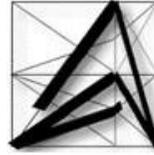




University of Maribor

Faculty of Arts



**Department of Philosophy, Faculty of Arts, University of Maribor and  
Slovenian Society for Analytic Philosophy and Philosophy of Science**

**Program/Programme**

**MEDNARODNI SIMPOZIJ:  
ŠTUDENTSKI FILOZOFSKI SIMPOZIJ 2021**

**INTERNATIONAL SYMPOSIUM:  
PHILOSOPHY STUDENTS' SYMPOSIUM 2021**

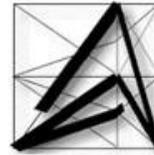
[Day 1 Classroom Link](#)

[Day 2 Classroom Link](#)

**MS Teams, 16.–17. November 2021**



Faculty of Arts



## Program/Programme

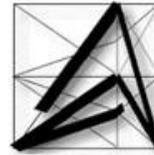
### Torek/Tuesday

#### MS Teams

- 09:00 – 09:15 Pozdravni govori/Welcome speeches
- 09:15 – 09:40 Jonah Schumacher, University of Illinois at Urbana-Champaign: Was Plato a Virtue Ethicist? Modern Deontic Reasoning in Plato's Republic
- 09:40 – 10:05 Mary Peterson, University of Edinburgh: Drawing the Line Between Contractualism and Particularism
- 10:05 – 10:30 Benedikt Namdar, University of Graz: Individuals, Existential Risk Prevention, and Moral Theories
- 10:30 – 10:45 Odmor/Coffee Break
- 10:45 – 11:10 Jennifer Bass, Pacific University: Individual and Institutional Change: How to Maximize Advocacy Efforts
- 11:10 – 11:35 Alina Ahmed, University of Georgia: Voluntourism and Epistemic Injustice: How Educational Institutions Benefit Credibility Excess
- 11:35 – 13:00 Kosilo/Lunch break
- 13:00 – 13:25 Dominika Palka, University of Warsaw and University of Maribor: Reasons-Responsiveness: One More Look at the Agent-Based and Mechanism-Based Approaches
- 13:25 – 13:50 Shonkholen Mate, Indian Institute of Technology: Understanding Scientific Normativity as Social Convention
- 13:50 – 14:15 Giuliano Rosella, University of Turin: Expanding Causal Modelling Semantics
- 14:15 – 14:30 Odmor/Coffee Break
- 14:30 – 14:55 Oushinar Nath, University College London: Wisdom, Action, and Knowledge
- 14:55 – 15:20 Niko Šetar, University of Maribor: Virtue Responsibilism and a New Kind of Responsibility



Faculty of Arts



## Sreda /Wednesday

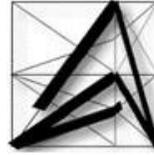
### MS Teams

- 09:15 – 09:40 Edoardo Vaccargiu, University of Genoa: Is Pragmatics a Mental Module?
- 09:40 – 10:05 Daniele Conti, University of Manchester: Probabilities of Actions and the Kinematics of Motivation
- 10:05 – 10:30 Jace Snodgrass, University of St Andrews: The Leibnizian Criterion and Property Individuation
- 10:30 – 10:45 Odmor/Coffee Break
- 10:45 – 11:10 Tadej Todorovič, University of Maribor: Arguments Against AI Art and How to Answer Them
- 11:10 – 11:35 Jaroslav Malík, University of Hradec Králové: Superintelligence and the Control Problem: Real Problem or Pseudo-Problem?
- 11:35 – 13:00 Kosilo/Lunch break
- 13:00 – 13:25 Marina Bajić, University of Maribor: Religious “Knowledge” as a Post-Truth Concept?
- 13:25 – 13:50 Teemu Tauriainen, University of Jyväskylä: Simultaneous Realism and Anti-realism, and Transcendence and Immanence of Truth
- 13:50 – 14:15 Sofia Melendez Gutierrez, University of Cambridge: Higher-Order Fictionalism and the Infinite Regress
- 14:15 – 14:30 Odmor/Coffee Break
- 14:30 – 14:55 Pak-Him Lai, Texas A& M University: A Proposed Solution to the Preface Paradox
- 14:55 – 15:20 Lucija Duda, University of Manchester: Material-Origin Essentialism or Necessity of Sex; But What About Gender?
- 15:20 – Zaključek/Conclusion



University of Maribor

Faculty of Arts



## Povzetki/Abstracts

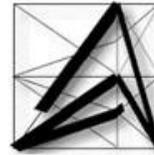
### **Was Plato a Virtue Ethicist? Modern Deontic Reasoning in Plato's *Republic***

*Jonah Schumacher, University of Illinois at Urbana-Champaign*

In the anglophone world of philosophy, Elizabeth Anscombe's *Modern Moral Philosophy* marks the birth of a line of reasoning which takes philosophers from classical Greek antiquity, especially Aristotle, to be engaged in a form of ethical reasoning that is fundamentally different in kind from that at work in later thinkers, such as John Stuart Mill or Immanuel Kant. She has not been alone in her view. Nearly the entire domain of contemporary ethical reflection takes her distinction between modern moral reasoning and ancient Greek ethical reasoning to be, in some sense, accurate. In the face of this consensus, this paper forwards the view that the distinction is unfounded, that deontic forms of ethical reasoning, functioning in a primary and not derivative capacity, can be found among ancient Greek philosophers, and that the view was not obscure. One finds evidence of it, I argue, in the ending myth of Er of Plato's *Republic*. I make the case that the Platonic myth of Er shares the three main characteristics of Immanuel Kant's Categorical Imperative Procedure. These consist of an aspect of universalizability, an adjusted social world, and an overall test of consistency instead of the maximization of an end. I begin by clarifying just what is at stake in deontological modes of reasoning, especially modern deontological modes, before turning to an analysis of Plato's myth of Er. I conclude by suggesting two ramifications of this thesis: one metaethical and one historical. Metaethically, the conclusion suggests that a single mode of ethical reasoning does not necessarily follow from holding a single metaphysics (and vice versa). Concerning the historical narrative, continuing to hold Kantian deontology as the cornerstone of modern philosophy could prove problematic, as the distinction between modern and ancient ethical reasoning is now less clear.



Faculty of Arts



## Drawing the Line Between Contractualism and Particularism

Mary Peterson, University of Edinburgh

This essay argues that Jonathan Dancy's particularism provides an epistemological but not a metaphysical critique of T.M. Scanlon's contractualism, and for its one-dimensionality, the critique fails. First, I detail Dancy's distinction between moral metaphysics and moral epistemology, showing that his attack on contractualism in *Ethics Without Principles* (EWP) touches the latter, not the former category. Next, I argue that Dancy misunderstands the context of justification that Scanlon develops in *What We Owe to Each Other* (WWO). Finally, I sketch a metaphysical critique of contractualism from the particularist position.

In an interview with Andreas Lind and Johan Brännmark titled "Particularism in Question," Jonathan Dancy distinguishes between moral epistemology and moral metaphysics. The context for the distinction is a description of particularism, Dancy's espoused position. As Lind points out, in the 1993 book *Moral Reasons*, Dancy considered particularism a position in moral epistemology. Later, working on *Ethics Without Principles* led the author to believe that particularism is primarily concerned with metaphysics—the properties of right and wrong actions—and only secondarily concerned with giving justification for morality. Of the distinction between metaphysics and epistemology, Dancy says:

I now take the view that there is a distinction between being a reason why one shouldn't do something and being a reason for believing that one ought not to do it. The latter topic is investigated in moral epistemology. "How do you know that it is wrong?" "What reason do you have to believe that it is wrong?" The former investigation is moral theory proper, where you investigate what makes an act right or wrong, which isn't an epistemic matter at all.<sup>1</sup>

Moral theory investigates reasons to do and not to do actions, without the further investigation into justifying beliefs about those reasons. At the time of the interview, Dancy considered moral metaphysics 'moral theory proper.'

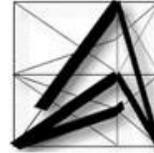
This paper takes Dancy's evolved description of his own work and applies the distinction between moral metaphysics and moral epistemology to the critique of T.M. Scanlon's contractualism in EWP. The distinction serves to illuminate Dancy's methodology, in that he takes metaphysics to be primary and epistemology secondary to a moral investigation. By Dancy's own standards, critiques of competing positions in moral theory ought to engage the metaphysical positions of the competitor, not merely what he would consider the secondary concerns of the theory. In attacking Scanlon, Dancy should respond to reasons not justification.

---

<sup>1</sup> Brännmark, Dancy and Lind "Particularism in Question: an Interview with Jonathan Dancy" *Theoria* Vol. 74, Issue 1, 2008.



Faculty of Arts



## Individuals, Existential Risk Prevention, and Moral Theories

*Benedikt Namdar, University of Graz*

A realization of an existential risk causes much harm. That harm includes the lives ended by such an event. But it also includes the potential value that would come into existence in the absence of an existential catastrophe. Combining these considerations with a significant overall possibility of existential catastrophes occurring leads to the conclusion that preventing such scenarios should be a major project of humanity.

The philosophical discussion surrounding existential risk prevention is mostly about incorporating it into policy making. However, what has yet been ignored is the role of individuals in the project of existential risk prevention. This presentation contributes to that part of the discussion. The first step of this talk is to motivate the claim that individuals should consider existential risk prevention in their moral deliberations as among possible consequences of individual acts are occurrences of existential catastrophes. For example, a single car ride could be the reason GHG threshold is exceeded, resulting in a climate catastrophe. Even given that such results are unlikely, it is important to consider such consequences, because the amount of value at stake is huge.

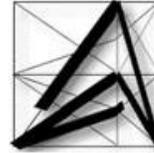
The second step of this talk is to show the complexity of incorporating existential risk prevention into moral theories. In Section II, I argue that simple consequentialism leads to overdemanding results in the context of existential risk prevention. After all, given the amount of value at stake in existential risk prevention, a consequentialist calculus will assess acts other than the one relevant for existential risk prevention as wrong in more cases than acceptable. In Section III, I discuss the question whether modified consequentialism informed by Scheffler's agent-centred prerogative can deal with existential risk prevention. I argue that the problem of defining the extra weight the agent-centred prerogative allows an agent to assign to personal projects is striking in the context of existential risk prevention. An agent-centred prerogative allowing to assign little extra weight will result in personal projects being outweighed by the amount of value at stake in existential risk prevention. And an agent-centred prerogative allowing agents to assign sufficiently large extra weight to personal projects to outweigh existential risk prevention will result in egoism.

In Section IV, I investigate if deontology can do a better job than consequentialism to deal with existential risk prevention. I argue that, firstly, well established problems of deontological theories are pressing in the context of existential risk prevention. Secondly, the intuition-based nature of deontological theories creates a problem. Intuitions are likely not apt to deal with existential risk prevention. The evolutionary nature of intuitions prevents these from being able to deal with issues that involve (temporally and spatially) distant people. Moreover, the amount of value at stake is likely to not be grasped by our intuitions.

In the last Section, I give action-guiding advice for individuals on how to pursue effective existential risk prevention. I argue that they should spend their resources on getting large-scale agents to do their job. And, lastly, I argue that the focus in the debate should still lie on institutional action as these are in the best position to pursue efficient existential risk prevention.



Faculty of Arts



## **Individual and Institutional Change: How to Maximize Advocacy Efforts**

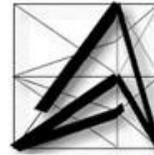
*Jennifer Bass, Pacific University*

As the climate crisis increases in severity, activists have started to shift their focus towards institutional or systemic levels of action rather than individual actions. In this paper, I consider how Henry, a young person with significant privilege, should address the climate crisis. To maximize his efforts, Henry hopes to consider effective altruism and a counter-argument known as the institutional critique. I argue that we can find common ground between the institutional critique and effective altruism, and there are strong moral reasons for individuals to spend their time and resources on efforts that promote institutional reform rather than directly to individuals in some scenarios.

Through thought experiments outlined by Brian Berkey, I suggest that individuals like Henry ought to partially dedicate their time and resources to individual actions and partially to systemic change. Individual-level actions include reducing personal plastic waste, while institutional-level activities include advocating for a new policy such as a plastic bag ban. Such a division of resources will result in the optimal course of action for future generations, even if it will result in less than optimal outcomes in each respective category as individual-level efforts will address the symptoms of issues and help facilitate a paradigm shift, while concurrent institutional level efforts will address the root causes of issues. I conclude that it is essential that all individuals partake in action in varying degrees, and like Henry, work to address the climate crisis by considering not only what outcome might maximize the immediate good, but also how their individual efforts might facilitate a cultural shift in environmental values and institutional change.



Faculty of Arts



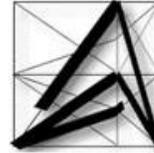
## **Voluntourism and Epistemic Injustice: How Educational Institutions Benefit Credibility Excess**

*Alina Ahmed, University of Georgia*

In this paper, I problematize the perceived expertise of western students who engage in voluntourism. I argue that they become epistemic “experts” by virtue of credibility excess, despite having little technical or academic training in the fields in which they volunteer. I aim to draw links between Wael Hallaq’s (2018) construction of Orientalism as a system of knowledge that is inextricably linked to colonialism, and modern day voluntourism that creates epistemic “experts” of white western college students who travel to countries across the world in order to ‘help’ communities of colour. I argue that these young western students become epistemic experts by virtue of little to no technical or academic training in the fields in which they are going to volunteer. This, I posit, is because of colonial and Eurocentric/Western systems of education which are, as a result of Orientalist conceptions of systems of knowledge, automatically seen as superior to the technical and indigenous knowledge systems of the rest of the world. I employ Miranda Fricker’s (2007) theorization of epistemic injustice to argue that this injustice is identity based and José Medina’s (2011) conceptualization of the importance of credibility excess to argue that voluntourism relies on both credibility deficit and excess. Lastly, I analyse educational institutions’ decisions to consider these experiences valuable as a method of gatekeeping education to those that can afford voluntourism. I argue that colleges’ and universities’ decision to consider these experiences as having any form of value is a method of gatekeeping education to those that can afford these excursions, and therefore necessitating credibility excess as a condition for educational and professional growth.



Faculty of Arts



## Reasons-Responsiveness: One More Look at the Agent-Based and Mechanism-Based Approaches

*Dominika Palka, University of Warsaw and University of Maribor*

According to Fischer and Ravizza, having alternative possibilities of action is not a necessary condition of moral responsibility; they are convinced by the Frankfurt example and motivated to hold this position due to the unresolved issue of compatibility of determinism and possibility to act otherwise. Instead, they propose that what we need for holding people morally responsible is that they possess a kind of control called "guidance control"; and this is dependent on the mechanism which leads to action's being moderately reason-responsive. The notion of moderate reasons-responsiveness (MRR) is explicated as follows: a type of mechanism *M* is MRR if and only if (1) in many possible worlds where it operates and *S* has a sufficient reason to do otherwise, *S* recognizes that reason as sufficient; (2) *S*'s disposition to recognize different reasons forms an intelligible pattern; (3) in at least one of those worlds, *S* acts on the recognized sufficient reason. After the authors presented their account in their book *Responsibility and Control: A Theory of Moral Responsibility* (1998), a lot of discussion and criticism emerged. One of the main points of criticism was the vagueness of the notion of mechanism and the lack of identity criteria for different types of mechanisms. This led some philosophers, Carl Ginet and Michael McKenna among them, to adopt an agent-based reason-responsiveness theory. This modification faces another problem: it has been suggested that it is vulnerable to Frankfurt-like examples, since the agent with whom the counterfactual intervener is connected is not reason-responsive, although he is morally responsible. This charge is rejected by both Ginet and McKenna; however, the latter admits that there are possible examples with some kind of "intervener" who is internal to the agent that may threaten the agent-based theory. One of his undeveloped proposals for such an example is the situation in which the mentally ill agent acts while his sickness is not active but it would activate and prevent him from acting otherwise in alternative scenarios. In my presentation, I will develop McKenna's type of example and contrast the ways in which the mechanism-based approach and the agent-based approach could deal with them. I will also present a possible counterexample of my own, concerning a recovered addict who refuses to take a drug, seems to be morally responsible (in this case - praiseworthy) for his resistance, and yet is not reason-responsive in the right way.



Faculty of Arts



## ***Understanding Scientific Normativity as Social Convention***

Shonkholen Mate, Indian Institute of Technology

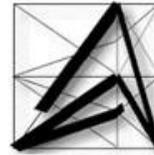
Philosophers of science have long been concerned with the foundational questions of scientific normativity such as the categorical versus hypothetical, or universalist versus pluralistic understanding of scientific normativity. This paper investigates the social foundation of scientific *normativity* to arrive at the conventional conception of normative perspective that governs our methodological choice in science. We build upon Lewis' game theoretic account of convention, which has been seen as providing the basis for the analysis of (social) norms. By stressing the implicit social dimension of Lewis' convention, we attempt to evaluate the social conventional account of normativity. We propose that scientific norms that govern methodological choice in science are conventions that arise out of human social interaction, values, and goals. This will enable us to assess how Lewis' analysis of convention can help us understand norms governing the methodological choice in science as conventions and how they emerge.

The paper is divided into two parts. The first part deals with Lewis' theory of convention and then proceeds to develop the social conventional framework for normativity (SC<sub>NF</sub>). Some key features of Lewis' convention will be examined and further modified, if necessary, to suit scientific practices. We use, for example, rule rather than regularity when talking about scientific normativity. Assuming a constitutive view, we equate rules or norms with conventions. Along the lines of Robert Sugden, Skyrms, Elster, and others who have followed Lewis and provided their game-theoretic and individualistic accounts of convention to understand social norms, we will use Lewis' convention to understand scientific normativity. By stressing the social foundation of scientific knowledge, our social conventional account avoids both extreme positions, namely, radical methodological relativism and universalist view of methodological rules or norms. Our thesis is that we create norms as conventions through interactions of humans with one another and their environment. The contingency feature of Lewis' conventions can also be observed in our understanding of scientific norms as conventions. Then, we will discuss the role of empirical investigation in the formulation and justification of methodological norms and the resolution of methodological disagreements in scientific practices. This will enable us to appreciate how our SC<sub>NF</sub> can help to resolve the methodological disagreements, such as that between Laudan and Siegel, by empirically appealing to the differing values, interests, and goals of the individual or groups in question. It will also enable us to appreciate how both methodological choices are justified accordingly.

The second part presents two main arguments to defend SC<sub>NF</sub>. We will attempt to show how SC<sub>NF</sub> can effectively reconcile relativism with realism. By maintaining conventionalism only at the level of normativity, our account is not susceptible to threats from radical relativism. SC<sub>NF</sub> does not necessarily entail radical relativism. Accordingly, we maintain framework relative conception of scientific progress. Second, we will provide some arguments for normative pluralism to defend our conventional account. We will argue that SC<sub>NF</sub> does not support a universalist view. We will also argue how variegated understanding of human experiences (aesthetics, legal, ethical, etc.) and multiculturalism lend support to our account.



Faculty of Arts



## Expanding Causal Modelling Semantics

*Giuliano Rosella, University of Turin*

In the present paper, we aim to expand Causal Modelling Semantics of Counterfactuals (CMS) introduced by Pearl (2000) and Galles&Pearl (1998) in order to account for the probability of counterfactuals with complex antecedents. Standard CMS, in fact, is limited to a class of counterfactuals whose antecedent is an atomic formula or a conjunction of atomic formulas. The idea underlying CMS is that a counterfactual  $A > B$  is true at a causal model  $M$ , if the consequent  $B$  is true in the submodel of  $M$  obtained by the intervention  $do(A)$  on  $M$ , where performing the intervention  $do(A)$  on the model  $M$  amounts to a modification of the structure of  $M$  so to make the antecedent  $A$  true at  $M$ . The limited expressive power of CMS is due to the fact that disjunctive interventions of the form  $do(A \text{ or } B)$  are not defined.

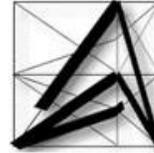
Our work builds upon Briggs' Interventionist Counterfactuals (2012) that extends CMS via Kit Fine's Truthmaker Semantics (Fine, 2017) in order to assign a truth-value to counterfactuals with disjunctive antecedents (of the form  $(A \text{ or } B) > C$ ). Briggs' main contribution is in the use of truthmaker semantics to select the causal (sub)models that make a disjunctive formula true, namely the (sub)models generated by a disjunctive intervention of the form  $do(A \text{ or } B)$  on a model  $M$ . We will also employ some results from the work of Eva, Stern and Hartmann (2019) that allow us to measure the similarity distance among causal models.

Our intuition is inspired by Lewis' semantic account of counterfactuals (1973), according to which a counterfactual  $A > B$  is true at a possible world  $w$  if and only if  $B$  is true at the most similar worlds to  $w$  where  $A$  is true. Intuitively, the probability of a counterfactual  $A > B$  amounts to the probability that  $B$  is the case under the circumstances that  $A$  obtains. So, our idea is that we can calculate the probability of a counterfactual  $A > B$  at a causal model  $M$  by looking at the probability that  $B$  is the case in the causal models where  $A$  is true and that are *the most similar* to  $M$ . And the most similar models to  $M$  are those that are the least distant from  $M$ , according to Eva et al.'s notion of similarity distance. Hence, we set the weight (i.e., likelihood) of the submodels of  $M$  to be inversely proportional to their similarity distance from  $M$ : the more distant from  $M$  a model is, the less *heavy* (i.e., likely) it is. Hence, we propose that the probability of  $A > B$  at  $M$  is the weighted average probability of  $B$  with respect to the submodels of  $M$  generated by the intervention  $do(A)$  on  $M$ .

In conclusion, Briggs has provided us with a procedure to select the relevant models generated by a disjunctive intervention, Eva et al. have shown a way to equip causal models with a similarity order, and by combining the work of the two and borrowing the idea underlying Lewis' similarity semantics, we introduce an expanded version of CMS that allows us to calculate the probability of counterfactuals with disjunctive antecedents with respect to a causal model.



Faculty of Arts



## Wisdom, Action, and Knowledge

*Oushinar Nath, University College London*

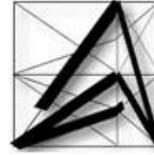
In this talk I attempt to explain the nature of wise actions. I take that a wise action is a kind of successful action – an action, the performance of which, somehow achieves a certain goal. The goals are the goals of certain domains which are certain states of affairs. The relevant actions are any actions that directly realize those states of affairs or are conducive to realizing them. And finally, such actions are successful only if they reach or somehow help in reaching the goals.

Given this characterization of wise actions, I explain them by taking a two-pronged approach. Firstly, I individuate two modal characteristics relevant for wise actions: (a) *counterfactual robustness*, which says that an agent  $S$ 's action  $\varphi$  is wise only if it is performed for the sake of reaching a goal  $G$ , and in all the nearby possible worlds, where  $S$  performs  $\varphi$  while believing that the conditions are normal,  $S$  succeeds in reaching  $G$ , and (b) *rational robustness* which says that  $S$ 's action,  $\varphi$ , is wise only if in all the nearby possible worlds where  $S$  has sufficient reasons to perform  $\varphi$ ,  $S$  performs  $\varphi$ .

Secondly, I argue that these two modal properties are best explained in terms of two central features of knowledge. (i) *Safety*: knowledge is 'safe', i.e., if  $S$  knows a proposition  $p$ , then  $S$ 's belief that  $p$  could not have been easily false. Safety explains counterfactual robustness, for if  $S$  knows that  $\varphi$  will help her successfully reach  $G$  under normal circumstances (while believing that the circumstances are in fact normal), then in all the nearby possible worlds,  $S$ 's belief that  $\varphi$  will help her successfully reach  $G$  under normal circumstances is true. Thus, in all the nearby possible worlds, whenever conditions are normal (and she believes truly that they are normal), and  $S$   $\varphi$ -es on the basis of her belief, she succeeds in reaching  $G$ . (ii) *Stability*: knowledge is 'stable', i.e., if  $S$  knows that  $p$ , then  $S$ 's belief that  $p$  could not have been easily rationally undermined by misleading evidence. Stability explains rational robustness, for  $S$ 's knowledge that  $\varphi$  will help her successfully reach  $G$  in normal circumstances (plus,  $S$ 's belief that the circumstances are in fact normal) entails that  $S$ 's belief that  $\varphi$  will help her successfully reach  $G$  in normal circumstances is not based on any reason that could easily have been defeated by misleading evidence. Thus, in no nearby possible worlds, where circumstances are normal, is  $S$ 's reason for holding the belief that  $\varphi$  will help her successfully reach  $G$  lost due to misleading evidence. Accordingly, in those worlds if  $S$  has sufficient reason to  $\varphi$  then  $S$  will  $\varphi$ .



Faculty of Arts



## **Virtue Responsibilism and a New Kind of Responsibility**

*Niko Šetar, University of Maribor*

Virtue epistemology, a subfield of epistemology concerning cognitive, intellectual, and social behaviours that are conducive to knowledge, is commonly divided into two theoretical approaches. One of them is virtue reliabilism, which advocates that an epistemic agent relies on certain faculties of mind they are sufficiently competent in exercising to obtain knowledge (e.g., Sosa, Greco). The other is virtue responsibilism, which in turn argues that successful acquisition of knowledge happens if one had responsibly and correctly developed certain traits and characteristics (e.g., Zagzebski). While it does well to point out there are ample compatibilist approaches and the two mentioned above need not be divided, we will herein focus mainly on responsibilism.

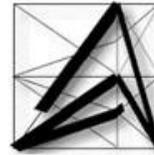
While studying epistemic vices, i.e., the opposite of epistemic virtues, Cassam (2019) defines two variants of responsibility: acquisition responsibility and revision responsibility. The former requires the epistemic agent to strive to acquire virtuous traits, which lead her to obtaining true belief and knowledge, while the latter states that the epistemic agent also has the responsibility to revise and 'get rid' of vicious traits and attitudes that prevent her from gaining knowledge when such traits and attitudes are brought to the agent's awareness.

It is, however, apparent that in certain environments one cannot adequately acquire virtuous traits and attitudes, and in numerous cases the epistemic agent herself cannot be held responsible, and revision might come well too late. Where, then, does the responsibility lie?

We will argue that there is a third kind of responsibility – provision responsibility – which does not pertain to the epistemic agent herself, but rather on figures of authority that guide her acquisition process, such as parents, teachers, and even politicians. The aim of our contribution here will be to explain why acquisition and revision responsibility alone are inadequate, define provision responsibility, and show where such responsibility is most applicable.



Faculty of Arts



## Is Pragmatics a Mental Module?

Edoardo Vaccargiu, University of Genoa

Cognitive approaches to the study of human communication build on the assumption that verbal understanding relies on mind-reading or “Theory of Mind” (ToM), that is, the ability to interpret others’ behaviours in terms of mental states (e.g., intentions, desires, beliefs).

Post-Gricean accounts, like Relevance Theory (RT), elaborate this assumption in a modular view of the mind. RT describe the capacity to interpret speakers’ utterances in conversation (shortly, “pragmatics”) as a sub-module of the ToM-module (Sperber & Wilson, 2002).

Recently, the modular view of pragmatics has been challenged, both empirically and theoretically. Empirically, critics argue that data on preserved pragmatic abilities in people with impaired ToM question the complete overlap between pragmatics and ToM (e.g., Bosco et al., 2018).

Theoretically, critics dwell mainly on RT’s loose formulation of the modular hypothesis. Among these, Mazzone (2018) advances two criticisms: (i) RT endorses a trivial notion of modularity, (ii) RT lacks an explicit account of the relationship between mind-reading in pragmatics and general mind-reading.

In this paper, I address Mazzone’s critical remarks from a relevance-theoretic perspective and I try to theoretically refine the modular hypothesis to provide a novel framework for properly assessing its empirical significance.

Firstly, I spell out the methodological problems underlying the present debate. Then, while partially accepting (i), I propose to further constrain RT’s hypothesis to provide a non-trivial notion of modularity which can be reconciled with existing data from developmental research (e.g., Csibra, 2010).

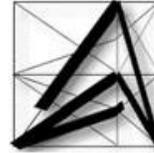
Secondly, I reject (ii) by drawing on Sperber’s (1994) early account. Specifically, I build on the hypothesis that, even when unfolding in an *egocentric* way (i.e., without taking into account the speaker’s beliefs), language interpretation always requires the interpreter to form a metarepresentation involving an attributed intention (i.e., “the speaker intends to communicate that  $p$ ”). Then, I build on the literature on mind-reading (e.g., Baron-Cohen, 1995; Scott & Baillargeon, 2017) to outline a working distinction between low-level metarepresentational skills and proper ToM.

Specifically, I describe low-level metarepresentational skills as aimed at detecting the agent’s *volitional* states (e.g., goals, intentions, purposes) while circumscribing the domain of proper ToM to the realm of *epistemic* states (e.g., beliefs, doubts, guesses). My main claim is that the former - and not the latter - should be regarded as necessary for verbal understanding.

Finally, I provide a wider account that reconciles data on early metarepresentational skills with studies on infants’ biases to communicative behaviours, supporting the hypothesis of a pragmatic module dedicated to mind-reading in intentional communication.



Faculty of Arts



## Probabilities of Actions and the Kinematics of Motivation

*Daniele Conti, University of Manchester*

Libertarians about free will maintain that a necessary condition for freedom is the 'openness of the future': whenever we face a free choice, there are multiple possible actions open to us. Often, however, libertarians hold a further thesis, according to which possible actions have objective probabilities of being performed. Call such thesis 'Objective Probability Thesis' (by objective probabilities we mean probabilities that, broadly speaking, refer to facts or properties of the world, independent of anyone's credences or expectations). According to it, for instance, not only it is undetermined whether tonight I will go to the local pub or I will stay home, but there is also (say) a 0.8 probability that I will undertake the first course of action, and a 0.2 probability that I will undertake the latter. But why posit such probabilities? One often-cited reason rests on phenomenological considerations. Typically, when we face a choice, we are not inclined in the same way towards the options open to us. Rather, we are more motivated to choose some of them over others. Many libertarians hold that a consequence of this phenomenon is that some possible actions are objectively more likely to occur than others.

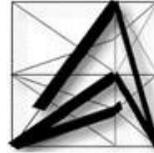
In this paper I will argue that such reasoning, albeit intuitive, is mistaken, and that phenomenological considerations about motivation cannot ground the Objective Probability Thesis. Motivation and probability, in fact, come apart when it comes to their evolution over time. Sure enough, both motivation of agents and objective probabilities of events are the kind of things that can change with time. We become motivated to do things that at a previous time we were not motivated to do, as well as we cease to be motivated towards things we once wanted to do. Similarly, the probability at  $t_0$  of a certain event may differ from the probability at a later time  $t_1$  of that same event. However, the way motivation and probabilities can evolve over time is different in an important respect. It is widely accepted that probabilities evolve according to David Lewis' kinematics of chance, according to which "the endpoint chance of the complete history of [an] interval is the product of the endpoint chances of the complete histories of [its] subintervals".<sup>1</sup> Broadly speaking, the probability at  $t_0$  of a certain event is a function of probabilities of other events recorded after  $t_0$ . Nothing of this sort is true about motivation. My current motivational state is not a function of, nor does depend on motivational states I may have in the future. How much I am motivated now to perform a certain action is something intrinsic to my current psychology in a way that the probability that I will perform that action just is not. Therefore, I will contend that the phenomenon of being more motivated towards certain actions than others cannot support the claim that such actions are objectively more probable than others.

---

<sup>1</sup> Lewis, David. 1986. 'A Subjectivist's Guide to Objective Chance.' in David Lewis (ed.), *Philosophical Papers*. Vol. II (Oxford University Press: Oxford).



Faculty of Arts



## The Leibnizian Criterion and Property Individuation

*Jace Snodgrass, University of St Andrews*

How should we assess the identity of properties? One view says, roughly, to appeal to the properties of properties. I will refer to this view as 'the Leibnizian Criterion'. Here is a simple formulation of it:

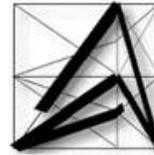
**Leibnizian Criterion:** For any property F and G, F is identical to G if and only if for any property H, F instantiates H if and only if G instantiates H.

The Leibnizian Criterion, in one form or another, has been used to assess the identity of *ante rem* universals (Michael J. Loux 1978: 100-01 and J.P. Moreland 2001: 117), the identity of causal properties in general (Elliot Sober 1982: 183), and most recently, this criterion has been a matter of dispute in debates about the identity of moral properties (Graham Oddie 2005: 149-51, Jussi Suikkanen 2010: 93-4, Bart Streumer 2013: 325-28, 2017: 19-22, and Frank Jackson 2017: 198-201).

However, one question that has received no serious attention in contemporary discussions of the Leibnizian Criterion is whether properties are typed or untyped. This is an unfortunate oversight, given that it has important implications for the metaphysics of the properties of properties; something that has not been forthcoming from proponents, or even opponents, of the Leibnizian Criterion. In order to draw out and explore these implications, I address two issues for the Leibnizian Criterion that arise when typed and untyped properties are taken into account.



Faculty of Arts



## **Arguments Against AI Art and How to Answer Them**

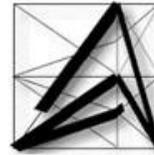
*Tadej Todorović, University of Maribor*

Artificial intelligence (AI) has come a long way since the last century, and even though AI was already impressive at the very start, e.g., the Logic Theorist came up with a shorter mathematical proof for a theorem (Russell and Norvig, 2016: 17-18), few believed that it would become so ubiquitous and necessary in human lives. Today, there are hardly any doubts that AI can outperform us in many domain-specific tasks, like playing games, driving cars etc. And as AI is getting better and better at these domain-specific tasks, we, with more and more uncertainty, diligently move the goalposts, stating that AI will surely not be able to beat us at the next mark. Currently, this mark is general-purpose reasoning abilities, even though experts from the field claim that there is a 50% chance that computers will reach human-level intelligence by 2040 (Bostrom 2014). Of course, Turing, arguably also an AI expert at his time, also claimed something similar in 1950, so this should be perhaps taken with a grain of salt (see Turing 1950). It comes as no surprise then that the one of the last bastions of human uniqueness, art and the creation of art, is fiercely defended against the possibility of AI creating art.

This paper will address arguments from selected authors that wrote about this issue (e.g., Audry and Ippolito 2019; Still and Inverno 2019; Hertzmann 2019; Manovich 2018 etc.), focusing on the following objections against AI art. First, the argument from creativity: (1) could we plausibly argue that creativity is either a necessary, sufficient, or necessary and sufficient condition for art, and (2) can AI be creative or is it already creative? Second, the argument from social agency: could it be claimed that AI cannot produce art because it is not a social agent? Third, the argument from intentionality: is intentionality a relevant condition for art production and can art be produced without intentionality? We will argue that the first and second objection can present a serious argument against AI art only if they boil down to intentionality and that intentionality therefore presents the best argument against AI art. Finally, we offer some arguments against the claim that intentionality should be understood as a necessary condition for art production.



Faculty of Arts



## **Superintelligence and the control problem: Real problem or pseudo-problem?**

Jaroslav Malík, University of Hradec Králové

The goal of this presentation is to judge the legitimacy of the control problem. According to a group of AI theorists, we will soon face the possibility of superintelligence (SI). We have to understand SI as a general intelligence that exceeds the cognitive performance of humans in all relevant domains. People like Nick Bostrom stress that we have to face the danger that SI poses. They conceptualize this threat as the control problem. The idea is that when we construct SI, we might lose control over its behaviour. Bostrom and others argue that we will most likely give SI one final goal. Because SI might have alien motives, it could interpret its goal differently than we would. That presents an existential risk to us. However, there are problems with how they conceptualize this threat. This presentation concentrates on two of their claims: (1) The SI's interpretation will stem from its motivations, (2) SI will be radically different from us. I argue that these positions are mutually exclusive.

The control problem is, at its core, a problem of interpretation. I follow Donald Davidson's account of its practice. When we interpret another person, we do so by projecting our attitudes and beliefs onto their statement. This idea aligns with (1), but there is a caveat. To properly understand an utterance, we have to accept the speaker's beliefs as mostly true. Otherwise, we would not be able to understand others. This requirement somewhat limits the nature of the given motives. If SI's motivations will be so different from our own, should we assign to it a role of an interpreter? Even more so, could we even talk about it as having motivations at all? Therefore (2) needs to be put into question. If SI is to be an interpreter and have motivations, it cannot be too dissimilar.

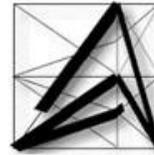
Justification of (2) is based on the concept of the space of possible minds. This space may be more expansive than we currently imagine, but there have to be limitations. The idea is that general intelligence must have specific characteristics. I base my arguments on the ideas of Hubert Dreyfus. I show how his critique of AI still applies to the concept of SI. The reason is that Bostrom and others still depend on assumptions that Dreyfus criticized. They still imagine that a disembodied entity can have a general intellect. The body is beneficial for intelligence because it provides a frame of reference. Without this frame, we could not cope with the world. I argue that embodiment is a necessary condition of general intelligence. In such a case, SI could not be drastically different from us.

What aligns Dreyfus and Davidson is the holistic conception of intellect. If intelligence is to be general, it has to be holistic. One of the consequences is that such intelligence cannot have one final goal. We can observe this when we examine human motivations. Our goals are always understood in the context of other goals. The same applies to interpretation. We can understand others because of the context we all share. For these reasons, I conclude that Bostrom and others have to make a choice. Either they have to stop using the terminology of motivations and concentrate the control problem on limited intelligence or give up the idea of SI being radically different. There is still a problem that we have to solve, but we cannot conceptualize it in this manner.



University of Maribor

Faculty of Arts



## Religious “Knowledge” as a Post-Truth Concept?

*Marina Bajić, University of Maribor*

The word ‘post-truth’ first surfaced when Donald Trump became president of the United States back in 2016. Since then, it has gone on to have many definitions and understanding, such as “where some feel emboldened to try to bend reality to fit their opinions”, “a deliberately complicated relationship with the truth”, or the simple Oxford Dictionary one: “relating to circumstances in which people respond more to feelings and beliefs than to facts”.

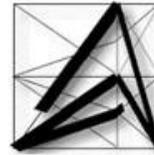
Though many people have taken on many angles on the post-truth subject, I have decided to focus on its relation to religion, or more specifically, religious ‘knowledge’, which is defined as “religion per se is created by God but religious knowledge is human-made. The sacred law is divinely created but its understanding is a human enterprise.” It seems that, almost by definition, religious ‘knowledge’ falls under the category of post-truth, that is why I use the term loosely; the word ‘knowledge’ implies ‘truth’ and in fact, it might be quite the opposite.

Leaning heavily on Morteza Hashemi and Amir R. Bagherpour’s *A Theory of Evolution of Religious Knowledge in a Post-Revolutionary Iran: And a New Frontier for Sociology of Knowledge*, I examine how many definitions of the term post-truth can play into our understanding of religious ‘knowledge’, focusing mainly on Islam, since that is the religion that Hashemi and Bagherpour focus on – that is not to say, of course, that these findings could not be applied to other religions. The main argument that supports my thesis is the fact that religious doctrines have, and continue to have, such divergent readings. For example, in the Quran, a hijab is described as a veil separating the earthly from the divine, while today we all know it refers to the fabric that women in Islam wear to veil themselves from the world. Upon asking ourselves who it was that read the Quran and interpreted that a mystical veil should mean that women must veil themselves, is asking exactly the questions that I have been asking: does religious ‘knowledge’ play by the rules of post-truth?

The answers in my paper are by no means definitive. However, it is a start in viewing an old term, religious ‘knowledge’, in correlation with a more modern concept, post-truth. Indeed, the concept of religious knowledge may be very old, and the concept of post-truth relatively new, but the latter gives us the opportunity to observe the former from a different angle.



Faculty of Arts



## Simultaneous Realism and Anti-realism, and Transcendence and Immanence of Truth

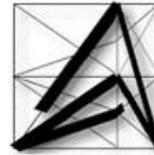
Teemu Tauriainen, University of Jyväskylä

Debates on whether truth is dependent or independent of human concerns are present throughout the history of western thought. Following Aristotle's explication, a popular position is to see that truth is mind independent. In this sense, truth is objective, and to speak truly is to do so in correspondence with how things stand. Others hold that truth is mind dependent. For example, truth is determined by what science has any given time proven to be the case. Further, one might hold that truth is a historically constructed notion and as such intimately tied to our stance regarding its nature. This persistent debate between mind-independent and -dependent understandings of truth has more recently been labelled as the realist and anti-realist debate. For the former, truth depends on what exists independent of minds, so that each truth-apt sentence is necessarily true or false, independent of any human concerns, such as our knowledge about it. For the anti-realists, truth is dependent on human concerns, such as our knowledge about it, or whether something has proven to be useful to believe. In this sense, beliefs can be not only true or false, but true enough, or unknown. This debate has seen a rise in popularity after the emergence of truth pluralism, according to which our understanding of truth should allow for both realist and anti-realist ways of being true. This is only intuitive, for surely there are truths about mind-independent states of affairs, like stones and chairs, and truths about ethics and law, such as what is permitted or prohibited. Despite the intuitive appeal, the truth-pluralist program has thus far failed to attracting widespread popularity. Interestingly enough, an analogous and thus far neglected debate to the realist and anti-realist dispute concerns truths *immanent* and *transcendent* nature. While this debate is prominent in historical literature, it has been largely bypassed by contemporary scholars. According to the claim of truths immanence, found explicitly in the works of Kant and Quine, truth is *necessarily* contained in some framework of understanding. In this sense, truth is dependent on and relative to the framework that gives rise to it. According to transcendentalists, truth rests outside of our frameworks of understanding, exceeding our knowledge about it. Thus, there is clear similarity between truths immanence and the anti-realist stance, and its transcendence and the realist stance.

I argue that the analogy between these distinctions proves illuminating for contemporary debates on the nature of truth. My core argument is that the general interest of accommodating both mind-independent and -dependent aspects of truth is supported by historical literature. After this, I propose an answer to a pressing issue that the current truth pluralist models face in accounting for the unifying feature between the mind-independent and dependent aspects. Truth pluralists are free to argue that the concept of truth is an *immanent feature* of our theories and as such inherently human-dependent, but that the function of this feature is to act as an intermediary between our theories and the theory-independent aspects of reality. Finally, I critically discuss this idea by noting that as an inevitable and for some unfortunate result, anti-realism gets the upper hand, for even the idea of a transcendent and fully objective truth is rendered an immanent feature of our theories.



Faculty of Arts



## Higher-Order Fictionalism and the Infinite Regress

Sofia Melendez Gutierrez, University of Cambridge

Fictionalism about a certain discourse is the view that sentences pertaining to that discourse are useful fictions: they are false because the entities that they quantify over do not exist; and, yet, they are informative—much like metaphors, they non-literally express truths in very convenient ways; and, hence, we should keep using them as we always have. The rationale behind fictionalism, then, is to make irrealism compatible with a wholly non-revisionary stance towards our discourses.

Some of the most popular fictionalist theories concern sentences that quantify over abstract entities, such as mathematical objects and universals. These theories are all subject to an interesting objection that has not received much attention yet, and states that fictionalist theories about *abstracta* are not genuinely nominalistic: they inadvertently quantify over abstract entities, such as sentence-types and fictions; and, hence, they are futile—they amount to exchanging commitment to certain sorts of *abstracta* for commitment to other sorts thereof. Ultimately, then, they commit their advocates precisely to the kind of entities that they are set to reject.

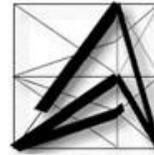
Fictionalists could resist this dreadful conclusion by formulating a higher-order fictionalist theory—a fictionalist theory about fictionalism. Fictionalists could claim, that is to say, that fictionalist theories themselves are useful fictions; that they are false because they quantify over abstract objects, but informative nonetheless. Fictionalists would thereby very effectively avoid committing to fictions and sentence-types in exactly the same way in which they avoided committing to all other sorts of *abstracta* before. As a matter of fact, endorsing a higher-order fictionalist theory is what Balaguer (1998, p. 13) recommended after noting that fictionalism is not genuinely nominalistic; and, in my paper, I will argue that higher-order fictionalism is the only response available to the fictionalists on the face of the nominalistic impurity charge.

Daly (2008, pp. 435-436) argued that this proposal is ineffective: he rightly pointed out that it generates an infinite regress: if first-order fictionalism cannot be formulated in nominalistic terms, then there is no reason to think that higher-order fictionalist theories could: they too will quantify over fictions and sentence-types. Consequently, if fictionalists are to remain nominalistic, they are bound to accept that, for every fictionalist theory  $T$ , there must be a fictionalist theory about  $T$  that allows fictionalists to elude the unacceptable commitments otherwise brought by  $T$ . Merely accepting that this is the case, Daly maintains, fact precludes fictionalists from remaining uncommitted to abstract entities.

I will contest Daly's latter point: I will contend that the infinite regress poses no threat to the nominalistic purity of fictionalism, as fictionalists can adhere to a fictionalist hypothesis even about the sentence that asserts that such an infinite regress exists. However, I will argue, this infinite regress is theoretically vicious, as it precludes fictionalists of all strands to fulfil their aim of establishing that a non-revisionary position in relation to our discourses is compatible with irrealism. To conclude, I will show that, independently of this viciousness diagnosis, the infinite regress makes the whole project of fictionalism about *abstracta* futile.



Faculty of Arts



## A Proposed Solution to the Preface Paradox

*Pak-Him Lai, Texas A& M University*

The preface paradox illustrates that the normative Lockean thesis, the conjunction principle, and the consistency principle can't be consistent with each other.<sup>1</sup> In this talk, I propose a new solution by rejecting the consistency principle and argue for a type-theoretic model of belief. I argue that the crux of the preface paradox is the self-evident but questionable assumption that all our beliefs belong to a single type. Because of this assumption, we are tempted to think of a person's overall doxastic state as a single set and thus inconsistent beliefs are always considered as irrationality. Instead of thinking all beliefs of a person as the members of a huge, single set, I propose to allow that there is a stratified hierarchy of belief-types. On this view, beliefs belong to different *types* or *levels* rather than a single set. If one's overall doxastic state is understood as a stratified hierarchy of belief-types, then one can be justified in holding first-order belief 'p' and higher-order belief '~p' at the same time without inconsistency. If belief is so understood, we may be able to overcome the problem of the preface once and for all.

In a type-theoretic model, there are at least two types of beliefs. Following the convention of type theory, we may say that first-order beliefs belong to type-0 level, and second-order beliefs belong to type-1 level. A higher-order belief is a belief about what one should believe. In the preface case, one has a first-order belief that (i) every claim in the book is true and a second-order belief that (ii) some claim in the book is false. According to the normative Lockean thesis, one is justified in holding both types of beliefs since one is rationally highly confident of both. Apply the conjunction principle in the usual manner. Now, one is justified in holding both first-order belief (i) and second-order belief (ii) at the same time. But there is no inconsistency at all. For we don't presuppose that her beliefs (i) and (ii) belong to a single set and thereby violate the requirement of logical consistency. According to the type-theoretic model, one's overall doxastic structure is characterized by a stratified hierarchy belief-types. So, one can rationally believe that (i) in type-0 level and that (ii) in type-1 level without being logically inconsistent. If the type-theoretic model of belief is plausible, one can rationally doubt what one believes. It sounds paradoxical in a set-theoretic model of belief, but it makes good sense in a type-theoretic model. After all, epistemic self-doubt and reflection are ordinary epistemic activities. This type-theoretic model allows us to resolve the preface paradox and accommodate the epistemic importance of self-doubt and reflection.

---

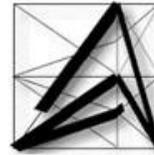
<sup>1</sup> 1. The normative Lockean thesis: A person X ought to believe that p iff X rationally has a degree of belief x in p such that  $x \geq T$ , where T is a threshold.

2. The conjunction principle: For any beliefs p, q, r, etc., if it is rational for a person X to believe that p, q, r, etc., then X is rational in believing the conjunction (p & q & r & ...).

3. The consistency principle: It can't be epistemically rational for a person to hold inconsistent beliefs.



Faculty of Arts



## Material-origin essentialism or necessity of sex; but what about gender?

Lucija Duda, University of Manchester

When it comes to personal identity, most analytic philosophers affirm to material-origin essentialism or gametic essentialism, according to which a person cannot fail to come from the exact sperm and egg which produced her essentially. This well-known thesis, introduced by Saul Kripke (1980), is by so far only supported by the intuition that an ordinary language speakers find that counterfactually “it is harder to imagine /.../ her [the Queen] being born from different parents”<sup>1</sup> than imagine her having a different subsequent history. On the contrary, anti-essentialists who oppose Kripke’s necessity of origin defend the view that persons do not have any of their non-logical properties essentially in the sense that they could not fail to have them and remain the same person (Janssen-Lauret, 2021; Janssen-Lauret and MacBride, 2015). In fact, they appeal to modal relativity about properties, claiming that essential and accidental properties are relative to a given inquiry, not fixed absolutely. Therefore, they contradict material-origin essentialism by showing that counterfactual statements which in antecedent ascribe alternative material origin to a referent (i.e., AMO statements), such as different DNA or genes, while holding some other characteristics fixed, say social or psychological, are regularly used by ordinary language speakers. Therefore, an anti-essentialist like Janssen-Lauret (2021) would assess my counterfactual “If I was a cis-man, I would not take the longest route home just because it is better lit” as a literally true statement about me *qua* me.

In this paper, I reject material-origin essentialism and embrace anti-essentialism and modal relativity. On these grounds, I claim that from Saul Kripke’s (1980) necessity of origin follows that biological sex is an essential property of an individual. If every person is coming from the exact sperm and egg which produced her essentially, and biology shows that sperm and egg carry genetic information about development of all the layers of biological sex, then a person possesses her biological sex essentially. If that is true, then self-identification or gender is a necessary contingent property and persons are reduced to their biological substratum. And this does not accord with the use of gender and sex terms in ordinary language, especially in the trans\* contexts. Therefore, I will support my thesis with real-life examples of AMO statements about sex delivered by trans\* individuals, whose gender does not align with sex assigned at birth, such as “If I had a man’s genitals, I would be much freer with my potential lovers.”<sup>2</sup> In this way, I show that biological sex can be an inessential property, whereas an individual can hold their gender necessary. Finally, my thesis show that anti-essentialism and modal relativity are socially aware alternatives to material-origin essentialism, which reduces a person’s identity to a biological substratum which might be the cause of their sufferings.

---

<sup>1</sup> Kripke, S. 1980. *Naming and Necessity*. Harvard University Press, 1980, p.113.

<sup>2</sup> Jury, H. 2012. *Testimonies of Genital Surgery by Transsexual Men*. Ed. Cotton, Tristan. Transgender Press, p. 123.