative inquiry and then goes on to make a decision, they may then settle their inquiry, too – even if they are extremely uncertain of what is true.

---

**William Gopal** (Amsterdam University): “Toward An Account of Epistemic Trust Suitable for Humans and Artificial Epistemic Agents”

Accounts of trust (Baier, 1995; Jones, 1996) typically focus on cases of human-to-human trust. However, the digital revolution has engendered a shift in philosophical accounts of trust toward cases of trust including technologies (Ess, 2020; Weckert & Soltanzadeh, 2020; Coeckelbergh, 2012) as well as conceptualisations of moral trust which sufficiently capture human-to-artificial agent relations

(Taddeo & Floridi, 2011; Taddeo, 2009; Taddeo, 2010; Grodzinsky et.al, 2010; Grodzinsky et.al, 2020). However, there is little discussion of the notion of *epistemic* trust in such contexts. Whilst some examine the role epistemic trust plays in science (Rolin, 2020, Wagenknecht, 2016; Almassi, 2012; Grasswick, 2010) comparatively less attention has been paid to whether epistemic trust obtains between humans and artificial agents. Consequently, I suggest that redesigning the concept of epistemic trust is necessary given the ubiquitous presence of artificial epistemic agents (AEAs) in epistemic contexts. In this paper I provide a novel account of epistemic trust from an informational approach.

In the first part of the talk, I argue in favour of examining this issue from within the philosophy of technoscience and technoscientific research contexts. Then, I introduce an example in which human-to-AEA epistemic trust may occur; that of academic search-engines. Following this, I suggest that examining epistemic trust in technoscience aids in explicating shortcomings of social epistemological approaches. Specifically, that social epistemology fails to consider the presence of technological artefacts central to contemporaneous methods of the production and dissemination of knowledge. Then, I turn to definitions of artificial agent (AA) and artificial epistemic agent, and highlight the strengths of accounts which describe human-to-artificial agent trust (Grodzinsky et.al, 2011, 22-23; 2020, 305; Taddeo, 2009, 24-25). Following this, I critique McCraw's (2015) notion of epistemic trust if applied to human-AEA relations. I then turn to the *pars construens* of my talk in which I provide, from an informational approach, the necessary and sufficient conditions for epistemic trust to obtain between humans and AEAs. Thereby providing a novel account of epistemic trust. Following this, I specify the conditions under which one is more, or less, justified when placing epistemic trust in an AEA. Finally, I address whether AEAs possess any novel epistemic responsibilities in virtue of epistemic trust obtaining between humans and AEAs; suggesting that whilst AEAs themselves do not possess epistemic responsibilities, those developing AEAs do possess additional epistemic responsibilities pertaining to the epistemic function of the AEA in question.

---

**Hagge Cadenas** (University of San Diego, California): “Cultural Evolution and Epistemology”

Introduction Here I argue that epistemic practices (i.e., the human-structured learning environment) are cultural-evolutionary adaptations that function to solve the puzzle of communication.

§1 The Puzzle of Communication Communication requires adaptations from the sender and receiver. Now, organisms can use signals to deceive. If signals are unreliable, then evolution should select against those very adaptations. How is it, then, that communication is stable? Mercier (2020) proposes that humans developed “open vigilance mechanisms” to keep communication stable. These include: checking for plausibility, attending to reliable sources, and checking for cues of trustworthiness.

§2 Michaelson’s Objection Michaelson’s (2018) philosophical objection to Mercier is that open vigilance presupposes the reliability of testimony. For example, plausibility checking vets sources based on what one knows, but what one knows depends on what others have told us. I raise two problems with Michaelson’s argument. First, Michaelson’s conclusion leaves it unclear what role open vigilance plays in knowledge. Second, Michaelson’s argument presupposes an overly rigid ordering of justification. The next section builds an account based on what we learn from our critique of Michaelson.

§3 The Human Structured Learning Environment Here, I introduce Sterelny’s (2012) concept of the human-structured learning environment. I then argue that open vigilance mechanisms can structure the learning environment. The structuring is complex because it involves various individuals with differing motivations to vet each other; it is dynamic because no one person occupies a fixed position. Michaelson was right to say that open vigilance requires trust, but he was wrong to neglect that trust requires vigilance—though it is not necessarily the trustor who must be vigilant. One way of structuring the environment is a kind of cultural teaching. From cultural toys (small hunting bows), practice making animal tracks, and semi-wild pets, to accompanying adults on hunts and seeing examples of edible mushrooms, children’s environments are richly structured for cultural learning. Sterelny’s solution also involves network characteristics. For example, if communication is between one-to-many agents, the chances of successful deception are reduced. Information pooling increases overall accuracy. In complex worlds, an epistemic division is efficient. In these circumstances, the benefits of deception are reduced.

§4 Culturally Inherited Vigilance In the previous section, I argued that open vigilance can help structure the learning environment. In this section, I argue that the human structured learning environment is the site at which open vigilance is culturally inherited.

§5 Costly Signals Costly signaling theory can capture my unification of Sterelny and Mercier’s solutions. The idea is that it is relatively costly to send unreliable signals in human structured learning environments, and this explains the reliability of communication. Sterelny rejects signaling theory, but

he does so because he misses the distinction between absolute and relative costs (Penn and Szamado, 2020).

Conclusion I argued that humans structure their environments in complex and dynamic ways to solve the puzzle of communication. In my view, these are epistemic practices that result from cultural evolution.

---

**Thomas Mitchell** (University of Edinburgh): “Norms of trust”

The purpose of this paper is to give an overview of the norms of trust. What can rationally justify trusting someone to do something?

It can be tempting to think that trust cannot be rational, that it is not responsive to reasons in the way that beliefs are. If there is little evidence supporting one’s trust, then it is not rational. But if there is sufficient evidence, then one merely forms an ordinary justified belief and there is no place for trust. Epistemic reasons might therefore seem inappropriate for justifying trust.

In response, some argue that all and only the evidence should be taken into consideration (e.g. Simpson 2017), whereas others think that we should look also to non-evidential reasons to fulfil the justificatory role (e.g. Baker 1987; Holton 1994). I will argue that trust can be rationally justified by certain epistemic and pragmatic reasons. I take trust to be reliance on another to be trustworthy, which is a common theme in theories of trust – although there are substantial differences in views on what it means to be trustworthy.

Another crucial point is the distinction between the right and the wrong kinds of reasons for an attitude. This is often made in the context of belief. Being offered a reward for believing something is the wrong kind of reason for believing it, though it may motivate one to inculcate the belief in oneself. The right kind of reason for belief is evidence. Further, I take it that one indication of a reason being of the wrong kind is the fact that it is state-given, rather than object-given. In the case of being rewarded for a belief, the reason is concerned with the belief itself – the mental state – rather than about the object of the belief, whereas evidence counts in favour of the object – the proposition to be believed.

I will argue that, as a kind of reliance, the norms of reliance apply to trust. These are similar to the norms of betting – considerations that make an outcome more desirable or more likely count in favour

of relying on it, as well as considerations that mitigate the bad outcome. Thus, both epistemic and pragmatic considerations are included. However, since trust is reliance on trustworthiness, the evidence that counts in favour of it cannot just be evidence that they will perform, but that they will be trustworthy in their performance. Similarly, some pragmatic reasons for trust are of the wrong kind. In therapeutic trust, for instance, one trusts another in order to encourage greater trustworthiness in them. This is a state-given reason; it favours having an attitude of trust, rather than counting in favour of the trustee's actual trustworthiness. An example of an object-given pragmatic reason would be there being some advantage to someone's proving trustworthy when trusted.

I believe that this approach is flexible and can be used in conjunction with any plausible theory of trustworthiness. However, I will end by demonstrating how it can be applied to a commitments-based view.

---

**Christian O. Scholz** (Amsterdam University): "Time in Dreams"

Dreams are as integral to our nightly experience as sleep itself. While mysteries of dreaming have been discussed widely throughout the history of philosophy (and psychology), the investigation into the chronometric properties of dreams specifically is a comparatively recent topic of interest. Arguably beginning with Maury's guillotine dream (discussed among others by Freud and Kramer) and leading up to more recent empirical investigations conducted in sleep laboratories, both theories and experimental data on our experience of time in dreams are steadily accumulating. Especially in light of a current resurging popularity of dream research in cognitive science and philosophy, as well as the recognized potential for dreams as a model of (minimal) consciousness, reviewing and synthesizing the present data is more relevant than ever.

In this paper, I will conduct a philosophical analysis aimed at harmonizing the seemingly incompatible results derived from research on time in dreams. In the first part, I will review empirical data from research conducted on lucid dreaming, duration estimation and REM behavior disorder. In this part I will also distill three different and partially incompatible hypotheses, namely the emotionality hypothesis, the synchronicity hypothesis and the modality hypothesis, all aimed at capturing our experience of time in dreams.

In the second part, I will use Vogel's structure-flow model, in order to define the concepts of dream flow and dream structure. I will subsequently investigate these concepts and will derive logically motivated insights, such as the logical equivalence of time passage in dreams with dreaming itself, as well as the exclusively present-oriented structure of dreams. I will then return to the three previously distilled hypotheses and argue that a distinction between two selves, namely between the dreaming self and the dream self (a distinction inspired by Augustine's distinction between dreaming and waking self) can help us in harmonizing the seemingly incompatible hypotheses derived from the empirical literature. Specifically, I will define the dreaming self as the physical self in the state of dreaming and the dream self as the mental representation of the physical self (crucially not the mental self). I will then show how dreaming self and dream self stand in relation to the previously discussed modes of dreaming and the results on dream flow and dream structure. I will conclude that while it is certain that we experience time in dreams, how we experience it depends largely on our level of lucidity. In connection with this conclusion I will stress the importance of distinguishing between different modes of dreaming and the relevance of investigating the whole range between the two extreme poles of non-lucid and lucid dreaming.

---

**Angelos Sofocleous** (University of York): "The depressed individual as a 'non-participant spectator'"

In the field of the phenomenology of depression, the depressed individual has been described as being disconnected, isolated, incarcerated, detached and alienated from other people and the world. I argue that such descriptions of one's experience of depression can be examined through understanding the depressed individual as a 'non-participant spectator' in the world.

I begin by examining the structure of interpersonal relationships and suggest that an interpersonal relationship is characterized by the following constitutive aspects: i) Reciprocity, ii) Interaffectivity, and iii) A sense of the other as a person. An interpersonal relationship involves turn-taking and is established as a reciprocal interaction in which agents respond to the other's actions and behaviour while also experiencing themselves as subjects. Such interactions are also interaffective in nature as participants have the ability to 'affect and be affected'. An interpersonal relationship must also involve a sense of the other as a person - that is, as an individual who offers the possibility of engaging in contingent interactions, with whom one can interact in a reciprocal and interaffective manner. I further

suggest that interpersonal relationships of this kind are fundamental in providing us with a sense of belonging in a world in which we feel ‘at home’.

Focusing on experiences of depression, I examine how interpersonal relationships are disturbed in depression and how this subsequently affects one’s being-in-the-world. By focusing on the three constitutive aspects of an interpersonal relationship (i.e. reciprocity, interaffectivity, and a sense of the other as a person), I demonstrate how each of these is disturbed in depression, subsequently affecting the individual’s being-in-the-world. Due to these disturbances, depressed individuals describe themselves as being alienated, isolated, and incarcerated in the world, and also experience a diminished sense of being-with other people. Consequently, some describe that they feel as if they inhabit ‘another world’ due to their inability to establish meaningful interpersonal relationships.

The depressed individual’s experience of being alienated from the world, I argue, can be analyzed through the analogy of being a ‘non-participant spectator’ - that is, someone who feels unable to engage, intervene, or participate in the world, as she adopts a detached third-person perspective toward the world. The depressed individual is then not an engaged participant in the world but a non-participant spectator who witnesses other people, with whom she feels unable to meaningfully interact, getting on with everyday life.. The world, then, takes a form of being a world-for-others.

First-person testimonies from depressed individuals demonstrate that the depressed individual feels as if they are a detached observer of worldly events and other people and lacks the feeling that they can get actively involved in what they perceive. Therefore, as the depressed individual cannot establish interpersonal relationships which are reciprocal, interaffective and involve a sense of the other as a person, she is alienated and isolated from the world, becoming a non-participant spectator in it.

---

**Karl Landström** (Coventry University): “Research funding as domination: Epistemic Freedom and Resistance”

In March 2021, UK Research and Innovation (UKRI), a leading research funding body in the UK announced a substantial reduction in its international research development budget following the UK government’s cut to its overseas development aid budget (UKRI 2021). The £1.5 billion Global Challenges Research Fund (GCRF) was heavily affected and research projects which had been awarded funding but were yet to start were cancelled and an estimated 800 research projects that were already

undergoing had to face the consequences of the funding reduction in the form of either termination or reprofiling. GCRF funding has since been partially restored to some of those projects, but these decisions will undoubtedly have lasting effects on UK researchers, research institutions and their international partners (Imperiale & Phipps 2022).

The interference with projects who had already been awarded funding sparked plenty of debate pertaining to the ethics of research funding and academic freedom, a debate to which this paper aspires to contribute. In this paper I advance the view that it is not the interference itself that is the core issue, but rather the funder's ability to interfere. This view is inspired by neo-republican accounts of freedom as put forth by scholars such as Quentin Skinner (1998) and Philip Pettit (1997). I draw on the work of the decolonial theorist Sabelo Ndlovu-Gatsheni (2018) and the neo-republican epistemology of James Bohman (2012) to argue that certain forms of research funding constitute domination. That is, in such funding schemes the funder has the ability to arbitrarily interfere in the funded research projects, and arbitrarily assign duties and responsibilities to the recipients of funding without there being any meaningful possibility for contestation. Thus, such funding schemes limits the freedom of the recipients of funding from those schemes. In this paper I discuss the Global Challenges Research Fund as an example of such a research funding scheme. I argue for four theses: (I) Certain forms of research funding constitute domination. (II) Such domination risks producing significant negative ethical, epistemic and practical consequences, particularly when it reproduces colonial hierarchies. However, (III) despite the dominating position inhabited by the funder, there is still space and possibility for agency and resistance on the part of the dominated. Lastly, (IV) I argue that despite the agency and resistance on the part of the dominated, these funding schemes nonetheless constitute domination. I conclude by considering some aspects of what a non-dominating funding scheme would entail. Thus, this paper hopefully adds to the growing critical literature on the Global Challenges Research Fund, but also to wider discussions pertaining to the ethics of research funding, research policy and governance, and to wider debates pertaining to the decolonisation of knowledge and epistemic justice.

---

**Miguel Rudolf-Cibien** (Groningen University): “Economic Exploitation as Structural Domination”

What is economic exploitation? Against the grain of most contemporary theories of exploitation, which attempt to define the concept with regards to distributional criteria, this article argues for a relational

understanding of exploitation. Exploitation, it is shown, can be seen as the transfer of value that occurs on the basis of a relation of domination. It is the economic materialisation of domination. Such a view was chiefly defended by Nicholas Vrousalis (2013). However, this article rejects Vrousalis' conception of exploitation as a form of exclusively interpersonal domination. Instead, I propose that exploitation is best conceived as a systemic phenomenon. Exploitation, in my view, is a concept that applies to social relations if, and only if, they emerge from structurally constructed positions of power and of disempowerment and not from incidental positions of dominance. It is the evil at play in society producing its classes; the economic materialisation of structural domination.

That conception is superior in at least three ways. Firstly, it fares better at identifying cases of exploitation, whereas Vrousalis' theory generates false positives. Secondly, it is truer to the long-lived intuition in Marxist circles that exploitation is a feature of capitalism itself more than that of particular agents under capitalism. Thirdly, by operating a clear conceptual separation between systemic exploitation and comparable types of interpersonal wrongdoings it gives an explicit target to action: systemic change.

My conception of exploitation builds on the concept of structural domination in the labour market as it was recently formulated within a radical trend of republican philosophy: labour republicanism. (Cicerchia, 2019; Gourevitch, 2015) These views hold that all market-dependent workers, i.e. people who have no alternative to selling their labour for subsistence, are dominated even if they may not always be dominated by anyone specifically. This is because the structure of the labour market itself is a structure of domination. In my view, the argument that all trade of labour comes from a place of structural domination implies that all trade of labour is exploitative. Therefore, all relations of employment (of market-dependent workers) amount to relations of exploitation.

The article proceeds as follows: I firstly present Vrousalis' seminal dominationview as a welcome alternative to the mainstream of exploitation theory. Secondly, I reject Vrousalis' theory on the basis that it characterises exploitation in an unconvincingly broad manner, and I show that the defence he already iterated against similar criticisms is in fact tautological, therefore invalid. Then, I argue that the theory's central insight can be preserved and ameliorated by anchoring exploitation within a concept of structural, rather than interpersonal domination. Finally, I discuss what such an account would have to say about exploitation in the labour market by articulating it with the labour republican understanding of domination.

**Ervin Kondakciu** (Rijksuniversiteit Groningen & Universität Hamburg): “On Rawls’s model of disobedient politics: can there be disobedience under near just conditions”

Since the beginning of the previous decade disobedient politics in its various forms have, to this day, taken centre stage in the European and the North American world. As a result, many theorists began scrutinizing existing accounts of disobedient politics. John Rawls’s model of civil disobedience was soon criticized for being restrictive and impractical. Though valuable, these critiques expose the ‘external vulnerabilities’ of Rawls’s model, and, therefore, are faced with their own limitations. This paper advances an internal method for identifying the model’s shortcomings. The analysis unfolds at the level of political ontology. The paper aims to, first, show that when Rawls’s political ontology is properly scrutinized, internal incoherencies emerge, and, second, lay bare the sources of the identified incoherencies.

Section I introduces Rawls’s model of civil disobedience and scrutinizes his concept of a nearly just society. Rawls’s fleeting engagement with the nature of nearly just societies, which constitute the context for his model, calls for interpretative work. I maintain that nearly just societies are characterized by, first, a responsive and accountable representative political system elected through periodical elections, a codified system of inalienable rights, and the separation of the three branches of power; second, a basic institutional infrastructure organized through principles of justice; third, ‘constitutional patriotism’, the idea that social cohesion and solidarity are generated through discursively weaved identification with the society’s institutional achievements; and, fourth, social injustices, this is, violations of the principles of justice, may obtain as a result of illegitimate political decisions.

Section II introduces the conceptual tools of political ontology. Rawls’s political theory is a form of social critique with practical or ameliorative intent. Political ontology studies the background ontological assumptions or taken-for-granted concepts, ideas, representations, hypotheses and the like, through which political theorists understand the relevant social phenomena. Such background assumptions carry the potential of prioritizing or downplaying particular constellations of the social phenomena in question, and, thereby, enabling or foreclosing certain interpretations and/or problematizations.

Section III develops the ontological status of Rawls’s understanding of political personhood, as invested with the moral powers of reasonableness and rationality, and his deontological individualist conception of political society. It then shows that upon scrutiny his understandings lead to incoherent

results: either disobedience is warranted, but superfluous, or unwarranted, but vital – and thus, his model fails to serve his own social critical or ameliorative goals.

Section IV identifies the sources of the incoherence in: first, Rawls's exclusive focus on the prohibitive, negative aspects of power - and his associated disregard for its productive, positive side, which entails that he is incapable of capturing and problematizing pre-political systems of oppression without necessarily identifiable agents. Second, there is no conceptual space for a *political* understanding of politics in Rawls's model, this is, an understanding that departs from the ineradicability of social conflict, and apprehends politics' presence whenever the usual way of doing things is contested. Instead, in Rawls's model we find all political questions of significance settled, a consensus-driven functioning of the political system, and an impoverished political debate revolving around questions of technocratic management.

---

**Shervin Mirzaeighazi** (Manchester University): "Punishment and Revenge"

In this paper, I want to delve into the relation between (retributive) punishment and revenge to see if there is a fundamental difference between the two practices or, as some philosophers mentioned: punishment is a mask for our desire to take revenge—or it is just institutionalised way of taking revenge. Nietzsche famously believed that our belief in free will and therefore just punishment is a consequence of our desire to take revenge. As Moore mentioned, Nietzsche believes that retributive emotions are

truly a witches' brew : resentment, fear, anger, cowardice, hostility, aggression, cruelty, sadism, envy jealousy, guilty, self-loathing, hypocrisy, and self-deception—those 'reactive affects' that Nietzsche sometimes lumped under the French term *ressentiment*." (Moore, 1997, p. 120).

On the other hand, Kant, who were opposed to private revenge, claimed that

Even if a civil society were to dissolve itself by common agreement of all its members ..., the last murderer remaining in prison must be executed, so that everyone will duly receive what his actions are worth and so that the bloodguilt thereof will not be fixed on the people because they failed to insist on carrying out the punishment; for if they fail to do so, they may be regarded as accomplices in this public violation of legal justice. (1996, 158)

As it is clear by these quotes, while some philosophers believe that there is a real difference between revenge and (retributive) punishment, others reject such a difference. The importance of the issue lies in its implications for the debates of the justifiability of punishment. Intuitively we believe that taking revenge is wrong and should be condemned. If our practice of punishment is just another way of taking revenge, then punishment is also not justified and should be condemned. The argument would be like this:

1. Revenge is not justified
2. Punishment is a form of revenge, Therefore

Punishment is not justified

Proponents of (retributive) punishment usually take two different ways to undermine the conclusion. They either reject the first premise and claim that revenge can be justified in certain circumstances, and for this reason, punishment also can be justified in some cases. Or, they reject the second premise and claim that there are important differences between the two practices and, therefore, what is true about revenge is not necessarily true about punishment. In this paper, I will try to use a combination of these strategies to show that (retributive) punishment can be justified. The first step of my argument is to claim that there are critical differences between punishment and revenge, and this gives us a reason to reject the argument above. In the second step, I will try to show that even revenge is not always wrong or unjustified—to the extent that it resembles punishment, it can be justified. The belief that revenge is always wrong seems to be more of a Christian doctrine than a universal intuition. I further discuss—briefly—the implications of my thesis on our practices of punishment and holding others morally responsible.

---

**Nadia ben Hassine** (University of Cambridge): “Empathy and Impartiality: On the Limits of Perspective-taking”

Political thinkers often want their theories to be applicable to societies whose members hold a wide range of perspectives. In answering questions such as: “What is the best form of political organisation?” or “How should goods be distributed?”, the political theorist has to give answers which can reasonably be accepted by a wide range of positions. Due to their wide societal application, then, political theories have to be compatible with a plurality of viewpoints. This respect for plurality is what

motivates the widely endorsed aim for impartiality in political theory (Nagel 1995). The ideal of impartiality is often defined negatively: being impartial requires you to *not* privilege one perspective over another without sufficient grounds for doing so. Although impartiality is a widely endorsed ideal in political theory, it remains underdiscussed how impartiality, understood as the not privileging of one perspective over others, can be achieved. As I will argue, it follows from this understanding of impartiality that it requires gaining knowledge of the perspectives of others. The process of gaining knowledge of other perspectives bears a close resemblance to empathy as it is understood in ethics. In ethics, empathy is often discussed as an epistemic tool as it is one of the primary ways in which we gain insight into other perspectives. Empathy as a way of gaining knowledge of others' perspectives is therefore an important, underdiscussed topic for political theory. Although the role of empathy in political theory has been acknowledged, a lot more needs to be done to review the role of empathy in political theory (Okin 1989; Barry 1995). This paper focuses on the way in which impartiality relies on empathy and discusses the impact of empathy's limits on this aim for impartiality.

Increasingly, critical feminist theorists as well as philosophers of race have pointed out traditional political philosophy's inability to incorporate the perspectives of certain social groups (Anderson 2009; Harding 2009; Mills 2005). For this reason, it is important to look at what impartiality requires, and, whether the current tools used in political philosophy are sufficient. I argue that political theory's aim for impartiality, when reviewed through its reliance on the tricky process of empathy, faces serious problems. In this paper I focus on the epistemic limitations we face in perspective-shifting attempts. These reflections will show that empathy is an inadequate method for gaining full knowledge of the perspectives of others. I both aim to show that (1) empathy is a necessary starting point for impartiality, and that despite this, (2) we need to supplement this method if we are aiming for impartiality. Concluding this paper is the insight that empathy cannot fully overcome the distortions generated by our own perspectives, which is why political theory needs to be in active dialogue with a variety of underrepresented voices.

---

**Jonathan Seim** (University of Bochum): “Two conceptualisations of political responsiveness. An intervention”

The concept of political responsiveness is frequently assumed to be of central importance to the assessment of the quality of representative democracy. It is surprising how rarely an accurate definition is made explicit, as if the meaning of the concept is readily apparent. On the contrary, it must be noted that the conceptualisation differs from discourse to discourse. The contribution on hand identifies two distinct conceptualisations of the concept of political responsiveness (a), works out their conceptual differences (b) and shows that the confusion of both can lead to misjudgements in the normative analysis of representative democracy (c).

(a) The first conceptualisation of *responsiveness as congruence* refers to the congruence between the preferences of citizens on the one hand and the decisions of their representatives on the other hand. The more the decisions correspond to the preferences of the citizens, the more responsive the political system is. The second conceptualisation of *responsiveness as intentional orientation* refers to the orientation of elected representatives towards the citizens' preferences when making a decision. The more the representative takes the citizens' preferences into account and follows their and not her own preferences when making a decision, the more responsive the representative is.

(b) Although, both conceptualisations are oftentimes empirically linked insofar as the intentional orientation of the representatives towards the preferences of the citizens provides a plausible causal mechanism that may lead to a congruence between the decisions of the representatives and the preferences of the citizens, neither is logically or conceptually contained within or can be derived from the other. In fact, the mentioned causal mechanism is even empirically neither necessary nor sufficient for the mentioned congruence. Firstly, a representative may orient her decision-making towards the preferences of the citizens without this leading to congruence because her assessment of the preferences may be mistaken. Secondly, a representative's decisions may be congruent with the preferences of the citizens without her decision-making being intentional oriented towards the citizens' preferences. This is for example the case, when the preferences of the representative and of the citizens do coincide, so that the representative automatically promotes the citizens preferences by following her own preferences.

(c) Three problems do arise: Firstly, labelling two quite different but related phenomena with the same term carries the danger of miscommunication. This identical labelling may, secondly, also lead to the assumption, that only responsive representatives can contribute to the responsiveness of the political system. Insofar as other causal mechanisms of ensuring responsiveness of the political system are conceivable, this assumption leads to misjudgements concerning the quality of representation and to

limitations concerning the design of democratic innovations which could improve the responsiveness of the political system. Thirdly, the same applies to the narrow focus of both conceptualisations on the correspondence with the citizens' preferences, which prevents the analysis of further possibilities to take the citizens' preferences into consideration.

---

**Hrayr Manukyan** (Amsterdam University): “Enacting the new social contract: civil disobedience in Armenia”

In this paper I claim that the notion of constructive disobedience provides a better framework to explain the Armenian Velvet Revolution of 2018 than the four main theoretical models of civil disobedience (religious-spiritual, liberal, democratic, anarchistic).

In the first part I describe the political systems of Armenia before and after the Revolution, and I show that there is substantial institutional change as well as change in the sources of legitimation of the government. The institutional change concerns elections. As for the primary source of the legitimacy of the Armenian government, in the pre-revolutionary period it was the Nagorno Karabakh conflict, while after the Revolution, some democratic and social issues took priority over this conflict.

In the second part I argue that the Revolution should be considered as a case of civil disobedience. Then I discuss the four main theoretical models of civil disobedience in details and show why they cannot adequately explain the Armenian case. The religious-spiritual model (Mahatma Gandhi and Martin Luther King) cannot explain the Armenian case because religiousness and the concept of god play a crucial role in it. The liberal model (John Rawls) cannot explain the Armenian case because it considers civil disobedience as a tool for non-radical transformations in newly just societies. The democratic model (Hannah Arendt and Jürgen Habermas) cannot fully explain the Armenian case because it sees civil disobedience only as a “democracy-enhancing” mechanism, not as “democracy establishing”. The anarchist model cannot explain the Armenian case because it rejects the notion of fidelity to the rule of law as a core principle of civil disobedience.

In the third part I present the idea of constructive civil disobedience, and I show that it is a better framework to explain the Armenian Revolution. The idea of constructive disobedience was suggested by Leslie Green to describe the anti-globalization protests of the 1990s and early 2000s. According to

him, these protests should be considered as civil disobedience that aimed to establish the social contract and regulations in such areas or places where the situation is equivalent to the state of nature. The situation in Armenia before the Revolution cannot be considered as an equivalent to the state of nature because independent Armenia was always a state with a functioning government, including a functioning police and army. Therefore, constructive disobedience as presented by Green cannot fully explain the Armenian case. However, constructive disobedience can also be understood in a broader sense to include the cases when the new social contract is attempted to replace the old social contract. Constructive disobedience perceived in this broader sense can also apply to the Armenian case. The institutional change and the change in the primary sources of government's legitimacy have been so radical in Armenia that they can be adequately understood only in terms of the old and the new social contracts. This notion of relation between civil disobedience and the new social contract is precisely the reason why constructive disobedience is the best explanation for the Armenian Revolution.

---

**Alexander Murphy** (University College London): “Why digital objects won't help virtual realists”

This paper discusses the (non-)reality of virtual objects – the objects we perceive, and interact with, in virtual environments. I object to virtual digitalism, the view that each virtual object is a digital object. Virtual digitalism is a form of virtual realism, the view that virtual objects are real. Identifying virtual objects with digital objects is meant to explain their reality, thus supporting virtual realism in the face of virtual fictionalism, the view that virtual objects are fictional.

I argue that virtual digitalists cannot individuate digital objects in a way that makes them plausibly identifiable with virtual objects.

I consider various proposals for how virtual digitalists could individuate digital objects. Many of these are inspired by David Chalmers' virtual digitalism. Chalmers tries to individuate digital objects by their grounding bases. These grounding bases are the physical manifestation of data structures containing data relevant to the objects in question. I argue that Chalmer's approach fails for two reasons that generalise.

Reason 1: Given current memory systems, virtual digitalists cannot individuate digital objects with acceptable modal profiles. They face a dilemma regarding cases involving multiple playthroughs.

Consider scenarios in which a virtual environment is reloaded to how it was at a certain moment (like in videogames). With regards to arbitrary virtual object P, virtual digitalists can say either that...

A) ...there are many distinct Ps, one for each playthrough.

OR

B) ...there remains only one P throughout the playthroughs.

Choosing A leaves the virtual digitalist with a fission problem. It's implausible that P (singular object) multiplied because one didn't complete the level in one playthrough.

Choosing B requires explaining how different data, from different playthroughs, can be grouped together into the grounding base of P even though data for each playthrough is wiped and overwritten.

Reason 2: computational data is too integrated for virtual digitalists to individuate digital objects using data structures.

Optimisation compilers use algorithms to fold information into a smaller number of data structures whilst keeping the resulting computational system semantically equivalent to the pre-optimised version. It's therefore possible that the same virtual objects can be generated whilst the number of data structures changes.

This points to a problem for virtual digitalism. Data structures don't track virtual objects as we experience them. Data structures are formed to increase data management efficiency, not to track objects.

The point isn't that virtual digitalists require a 1-1 mapping between data structures and virtual objects. Rather, it's that there is no natural object-like grouping of data. Consequently, virtual digitalists cannot successfully individuate virtual objects, or even uphold an object/law distinction for virtual environments, using data structures.

Virtual digitalists are unable to individuate digital objects that can plausibly be identified to virtual objects. As such, virtual digitalism doesn't aid the virtual realist project.

That said, my argument isn't directed against virtual realism in general, but specifically against virtual digitalism. I remain silent on the prospect of non-digitalist virtual realism. I note, however, that unless such an account presents itself, virtual fictionalism looks promising.

---

**Eve Judah** (École Normal Supérieure): "From Outside the Estate: The Symbolic Language of Philosophical Canon-Formation."

The aim of this paper is to examine the process of philosophical canon-formation at the level of its symbolism: its 'foundational metaphors'. My core contention is that the perceived unity or integrity of the philosophical field is maintained by genealogical, Eurocentric metaphors of succession, paternity and inheritance that require authors from beyond the pale find different symbolic languages with which to express themselves, and that, moreover, this has a crucial impact on conceptions of the 'foundation'. We can get a sense of the emergence of that so-called 'different language' in the work of three very different post-structuralists, Sarah Kofman, Kamau Brathwaite and Jacques Derrida, by examining their differing uses of spatial, genealogical imagery as it expresses their belonging, or rather non-belonging, to the European tradition.

This paper will delve into a comparative reading of the territorial and oceanic metaphors which make up the common language of Kofman, Brathwaite and Derrida as they navigate new directions in poststructuralist ontology. In doing so, it will engage in a discussion of the role which metaphor and the forms of poetry can play in the development of poststructuralist approaches to processes of becoming. In point of fact, while Jacques Derrida characterises the inheritance of substance ontology as a 'private estate', and himself as a trespasser on the land, his Parisian contemporary Sarah Kofman deploys classical oceanic imagery to express her sense of having been cut adrift from this patrilineal tradition, figuring herself instead as an orphaned and unwilling oceanic mother, a kind of watery 'womb-tomb'. Meanwhile, Kamau Brathwaite stages his symbolic escape from the private estate, which in a colonial context morphs into a paternalistic plantation system, as a flight forth into the poetic region of the Caribbean Sea, which, in contrast to Kofman's Aegean, stands for liberation, self-determination and recovered ancestral inheritance. In a conjoined attempt to move away from the paradigm of philosophical paternity and filial succession and its qualitative impact on post-structuralist critiques of substance ontology, Brathwaite and Kofman make metaphor the ground (or waters) of their ontological analyses to pursue an understanding of the 'foundation' which escapes rational schematisation and traditional conceptualisations. Moreover, this paper will open onto a discussion of post-structuralist post-foundationalism through an analysis of Judith Butler's notion of 'contingent foundations'. Butler's model is adapted to our ends precisely in that she enacts a shift towards a processual, relational and genealogical perspective, by positing the idea that foundations attain quasi-transcendental status precisely because they are constantly being subjected to constestation, succession and change. Finally, we hope that our examination of canon-formation from the perspective of authors who have traditionally been excluded from that canon, notably Brathwaite and Kofman, will

demonstrate the importance of listening to minority post-colonial and feminist voices for the sake of our joint philosophical futures.

---

**Santiago de Arteaga Gallinal** (Pontificia Universidad Católica de Chile): “Søren Kierkegaard and Carl Jung on freedom to become oneself”

Søren Kierkegaard and Carl Gustav Jung have in common the idea that the *self* must be posited, it is a reality that must be constituted. The human being must become himself from a concrete ontological/psychological structure that presupposes precisely the need to realize himself as constitutive-of-himself, that is, the *existential* character of the self. The self is the telos of human existence. For Kierkegaard, it means a synthesis of various elements that have been pre-established and that only obtain unity when they are consciously related to each other: the self, says the Dane, is the relationship that is related to itself. For his part, Jung maintains that it is necessary to integrate the various elements of the psyche, conscious and unconscious, in the unity that signifies the totality of the personality. However, it is not clear what place freedom occupies in this constitution/integration.

What is the conception, broadly speaking, of Kierkegaard and Jung about freedom and the meaning it has for the becoming of the self? Is there freedom in this constitution of the self? Could it be said that Jung and Kierkegaard agree that freedom is rather paradoxical, that is, it has both a sacrificial and an active dimension? In other words, does freedom have a moment of surrender-of-itself for the constitution of the self?

Jung states, on the one hand, that there must be a denial of freedom at the level of the ego (consciousness), while, on the other hand, he maintains that this means a broader freedom of the self. It is only when certain aspects of the personality are sacrificed that the larger personality can come into being; now, this necessarily requires a relationship with the archetype he calls the Self, and a (partial) sacrifice of ego to its deeper claims. For Jung, the freedom of the ego is sacrificed to the needs of the Self, from whence results the broader freedom of the personality. According to Kierkegaard, freedom is a necessary condition for the constitution of the self, since only through freedom can the synthesis be achieved in which the relationship that is the self is related to itself. Existence is a gift and a task in such a way that the task is contained in the very structure of the gift but must be carried out personally by the choice and the conscious relating of the relationship to itself. Now, this requires understanding the radical dependence of freedom on its foundation. So, it seems we have two interpretations.

Kierkegaard maintains that it is necessary to recognize the dependency with respect to the ground so that the self can become what it is. Jung argues that only the sacrifice of the ego to the necessary claims of the Self can result in true, broader freedom. For both, the resulting freedom is the self.

Kierkegaard and Jung are two thinkers who have rarely been put into dialogue, there are few texts that directly address the relationship between them; so, my intention is to contribute to this approach from the point of view of freedom in becoming oneself. This presentation will offer the general lines of the subject, which is part of my doctoral research and, therefore, is in process.

---

**Lucas Gronouwe** (Radboud University): “Doing justice to poetry: Gadamer and Derrida on reading Paul Celan”

What does it mean to do justice to a text? To a philosophical text, but also to a literary or poetic text, i.e. to poetry? What, in other words, is the task of philosophical hermeneutics, the discipline traditionally concerned with explanation and interpretation? This is the question that is at stake in the Gadamer-Derrida encounter: an intellectual debate that retains its relevance, not only because it has ‘animated contemporary philosophy’,<sup>1</sup> but also because the practice of reading and interpreting texts makes up a large part of philosophy as such. By revisiting the diverging hermeneutical strategies of Gadamer and Derrida, this paper seeks to determine how we can engage in this practice in a righteous or just manner.

Hans-Georg Gadamer and Jacques Derrida first met each other in April 1981, at a colloquium at the Goethe Institute in Paris. Their exchanges at this colloquium are collected in the edited volume *Dialogue & Deconstruction*,<sup>2</sup> and have been extensively commented upon.<sup>3</sup> Rather than adding another commentary to this already extensive literature, this paper aims to shed a new light on the Gadamer-Derrida debate by taking its starting point in the shared admiration of both philosophers for the twentieth century German poet Paul Celan. Since the debate between Gadamer and Derrida can be said to have continued throughout their respective readings of Celan’s poetry,<sup>4</sup> one can use one of these readings each as a way of entry into the differences, as well as similarities, between Gadamer’s and Derrida’s take on hermeneutics. In this endeavor, I focus on two texts that are closely related to each other. First, a book-length essay by Gadamer from 1973 called *Who Am I and Who Are You?*, which is a running commentary on Celan’s poetry collection “Atemkristall”,<sup>5</sup> and second, a lecture

presented by Derrida at a memorial conference in Heidelberg in 2003 in remembrance of Gadamer.<sup>6</sup> This text by Derrida is particularly relevant in this regard, because it contains a reading of several poems by Celan, which takes into account Gadamer's own comments on Celan in the aforementioned essay. In order to determine what doing justice to poetry means for Gadamer and for Derrida, I will discuss a number of issues, starting with Gadamer's and Derrida's shared rejection of the intentions of the poet as the decisive factor in interpreting poetry. This is then followed by a discussion of Gadamer's hermeneutical approach, as exemplified by his interpretation of Celan, in direct relation to Derrida's main objections to this approach. After subsequently having discussed the way in which Derrida demarcates his own hermeneutics from that of Gadamer, I first conclude that Gadamer's and Derrida's positions are sufficiently refined to be considered as complementary rather than mutually exclusive, and then argue that doing justice to a text means seeking to decipher its meaning, as Gadamer argues, whilst accepting that no articulated meaning can ever be final, certain or exhaustive, as Derrida emphasizes.