BALKAN JOURNAL OF PHILOSOPHY

Vol. 16, Issue 2, 2024

The issue is sponsored by the Bulgarian National Science Fund, competition “Bulgarian Scientific Periodicals – 2024”, project No: КП-06-НП5/80 from 08.12.2023

TABLE OF CONTENTS

IN MEMORIAM: Nenad Miščević (1950–2024)

On May 11, 2024, the community of analytic philosophers in Southeast Europe lost one of its most prominent representatives who had a profound impact on academia in Central and Eastern Europe and beyond.

Born on November 1, 1950 in Zagreb, Croatia (at that time part of former Yugoslavia), Nenad Miščević studied philosophy in Zagreb, Chicago (with Paul Ricoeur), Paris (with Jacques Derrida), and Ljubljana, where in 1981 he earned his PhD under the supervision of Prof. Frane Jerman. He began his philosophical journey engaged with the post-structuralism endorsed by thinkers such as Foucault and Derrida, before transitioning to analytic philosophy, inspired by Dummett's work on Frege.

Nenad Miščević first taught philosophy in Zadar (Croatia), but during the Yugoslav Wars in the early 1990s, he moved to the University of Maribor (Slovenia). In different periods of his career, he also taught and held positions at the University of Rijeka, University of Klagenfurt, University of Graz, and the Central European University in Budapest, being affiliated with the latter for more than 20 years.

Nenad Miščević's philosophical interests were extremely broad, ranging from epistemology, philosophy of language and philosophy of mind to ethics and political philosophy. He published 17 books, including Thought Experiments (2022, Springer), Curiosity as an Epistemic Virtue (2020, Palgrave Macmillan), Nationalism and Beyond: Introducing Moral Debate about Values (2001, CEU Press), and Rationality and Cognition: Against Relativism-Pragmatism (2000, University of Toronto Press). His work on thought experiments and nationalism is particularly noteworthy. Nenad Miščević's contributions to the study of thought experiments, as summarized in his recent book Thought Experiments (2022), offer a deep dive into the methodology and role of thought experiments in philosophy. The earlier book Nationalism and Beyond: Introducing Moral Debate about Values (2001), together with the entry "Nationalism" written for the "Stanford Encyclopedia of Philosophy" (first published in 2001, substantive revision in 2020), best represent Nenad Miščević's views on the ethical dimensions and the challenges of nationalism, which have been recognized as a significant contribution to contemporary political and moral philosophy.

Nenad Miščević's career was distinguished by his exceptional teaching and mentorship. He guided numerous undergraduate, graduate, and doctoral students, helping them navigate complex philosophical inquiries and fostering both their intellectual and academic growth. He was known for his unwavering commitment to helping students from different countries to get accepted into prestigious doctoral programs, mentoring them constantly and assisting them in preparing for high-level academic events and publications. Miščević played a crucial role in helping his students further secure academic positions, ensuring their successful integration into the global philosophical community. Many of his students have gone on to make significant contributions to the field, a testament to his influence and guidance.

Nenad Miščević made tremendous efforts for the spread and institutionalization of analytic philosophy in the Balkan region. He played a pivotal role in establishing the philosophy departments at the University of Rijeka and the University of Maribor, served as a co-director of two of the annual conferences held by the Inter-University Center in Dubrovnik, and was a member of the editorial boards of the newly established journals in the region, including Acta Analytica, Balkan Journal of Philosophy, Croatian Journal of Philosophy, and European Journal of Analytic Philosophy. In recognition of his contribution to the development of analytic philosophy in Europe, Nenad Miščević was elected as the president of the European Society of Analytic Philosophy (ESAP) in the period 1996–1999.

Probably few people know that Balkan Journal of Philosophy (BJP) owes its name to Nenad Miščević. He embraced the idea of a journal aimed at bridging the diverse traditions of philosophy in the Balkans and gladly accepted the invitation to serve as a member of the BJP Editorial Board. He stayed at this position from the founding of the journal in 2009 until his unexpected death. As a member of the BJP Editorial Board, Nenad Miščević contributed with ideas and publications to several thematic issues, including those on “The dialogue between analytic and continental philosophy” (2011), “The 21st century metaphysics” (2013) and “Tolerance and its limits” (2019).

Beyond academia, Nenad Miščević was an engaged public intellectual, advocating for multiculturalism, humanism, and social justice. His public writings and appearances enriched public discourse and highlighted his commitment to societal issues.

Those of us who knew Nenad personally were highly impressed by his intellectual presence, openmindedness, unlimited curiosity, and sense of humor. He was one of those people whom you would like to meet at least once in your life. Nenad Miščević spoke many languages fluently, but more importantly –he spoke everyone’s language. This is one reason why he was considered a very talented mentor and a desired interlocutor.

Nenad Miščević’s academic presence and his exceptional role as an educator and mentor will be sorely missed, but the memories of his enthusiasm and devotion to philosophical wisdom will continue to inspire the quests of many of his students and fellow philosophers.

Marina Bakalova, Editor-in-Chief
Institute of Philosophy and Sociology,
Bulgarian Academy of Sciences

Boris Grozdanoff
Institute of Philosophy and Sociology,
Bulgarian Academy of Sciences

Lilia Gurova,
New Bulgarian University

Articles

Invited paper

NATURALISM AND THE INTELLECTUAL LEGITIMACY OF PHILOSOPHY

Hilary Kornblith

University of Massachusetts – Amherst, U.S.A.
email: kornblith@philos.umass.edu
ORCID ID: https://orcid.org/ 0000-0003-3940-1074

Abstract:

There is a worry about the intellectual legitimacy of philosophy. Although the sciences have a progressive history, with later theories largely building on earlier ones, and a tremendous amount of agreement within the scientific community about the approximate truth of current theory, philosophy is different. We do not see a progressive history of philosophical theorizing, and there is little agreement within the philosophical community about which theories are even roughly correct. This not only encourages a certain skepticism about the possibility of philosophical knowledge, but a real question about whether the activity of philosophical theorizing is intellectually legitimate. It is argued here that there are features of a naturalistic methodology in philosophy which have the resources to respond to these concerns.

Key words: naturalism; philosophical methodology; disagreement; skepticism; intellectual legitimacy.

1. There is a striking fact about the history of science: the sciences, at least in the last few hundred years, show a progressive history (Boyd 1981; Putnam 1975). Later theories tend to make predictions which are both more accurate than earlier theories, and which cover a wider range of phenomena. They have greater explanatory power. In many cases, they allow for a wider range of technological applications. It is this fact about the progressive history of scientific theorizing that provides the best argument for realism about scientific theories. It provides the strongest reason for believing that our theories are at least approximately true. For this reason, there can be little doubt about the intellectual legitimacy of scientific theorizing.

Philosophy, however, does not seem to have a progressive history. Philosophy has certainly changed over time, but it has not shown the kind of progress that we see in the sciences. The problem is not, of course, that philosophy has failed to produce predictions or explanations of physical phenomena, since such predictions and explanations seem to be no part of the goal of philosophical theorizing. Still less is the problem that philosophy has not spun off any technological innovations. The worry is an epistemic one. When problems arise within scientific theories, by and large, the problems are resolved over time. When there are competing theories within the sciences, by and large, further investigation and the accumulation of further evidence tends to resolve the conflict in favor of one theory rather than the other. Philosophy, however, doesn’t seem to be like this. The conflict, for example, between consequentialist moralities, and deontological views, and virtue theories shows no sign of resolution. The conflict between views of compatibilism and incompatibilism is much the same. The conflict between internalist views in epistemology and externalist views continues unabated. There are developments within each of these different competing theories; the best current version of each of these views looks different in important ways from earlier versions. But the fundamental conflict remains entirely unresolved. And this is utterly different from what one sees in the sciences.

It will not do to suggest that the goal of philosophical theorizing is not to have approximately true theories but rather to achieve understanding. It is one thing to claim that understanding is *a* goal of philosophical theorizing; quite another to claim that having theories that are approximately true is not also such a goal. It is difficult indeed to see how understanding of knowledge, freedom, morality, and so on might be achieved without also achieving a good deal of knowledge about such phenomena. The claim that the goal of philosophical theorizing is understanding thus does not offer an alternative to the view that progress in philosophy requires progressively true theories.

Because the progressive character of scientific theorizing provides our reason for thinking that current theories are at least approximately true, and therefore provides our reason for regarding the scientific enterprise as an intellectually valuable one, the lack of such a progressive history in philosophy is especially worrisome. It suggests that, whatever theories we may favor within philosophy, there is no good reason to believe that they are actually true or even approximately true. And this, in turn, suggests that there is reason to question the very intellectual legitimacy of philosophical theorizing.

2. Some will argue that philosophy does have a progressive history. Scott Soames (2014, 2017) offers a history of philosophy in the 20th century which paints it in progressive terms. Timothy Williamson (2014) offers such a progressive view as well. Robert Brandom (2009) takes a longer view, situating his progressive account of the history of philosophy in ideas found in Kant and Hegel. In moral philosophy, Derek Parfit (2011) offers an account of the history of ethics which sees it in progressive terms, ultimately uniting moral theories which others had seen as competitors. Many philosophers throughout the history of the discipline have offered progressive accounts of its history, typically, of course, culminating in their own work. No doubt each of the readers of this paper will have views about the ways in which philosophical theorizing has developed over time, making progress on the issues which concern them most. As a naturalist, there is a story about the history of philosophy that I sometimes tell myself which culminates, unsurprisingly, in the triumph of naturalism. So, there is no shortage of accounts of the history of philosophy in progressive terms.

This plethora of progressive accounts of the history of philosophy does not resolve the problem I raise; it is, instead, a symptom of the problem. There is a progressive account of the history of the sciences, but we do not have fundamentally different stories about the history of physics, or chemistry, or biology, culminating in radically different theories within each of these areas. Within the sciences, there is broad agreement about a single progressive story, and it is because of that agreement that theorizing in the sciences can draw on earlier theories in ways that have led to the many sorts of progress — in prediction, explanation, and technological application — mentioned above. It is the lack of this sort of agreement within philosophy that raises doubts about what we are doing in this field, whether we are, in the end, latching onto the truth, or nearly so, on the various topics we investigate. And it is the lack of this sort of agreement, the lack of a single progressive account, that raises doubts about the intellectual legitimacy of philosophy.

3. In this paper, I propose a solution to this problem. I offer an argumentative strategy for philosophy which is recognizably naturalistic, which is well-represented in the literature, and which puts philosophy, or at least philosophy which adopts this argumentative strategy, in good epistemic standing. I present examples of a wide range of philosophers who have adopted this common strategy in order to show that my ideas here are anything but idiosyncratic. The strategy I propose is widely, though certainly not universally, adopted. I would not defend all of the conclusions that the various philosophers I cite have reached. One may agree that the strategy these philosophers adopt is the right one for legitimate philosophical theorizing, and yet disagree about the details of the particular implementation of that strategy.

I will not argue that this is the only strategy one might pursue if one wishes to defend the intellectual legitimacy of philosophy. Perhaps there are others as well. All the same, the legitimacy of philosophy is surely under threat for the reasons sketched above, and if we can find even one approach which will escape this argument, that will be a real achievement. If others wish to suggest that there are additional ways in which the threat can be avoided, so much the better.

3.1. Jerry Fodor's defense of the language of thought (1975) makes the following argument. The cognitive sciences constitute a rich body of theorizing about the mind, and work within this field, just in the two decades prior to Fodor's work, has shown a progressive history: the discipline has made progress in understanding how cognition works, offering ever greater and more accurate predictions and explanations. These theories offer computational accounts of cognition, and there is a presupposition underlying such theories which Fodor brings out here. Computation is not possible without mental representations; representation requires a representational medium; any such medium can only do the work required if there is a language of thought.

Let us forget, for the purposes of this paper, whether Fodor is right about the language of thought. What I am interested in here is his argumentative strategy. Fodor begins with work in a well-established scientific tradition and argues that his conclusion — that there is a language of thought — is presupposed by that work. The fact that the scientific tradition he takes as his point of departure is well-established, that it is not some out of the way wildly controversial approach to the topics it addresses, speaks for itself: work in this field has the authority of science. It is not that apparently successful approaches within the sciences cannot, in the fullness of time, lead to a dead end and need to be rejected. But to the extent that there is a well-established research tradition within some science that has a progressive character, there is currently every reason to view the results of that tradition as presumptively true. One can do no better than this. This provides one with excellent reason to believe that the results of such theorizing are at least approximately true. But now if these theories within the cognitive sciences share the presupposition which Fodor has brought to light, that is the best sort of reason one could have for believing that the presupposition is true.

The language of thought hypothesis, on this way of viewing things, is a well-confirmed scientific theory. What does it have to do with philosophy? There can be little doubt that this view is deeply relevant to a wide range of philosophical issues about the mind which have been the subject of philosophical theorizing for centuries. Questions about innateness, about the a priori, about the nature of inference and its justification, and, of course, the nature of thought, are all illuminated by the language of thought hypothesis. So the theory which Fodor offers makes an important philosophical contribution. By deriving this view from well-established work in the sciences, the intellectual legitimacy of the enterprise in which Fodor is engaged is thereby assured. Here we have work which has the epistemic standing we were looking for in philosophical theorizing, and it achieves that standing by being rooted in independently well-established science.

3.2. One more example from Fodor may be found in his work on the modularity of mind (1983). Fodor's work here is influenced in a very direct way by Chomsky's account of language acquisition (1957, 1964). Earlier theorists, such as Skinner (1957), viewed the learning of language as a product of domain-general learning strategies. On this view, we learn language by way of the very same learning strategies that allow us to learn to tie our shoes, to learn physics, to learn auto mechanics or plumbing, or to learn anything else. There is a set of learning strategies which work equally well on any of the topics we are capable of learning. Chomsky, famously, argued that language learning is not like this. We have an innate set of cognitive capacities which allow us to learn language and nothing else. Chomsky's devastating arguments against Skinner's account of language learning (1959), together with the positive development of his account of how language learning proceeds, led to a revolution in linguistics and a highly developed research tradition with an impressive record of progress.

Fodor argues that not only Chomsky's account of language learning, but available accounts of perceptual input systems, offer a picture of the large-scale structure of the mind. Input systems are cognitive modules which are able to perform the rapid processing of information required of them precisely because of their modular structure. The output of these modules is then delivered to central processing where the results of the different input systems may be brought to bear on one another. Fodor offered a rich and detailed account of the structure of the mind, one which had obvious implications for a wide range of philosophical issues.

Once again, my interest here is not in the details of Fodor's argument or in defending his account of modularity. Rather, my interest is in the argumentative strategy which Fodor employs in developing the philosophical conclusions he reaches. By starting with a well-developed scientific theory, and arguing that his philosophical conclusions are consequences of that theory, Fodor assures the intellectual legitimacy of the enterprise he is engaged in.

I will be far briefer in my discussion of other theorists.

3.3. Ruth Millikan's work (1984, 1993, 2005, 2006, 2019) takes a different approach to theorizing about the mind. Throughout this large body of work, Millikan situates her ideas about the nature of mind in biology, showing its implications for an extraordinarily rich body of philosophical issues. The epistemic standing of biological theorizing is beyond question. The status of her philosophical conclusions is then assured, at least if her derivation of these views from biology is accurate. These philosophical views earn their epistemic standing by being rooted in well-established scientific theorizing.

3.4. Tyler Burge presented a wide range of arguments in favor of an anti-individualist account of mental content, many of which did not employ the argumentative strategy I am concerned with here. (See, for example, his (1979) "Individualism and the Mental.") But Burge did present an argument (1986) for anti-individualism which used exactly the argumentative strategy with which I am concerned. Burge examined David Marr's theory of vision (1982), and he argued that one can only make sense of the explanations Marr offers by supposing that the content of the mental representations involved in visual processing have anti-individualist content. A well-confirmed scientific theory is used here, once again, as a source of philosophically rich ideas, and the philosophical views thus earn their legitimacy by way of the scientific theory in which they are rooted. Burge's more recent work in philosophy of mind (2010, 2022) makes extensive use of this strategy.

3.5. The argument for scientific realism due to Hilary Putnam (1975) and Richard Boyd (1981), with which I began this paper, itself involves an application of the strategy under discussion. In examining the history of successful sciences, Putnam and Boyd detect a crucial feature of theory change: later theories tend to entail the approximate truth of the theories they replace. Crucially, it is not just that observational predictions of later theories largely contain and expand upon the observational predictions of earlier theories. Rather, the unobservable objects, properties, and causal relations posited by the theories are approximately preserved in successor theories. The fact that successor theories, which are arrived at as a product of such a constraint, tend to be more successful than earlier theories in prediction, explanation, and technological application then provides us with a reason for thinking that the realist methodological presupposition which led to these theories must itself be true.

3.6. Putnam's (1973) defense of the causal or historical theory of reference depends on use of this strategy as well. It is important to distinguish Putnam's argument for this view from Kripke's (1980). Kripke, famously, sees his argument for the causal theory of reference as rooted in an appeal to intuition. Much as many others have expressed serious doubts about the epistemic status of the kinds of intuitions to which philosophers often appeal in support of their theories (see, for example, many of the papers in DePaul and Ramsey 1998), Kripke holds that intuitions about hypothetical cases provide an exceptionally secure foundation for philosophical theorizing.

Some philosophers think that something's having intuitive content is very inconclusive evidence in favor of it. I think it is very heavy evidence in favor of anything, myself. I really don't know, in a way, what more conclusive evidence one can have in favor of anything. (Kripke 1980, 42)

But Putnam's argument for the causal theory of reference is not rooted in appeals to intuition. Rather, Putnam's argument is part of his response to views in the philosophy of science, such as Feyerabend's (1975) and Kuhn's (1962), which raise doubts about the objectivity of science. Putnam argues that descriptions theories of reference fail to make sense of the history of science. In particular, scientists who hold differing views about some phenomenon, such as Newton and Einstein, are viewed as talking about different things, since the referents of their terms are determined by the descriptions they associate with them, rather than disagreeing about some single thing. When Newton says that mass is independent of velocity and Einstein says that it is, instead, dependent on velocity, a descriptions theory of reference makes the referent of Newton's use of the term 'mass' different from the referent of Einstein's use of that term. Instead of disagreeing, this view thus has the two of them talking past one another. As Putnam makes clear, any such view makes a hash of the history of science. A constraint on a theory of reference that makes sense of the history of science is that it allows for commonality of reference across theoretical disagreement, and it is this that leads Putnam to the causal theory. Putnam's argument for the view, unlike Kripke's, is an empirical argument rooted in the sciences.

3.7. My (2002), *Knowledge and its Place in Nature*, offers a defense of a reliabilist account of knowledge which makes use of this argumentative strategy as well. In substance, my account of knowledge is very much like Alvin Goldman's (1986), but while Goldman offers his account of knowledge as a matter of conceptual analysis — an analysis of the folk concept of knowledge — I argue that knowledge is a natural kind and Goldman's account offers the best available view of the nature of that kind. I argue that knowledge plays a crucial explanatory role in theories in cognitive ethology, and the fact that our best theories of animal behavior appeal to such a thing gives us reason to believe that knowledge is the very state about which theorists of animal behavior give an account. The account they give, I argue, bears a striking resemblance to the view which Goldman arrives at by other means. More than this, just as the referent of the term 'water' was H_2O even before the chemical composition of water was known, the referent of the term 'knowledge' was reliably produced true belief even before theorists of animal behavior developed their evolutionary account of the functional role of the mental states of animals. As I see it, such a foundation for the reliabilist account of knowledge provides a far more secure grounding of that view than appeals to conceptual analysis. This is just an application of the argumentative strategy under discussion here.

3.8. The KK Thesis, the view that knowing that *p* entails knowing that one knows that *p*, has been very widely rejected in recent years. Dan Greco (2014) has offered a striking defense of the view. Greco offers a variety of considerations in favor of the KK Thesis, but one important strand of his argument makes use of our argument strategy. Greco argues that the notion of common knowledge plays an important explanatory role in computer science, in game theory, and in explanations of social conventions in a variety of social sciences. Common knowledge involves the infinite hierarchies of knowledge required by the KK Thesis and which many who object to it have argued are never actually instantiated. By rooting these infinite hierarchies in successful theories in the sciences, Greco makes use of the strategy under discussion.

3.9. Well-being plays an important role in many moral theories, as well as in accounts of human flourishing. Some have sought to illuminate the nature of well-being by conceptual analysis, but a variety of recent theorists have offered accounts of well-being which are grounded in the social sciences. Work by Michael Bishop (2015), Daniel Haybron (2008), and Anna Alexandrova (2017) are representative. These philosophers see any defense of an account of well-being as requiring just such scientific backing. And this is just an application of our argumentative strategy.

3.10. The strategy has been employed, as well, in explaining the nature of the good and the right in moral theory by Cornell realists such as Richard Boyd (1988), Nicholas Sturgeon (1985), and Peter Railton (1986).

3.11. Work in philosophy of science, and especially in philosophy of physics and philosophy of biology, has long been a home for precisely this kind of inferential strategy. See, for example, Dupre

(1993), Kitcher (2003), Ladyman and Ross (2009), Maudlin (2011), Sklar (1977), Sober (1994, 2000), Sterelny (2001, 2014), Wallace (2014).

3.12. Penelope Maddy (2007) presents a version of this inferential strategy as a general method for philosophical theorizing.

4. I turn now to a consideration of some objections.

4.1. One might object that this argumentative strategy leaves philosophy hostage to the vicissitudes of science. Although it is true that science has a progressive history, it is also true that scientific claims which are well-supported at one time may be given up when new evidence arises; scientific progress is not a simple matter of adding new claims to old without any revision. As a result, building philosophy on our best available scientific theories leaves open the possibility that our philosophical views will be based on falsehoods rather than truths. It is for this reason that some would see the strategy proposed here as leaving philosophy a hostage to the inevitable revisions in current scientific claims. I would put this differently: this strategy leaves philosophy hostage to the facts. It would certainly be unwise to defend philosophical theories which are rooted in highly controversial scientific views,[1] but that is not what is being suggested here. Rather, the suggestion is that philosophical views should be grounded in our best available scientific theories. This provides an exceptionally secure, even if fallible, foundation for philosophical theorizing. Since even our best available scientific views may be revised with the passage of time and the availability of new evidence, our philosophical views are subject to revision in the very same way. But far from being a fault in the proposed strategy for philosophical theorizing, this is a source of its strength. If philosophy can be given a foundation which makes its views no less secure than views in our best available sciences, this provides us with a robust response to the worry that the history of philosophy fails to show the kind of progress that we see in the sciences. By adopting the proposed argumentative strategy, philosophy can achieve a record of progress that leaves no room for doubt about its intellectual legitimacy.

4.2. One might object that the kind of empirical grounding of philosophical theorizing, whatever its successes in some areas, will not cover all areas of philosophy. For example, it seems implausible as a strategy for explaining the nature of mathematical knowledge.

I do think that it is an open question as to how far the strategy I endorse may be pushed, and how much of philosophy may be grounded in such a strategy. At the same time, I believe there is some reason for optimism here. As indicated above, some areas which one might have thought would be especially resistant to this strategy, such as work in moral philosophy, have actually been developed by a number of philosophers in encouraging ways. It would simply be a mistake to suggest that the current strategy obviously fails for normative inquiries within philosophy. Any such claim would have to overlook not only the work of moral realists, but the work of naturalistic epistemologists as well. There are even promising accounts of mathematical knowledge which make use of this strategy. Philip Kitcher's (1983) makes a case for seeing mathematical progress as rooted in the empirical sciences, just the sort of thing which one might have thought was antecedently implausible. At the same time, it is important to acknowledge that it remains an open question as to just how much of philosophy may be amenable to this approach. If there are areas of philosophy which prove resistant to the argumentative strategy recommended here, then some other account of the intellectual legitimacy of those areas will be needed.

4.3. It may be objected that the proposed argumentative strategy doesn't answer distinctively *philosophical* questions. Instead, it changes the subject and simply offers us something to do *instead* of doing philosophy.

It's hard to know precisely how to respond to such an objection. The many illustrations I have given above all involve views not only put forward by people who are, by every reasonable account, philosophers, but also views which address familiar

[1] It is not at all unreasonable to consider what the philosophical implications of such views are so long as one recognizes the highly tentative nature of any such implications.

philosophical issues, such as the nature of thought, of the a priori, of the mental, of knowledge, of the good and the right, and so on. These traditional philosophical topics are, admittedly, addressed by non-traditional means,[2] but this does not make them any less philosophical. I believe that this objection relies on an overly narrow and contentious view of what philosophy is, but it deserves a more detailed response, to which I now turn.

5. There is a story about the origins of analytic philosophy which is widely told, one which would support the objection that the argumentative strategy I propose provides us, not with a way of lending intellectual legitimacy to philosophy, but, instead, with a way of doing something other than philosophy. Because the story is so widely told, and, indeed, so widely believed, it deserves some real attention.

The origins of the analytic tradition are frequently traced to Frege. One of the many important themes in Frege's work, and one which played a central role in the development of the analytic tradition, was Frege's attack on psychologism. Martin Kusch (1995) provides a fascinating and detailed history of the debate over psychologism in the German-speaking world at the close of the nineteenth century, with Frege and Husserl (1900/2001, 1901/2001) taking the anti-psychologistic side of the debate. As Kusch makes plain, exactly what psychologism was is often quite unclear in much of this debate, both in those who defend it, and in those, such as Frege and Husserl, who reject it. For some, psychologism was the view that the subject matters of logic and mathematics — the truth-makers of claims in those fields — were themselves psychological. For others, psychologism was a view about the nature of epistemic justification, one which tied questions of justification to psychological questions, rather than to features of the quality of arguments, which were viewed as entirely independent of psychology. It was widely held by partisans on both sides of the debate about psychologism that the resolution of this issue had important implications for the very nature of philosophy. Rejecting psychologism was seen as a necessary condition for defending the autonomy of philosophy, while the proponents of psychologism were seen as rejecting any such claim of autonomy.

As Kusch (1995) argues, there was a tremendously important sociological vector playing a role in this debate. The close of the nineteenth century saw the rise of experimental psychology as an independent science. In German-speaking universities, and in much of the world, psychology was housed in philosophy departments, and it was only late in the nineteenth century that psychology separated off from philosophy into a separate department. The rancor which characterized the debate over psychologism, Kusch argues, had a good deal to do with the intensely hierarchical structure of universities in the German-speaking world, with a very small number of professors in each department. As psychology was emerging as a separate discipline, with its own disciplinary standards and methods, but before it was housed in a separate university department, the issue of whether professorships would go to philosophers or psychologists was, understandably, extremely fraught. As Kusch tells it, the debate over psychologism, and over the autonomy of philosophy, was lent a good deal of fuel by concerns over the autonomy of philosophy *departments*.

This is not to deny that there is a legitimate question about the extent to which issues in epistemology about the nature of justification should be seen as dependent on psychological facts or, instead, independent of them. This is, in my view, an exceptionally important issue. As the analytic tradition emerged and began to flower in the work of Russell and Moore, and, still later, in the logical positivists, Frege was held up as the champion of an idea which was central to the development of that tradition, namely, the autonomy of philosophy, and the divorce between questions about epistemic justification and psychological questions about how beliefs are actually formed. Indeed, analytic philosophy became tied to a view about philosophy in general, which saw philosophical issues and their resolution as radically independent of empirical issues and their resolution.

[2] I would actually want to qualify this. As I will argue in the next section of this paper, the idea that philosophy is independent of the sciences, much as it was part of the self-image of analytic philosophy for much of the first half of the twentieth century, was, in my view, a historical aberration. For an extended defense of the kind of view of philosophy I favor, and to which I am much indebted, see Kitcher 1992.

The methodological strategy I have proposed in this paper is thus a break from that particular aspect of the analytical tradition. But it is important to see that the feature of the analytic tradition from which my proposed methodology does indeed separate itself is a feature of a relatively brief period in one recent philosophical tradition, not a feature of the history of philosophy generally. If I am right, the kind of empirical grounding I have suggested for philosophical claims can serve as one important way to defend the intellectual legitimacy of philosophy. Right now, I know of no other.

6. I have argued that there is reason to question the intellectual legitimacy of philosophy, and the problem arises from the fact that philosophy does not display the kind of progressive history that we see in the sciences. This is a problem because the central reason for believing that the claims of the sciences are at least approximately true is dependent upon seeing our current views as emerging from just such a progressive history. It is not the mere fact that current scientific theories provide us with roughly accurate predictions of the results of experiments that justifies our beliefs in scientific claims. Rather, it is the fact that, over time, scientific theories have become ever more successful in their experimental predictions, in their explanations of phenomena, and in their technological applications. The historical development of scientific theories has depended upon a methodology which builds on earlier theory, not only taking on its observational consequences, but its theoretical commitments as well. And the fact that such a methodology has been instrumental in producing a history of progress is reason for thinking that the presuppositions upon which that methodology was built are at least approximately true. It is this pattern of progress which is at the heart of any defense, not only of realism in philosophy of science, but of the truth, or approximate truth, of scientific claims, and, along with that, of the intellectual legitimacy of science.

The obvious progress of scientific theorizing has not left much room for serious questioning of the intellectual legitimacy of science. But the lack of progress in philosophy should, as I have argued, make us worry about the legitimacy of the philosophical enterprise. I have thus tried to lay out a methodology for philosophy which allows us to respond to that challenge by showing how philosophical issues may be addressed in a way that emerges from the sciences, thus allowing our philosophical work to inherit its legitimacy from the scientific theories on which it may be based. I have, in addition, tried to make out the case that the methodology I propose is one which is already present in a good deal of work in a wide range of areas in philosophy. I believe that philosophical theorizing has been slowly emerging, beginning in the mid- to late twentieth century, from the misguided myth of the autonomy of philosophy, a myth which was a crucial part of the founding of the analytic tradition.[3]

References

Alexandrova, A. (2017). *A Philosophy for the Science of Well-Being.* Oxford: Oxford University Press.

Bishop, M. (2015). *The Good Life: Unifying the Philosophy and Psychology of Well-Being*. Oxford; Oxford University Press.

Boyd, R. (1981). Scientific Realism and Naturalistic Epistemology (pp. 613–622). In P. Asquith and R. Giere, (Eds.). *PSA 1980*, vol. 2. East Lansing, MI, Philosophy of Science Association.

Boyd. R. (1988). How to be a Moral Realist (pp. 181–228). In G. Sayre-McCord (Ed.). *Essays on Moral Realism*. Ithaca, NY: Cornell University Press.

Brandom, R. (2009). *Reason in Philosophy: Animating Ideas.* Cambridge, MA: Harvard University Press.

Burge, T. (1979). Individualism and the Mental. *Midwest Studies in Philosophy*, *4*, 73–122.

Burge, T. (1986). Individualism and Psychology. *Philosophical Review*, *95*, 3–45.

Burge, T. (2010). *Origins of Objectivity.* Oxford: Oxford University Press.

Burge, T. (2022). *Perception: First Form of Mind.* Oxford: Oxford University Press.

Chomsky, N. (1957). *Syntactic Structures*. The Hague: Mouton.

[3] I am indebted to two anonymous referees and to members of the audience at the conference on Naturalism in Philosophy in Sofia for helpful comments on an earlier version of this paper.

Chomsky, N. (1959). Review of B.F. Skinner's *Verbal Behavior*. *Language*, *35*, 26–58.

Chomsky, N. (1964). *Aspects of the Theory of Syntax*. Cambridge, MA: MIT Press.

Depaul, M. and Ramsey, W. (Eds.). (1998). *Rethinking Intuition: The Psychology of Intuition and its Role in. Philosophical Inquiry*. Lanham, MD: Rowman and Littlefield.

Dupre, J. (1993). *The Disorder of Things*. Cambridge, MA: Harvard University Press. Feyerabend, P. (1975). *Against Method.* London: Verso.

Fodor, J. (1975). *The Language of Thought*. New York: Thomas Y. Crowell Company.

Fodor, J. (1983). *The Modularity of Mind.* Cambridge, MA: MIT Press.

Frege, G. (1884/1980). *Foundations of Arithmetic: A Logico-Mathematical Enquiry into the Concept of Number*. J. L. Austin translator. Evanston, IL: Northwestern University Press.

Frege, G. (1893/2022). *The Basic Laws of Arithmetic: Exposition of the System*. M. Furth, translator. Berkeley: University of California Press.

Goldman, A. (1986). *Epistemology and Cognition.* Cambridge, MA: Harvard University Press.

Greco, D. (2014). Could KK Be OK? *Journal of Philosophy*, *111*, 169–197.

Haybron, D. (2008). *The Pursuit of Unhappiness: The Elusive Science of Well-Being.* Oxford: Oxford University Press.

Husserl, E. (1900/2001). *Logical Investigations*, vol. 1. London: Routledge.

Husserl, E. (1901/2001). *Logical Investigations*, vol.2. London: Routledge.

Kitcher, P. S. (1983). *The Nature of Mathematical Knowledge*. Oxford: Oxford University Press.

Kitcher, P. S. (1992). The naturalists return. *Philosophical Review*, *101*, 53–114.

Kitcher, P. S. (2003). *In Mendel's Mirror: Philosophical Reflections on Biology*. New York: Oxford University Press.

Kornblith, H. (2002). *Knowledge and its Place in Nature.* Oxford: Oxford University Press.

Kripke, S. (1980). *Naming and Necessity.* Cambridge, MA: Harvard University Press.

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

Kusch, M. (1995). *Psychologism: A Case Study in the Sociology of Philosophical Knowledge.* London: Routledge.

Ladyman, J. and Ross, D. (2009). *Everything Must Go: Metaphysics Naturalized*. New York: Oxford University Press.

Maddy, P. (2007). *Second Philosophy: A Naturalistic Method.* Oxford: Oxford University Press.

Marr, D. (1982). *Vision*. Cambridge, MA: MIT Press.

Maudlin, T. (2011). *Quantum Non-Locality and Relativity: Metaphysical Intimations of Modern Physics.* Malden, MA: Wiley/Blackwell.

Millikan, R. (1984). *Language, Thought, and Other Biological Categories.* Cambridge, MA: MIT Press.

Millikan, R. (1993). *White Queen Psychology and Other Essays for Alice.* Cambridge, MA: MIT Press.

Millikan, R. (2005). *Language: A Biological Model.* Oxford: Oxford University Press.

Millikan, R. (2006). *Varieties of Meaning.* Cambridge, MA: MIT Press.

Millikan, R. (2017). *Beyond Concepts: Unicepts, Language, and Natural Information.* Oxford: Oxford University Press.

Parfit, D. (2011). *On What Matters,* vol. 1. Oxford: Oxford University Press.

Putnam, H. (1973). Explanation and Reference (pp. 199-221). In G. Pearce and P. Maynard (Eds.). *Conceptual Change.* Dordrecht: Reidel.

Putnam, H. (1975). Language and Reality (pp. 272–290). In H. Putnam, *Mind, Language and Reality: Philosophical Papers*, vol. 2. Cambridge: Cambridge University Press.

Railton, P. (1986). Moral Realism. *Philosophical Review, 95*, 163–207.

Skinner, B. F. (1957). *Verbal Behavior*. New York: Appleton Century Crofts.

Sklar, L. (1977). *Space, Time, and Spacetime.* Berkeley, CA: University of California Press.

Sober, E. (1994). *From A Biological Point of View: Essays in Evolutionary Philosophy*. New York: Cambridge University Press.

Sober, E. (2000). *Philosophy of Biology*. Boulder, CO: Westview Press.

Soames, S. (2014). *The Analytic Tradition in Philosophy, Vol 1: The Founding Giants.*

Princeton: Princeton University Press.

Soames, S. (2017). *The Analytic Tradition in Philosophy, Vol 2: A New Vision.* Princeton: Princeton University Press.

Sterelny, K. (2001). *The Evolution of Agency and Other Essays*. New York: Cambridge University Press.

Sterelny, K. (2012). *The Evolved Apprentice*. Cambridge, MA: MIT Press.

Sturgeon, N. (1985). Moral Explanations. In D. Copp and D. Zimmerman (Eds.). *Morality, Reason, and Truth.* Totowa, NJ: Rowman and Allenheld.

Wallace, D. (2014). *The Emergent Multiverse: Quantum Theory According to the Everett Interpretation*. New York: Oxford University Press.

Williamson, T. (2014). How Did We Get Here from There? The Transformation of Analytic Philosophy. *Belgrade Philosophical Annual*, *27*, 7–37.

HUME AND THE PROJECT OF NATURALIZING MODALITY

Daniel Dohrn

Dipartimento di Filosofia
Università degli Studi di Milano, Italy
email: Daniel.Dohrn@unimi.it
ORCID ID: https://orcid.org/0000-0002-0148-3070

Abstract:
The theory of modality deals with objective possibility and impossibility. Modality is an intensely discussed topic in current analytic philosophy. The discussion concerns both epistemology and metaphysics. How do we know modalities, and how are they grounded in reality? Naturalism is also an intensely discussed project in current analytic philosophy. This project gives rise to the subtask of naturalizing modality. Historically, modality has been a topic of interest to David Hume, who is often regarded as one of the forerunners of current naturalism. I aim at bringing together the topics of modality and naturalism in outlining two alternatives for naturalizing modality in the work of Hume.

Key words: Hume, naturalism, modal, possible, necessary.

Hume is often described as a forerunner of modern philosophical naturalism. I want to outline two dimensions in which he also can be described as a forerunner of naturalism with regard to the philosophy of *modality*. Modality (as I shall discuss it) concerns what is possible and impossible. Modality plays a huge role in current philosophical debate. As far as it concerns impossibilities and possibilities that are not actual, it seems especially problematic from a naturalistic perspective.

1. Naturalizing Modality

I shall begin with a brief characterization of naturalism. The project of naturalism may be motivated by the slogan that everything, including ourselves, is part of nature. Most generally, naturalism seems to presuppose privileging some A as natural and to account for some B in terms of A. ‘Accounting for’ stands for several relationships: reduction, elimination, explanation…. If such an account is not feasible, B is qualified as non-natural, and the project of naturalism about B is in jeopardy. Especially salient candidates for A are (the categorical facts about) the actual world, in particular the world of material objects, including ourselves.

In trying to further characterize naturalism, I shall draw further inspiration from Don Garrett’s (2007) seminal work on naturalism in Hume. According to Garrett, naturalism may exclude (i) supersensible entities, such as those that would be divine, immaterial, abstract, or universal; (ii) [entities] that would not be governed by ordinary causal laws, such as miracles, causally undetermined acts of free will, or spontaneous exercises of intellectual insight; and (iii) properties whose character is in principle outside the scope of scientific investigation, such as explanatorily basic intentional or normative properties... (Garrett, 2007, 3).

Filling these variants (i)–(iii) into my general scheme, naturalism would come with a commitment to accounting for the problematic items in terms of some natural A, and be it simply by eliminating them.

Modal properties may fall under all three categories, depending on our other views on them. They may (i) be divine, grounded in God's power of creating different worlds; immaterial in the sense of not being part of the material world; abstract, e.g., in the sense of being grounded in possible worlds, considered as sets of propositions; universal, as e.g., in a universalist metaphysics of properties. (ii) Certain possibilities may be counternomological, not governed by the actual laws. Free will may be taken to figure among the possibilities or to entail a choice between different possibilities. Likewise, non-natural epistemic capacities like intellectual insight may be possible or be postulated in order to explain knowledge of modality. Modal properties are often described as the objects of intellectual insight (Chudnoff, 2011, 330). (iii) Hence, they may be outside the scope of scientific investigation and only accessible to such an insight.

Garrett also distinguishes several versions of naturalism (Garrett, 2006, 301–302):

Doxastic naturalism: belief formation is an operation of nature.
Epistemic naturalism: beliefs can have epistemic authority and be rightful objects of assent in virtue of the ways in which they result from operations of nature.
Explanatory naturalism: phenomena are explained without appeal to anything outside nature.
Metaphysical naturalism: nothing exists outside of nature.

These alternatives further divide up into two pairs. The first concerns the ways we come to believe or know certain things, for instance modal claims to possibility or impossibility. The second concerns the metaphysical source of certain truths, what they are grounded in, for instance the metaphysical grounds of mere possibilities or impossibilities.

As we shall see, Hume's two alternatives for naturalizing modality lend themselves to spelling out any of Garrett's varieties of naturalism. The two Humean alternatives for a theory of modality concern both the *epistemology* and the *metaphysical foundation* of modality. The first is a *projectivism* about modality. Modality does not concern features of the world that are independent of our best modal judgements. Rather, they are projected on the world by our faculties of conceiving/imagining. The second is *skeptical realism*: modality may be grounded in the independent causal powers of real objects and accessible (if at all) by thorough scientific inquiry. While there is a concern that both alternatives are irreconcilable with each other, both are highly relevant in anticipating current naturalistic approaches to modality.

2. Hume's two Alternatives for Naturalizing Modality

2.1. Projectivism

The locus classicus in Hume's writings on how we come to form modal judgements is the following:

> Tis an establish'd maxim in metaphysics, *That whatever the mind clearly conceives includes the idea of possible existence*, or in other words, *that nothing we imagine is absolutely impossible*. We can form the idea of a golden mountain, and from thence conclude that such a mountain may actually exist. We can form no idea of a mountain without a valley, and therefore regard it as impossible. (THN 1.2.2.8, 32)

Here Hume treats conceiving and imagining as equivalent with regard to forming modal belief. I shall follow him in doing so. If one clearly conceives/imagines some object, one thereby conceives/imagines its possible existence. Moreover, whenever one conceives/imagines something, it is not absolutely impossible. One may read 'absolutely' as impossible in the widest sense. For instance, I cannot play the piano, but it is possible in a wider sense (upon proper training) for me to play the piano. Hume's example is our imagination of a golden mountain. There is no actual golden mountain, but we can imagine one, and we conclude from our imagining that a golden mountain could exist. In contrast, when we try to conceive of a mountain without a valley, we conclude from our qualified failure that such a mountain is impossible. From the negative version of the example, we can also derive that Hume accepts a link between our inability of conceiving/imagining something and its impossibility, viz. the necessity of its negation.

Hume seems to subscribe to something akin to the following biconditional:

BICONDITIONAL: It is possible that *p* if and only if we can clearly conceive viz. imagine that *p*.

The biconditional needs qualification to be plausible (see Yablo, 1993, 29). One important refinement is due to Hume himself. For conceiving/imagining to provide access to modal truth, the ideas viz. representations involved should be *adequate*:

> Wherever ideas are adequate representations of objects, the relations, contradictions and agreements of the ideas are all applicable to the objects… The plain consequence is, that whatever *appears* impossible and contradictory upon the comparison of these ideas, must be *really* impossible and contradictory… (THN 1.2.2.1, 29)

Ideas should be adequate to their objects. By adequacy Hume seems to mean a correspondence relationship. Ideas should represent their objects correctly and in relevant detail.

The adequacy requirement is only the beginning. Further idealization has to be added: only *ideal* conceiving/imagining entails possibility (Chalmers, 2002, 171). For instance, we should know all relevant facts–in the case of golden mountains, facts about the material mountains are usually built from, the forces actually acting on them, and so on. The same goes for inconceivability/lack of imaginability. Failures of conceiving/imagining may be due to lack of cognitive resources, ignorance, etc. Such failures do not bear on possibility. I shall not further discuss these refinements.

The key question for my topic of naturalism is how to interpret the direction of fit between conceiving and modal truth: (i) does conceiving/imagining have to track some independent modal truths, or (ii) do these truths in turn reduce to some idealized version of conceiving/imagining?

One may read the adequacy requirement as ruling out the latter alternative. For ideas to be adequate to their objects, the direction of fit must go from the objects to the ideas. Yet this consequence would be too rash. For instance, the direction of fit may go from idealized conceiving/imagining to modal truths, while the idealization conditions may just concern a more limited requirement of adequacy. Adequacy may not include modal features but only non-modal features of things. For instance, the adequate idea of a mountain may include that of a valley, but not yet the modal feature that there can be no mountain without a valley. Ideal conceiving/imagining may have to be adequate to the non-modal facts, but the modal facts in turn may be determined by ideal conceiving/imagining.

While the first (i) alternative leaves us with the task of accounting both for our access to and the metaphysical status of modal truths, the second (ii) gives rise to an amazing prospect of naturalization. The naturalization concerns both the metaphysical grounds of modal truth and our epistemic access to them. The grounds of modal truth lie in some part of nature, namely *ourselves* and our natural endowment, in particular the epistemic capacity of conceiving/imagining. Modal claims are *projected* from the results of our exercise of conceiving/imagining on the world. This also explains how we can have access to these truths. Being projected *from* our conceiving/imagining, they are accessible *by* conceiving/imagining.

Further passages in Hume indicate that he opts for the projectivist alternative:[1]

> This therefore is the essence of necessity. Upon the whole, necessity is something, that exists in the mind, not in objects; nor is it possible for us ever to form the most distant idea of it, consider'd as a quality in bodies. Either we have no idea of necessity, or necessity is nothing but that determination of the thought to pass from causes to effects and from effects to causes, according to their experienced union. (THN 1.3.14.20, 165–166)

At least with regard to necessity, Hume insists that it is not a feature of the independent world. For instance, if there is a mountain, there is also a valley.

[1] For examples of most recent literature on Hume's projectivism: Boehm (2021), Ehli (2024). According to Boehm, projection proceeds by mixing a feeling like pleasure or the felt determination in proceeding from one object to the other with sense impressions so as to create a complex sense impression that already has the structure of an aesthetic or a modal judgment. Such judgments may be true (Boehm, 2021, 832).

According to Ehli, Hume's diagnosis of modal projection is that it is an error committed especially by philosophers. The supposition of a causal connection does not depend on the supposition that it is necessary. The folk feel determined to proceed from one object to the other, but philosophers mistakenly project the non-spatial feeling of determination on extended objects (2024, 190). If Ehli is right, Hume's projectivism goes together with modal skepticism about the correctness of modal claims. It may still count as a naturalization of our attitudes towards modality.

This is a fact of the outside world. Yet there is no such fact corresponding to the claim that there *could* be no mountain without a valley. That claim is due to our inability to imagine a scenario in which there is a mountain but no valley.

As often with Hume, there is an uncertainty about how to interpret the quote. He seems to talk only of the necessity of an effect occurring if its sufficient cause occurs. Yet Hume here also talks of the essence of necessity more generally. Hence there is at least one salient interpretation in which Hume says something about the very category of necessity and not only about the modal status of the connection between causes and effects.

Hume here also suggests something akin to a semantic claim: it is not only the *truth* of claims to necessity that is grounded in our reasoning. The very *meaning* of modal notions like *necessity* is to be analyzed by connections like BICONDITIONAL or our disposition to make a transition from causes to effects. Hume here also anticipates a *semantic* naturalization of modal notions.

If necessity resides only in the mind, the same seems to go for possibility, at least if we accept the equivalence that *necessarily, p* is just the negation of *possibly, not p*. We thus get a general projectivist picture according to which possibility and necessity are grounded in certain ways of reasoning in terms of them.

Hume's projectivism about modality has inspired more recent positions as diverse as Blackburn's (1993, ch. 3 & 5) quasi-realism, Chalmer's (2002, 194) modal rationalism, and the modal conventionalism of authors like Sidelle (1989, 37) and Thomasson (2021, ch. 3). According to Blackburn and Chalmers, possibility reduces to ideal conceivability/imaginability and impossibility to ideal inconceivability/lack of imaginability. According to Sidelle and Thomasson, modal truths just reflect conventional boundaries of things. A striking example is Thomasson's (2021, S2088) analysis of modal terms as object-level statements of semantic conventions. For instance, the claim *necessarily, bachelors are unmarried men* states the semantic rule to apply 'bachelor' only to unmarried men. In Thomasson's theory, there is no sense in asking whether there are any features of the independent world for modal statements to correspond to. The features they correspond to are how we use our words. *Bachelor* is an easy example: our conventions settle that bachelors must be unmarried. More contentious examples are natural kind terms: our conventions settle that gold must have the deep structure identified by best science. To be sure, it is not a matter of course that conventional norms can be naturalized. If they cannot be naturalized, there are principled limits to the project of naturalism. Yet first, Hume also raises prospects of naturalizing norms (Garrett, 2006, 302). From a Humean perspective, a reduction of modality to conventions would therefore not be incompatible with the naturalist project. Second, even if norms cannot be naturalized, at least it is not modality as such that resists naturalization.

While a projectivist outlook promises a smooth naturalization, its main challenge is to tell what the very point of projection is. Why are the results of our exercises of imagining or conceiving so important to us, and why do we talk as if there were modal truths independently of our exercises of imagining? I take this issue to be the main concern about a projectivist naturalization of modality.

There are several strategies in the current literature of answering this challenge. I mention two exemplary ones: Chalmers (2002, section 12) argues that there simply are no candidates for possibilities that outlast ideal conceiving. As a consequence, the boundaries of our conceiving are the boundaries of reality as far as it matters to us. Thomasson (2021, 2087-S2088) argues that object-level statements of semantic rules make it easier to state such rules. For instance, *bachelors can't be married* is easier to digest than *'bachelor' does not apply to married men*. Projection contributes to cognitive economy.

While I shall not take a stance on such strategies, none of them has met with broad agreement. I shall therefore consider a further line of naturalism that takes inspiration from Hume.

2.2. Skeptical Realism

While Hume has been traditionally interpreted as a projectivist about modality, in the 'new Hume' reading there are resources for a very different picture of modality (Wright 1983, 150-155; Strawson, 1989, 201; critical Winkler, 1991, section 8). This picture is inspired by passages like the following, where Hume talks about the danger of a metaphysical determinism that may prove incompatible with our religious beliefs:

> …may it not happen, that, cou'd we penetrate into the intimate nature of bodies, we shou'd clearly see why it was absolutely impossible, they cou'd ever admit of any other Disposition? So dangerous is it to introduce this Idea of Necessity into the present Question! (*Dia-*

> *logues Concerning Natural Religion*, Hume, 1779, 218-219)

In this passage, Hume seems to suggest the following *epistemic* possibility: *perhaps*, if we were able to discern the intrinsic natures of material entities, we would find them to be necessarily the way they are. The material entities at stake may be 'the parts and particles of 18th century natural philosophy' (Winkler 1991, 548). I suggest the following modernized paraphrase of the necessity statement: the world of such entities would be deterministic. It could not be different from the way it actually is.[2]

The picture just sketched is very different from the projectivist picture drawn in the last section. Hume here speaks of a necessity that is grounded not in our minds but in the objects themselves. These objects do not admit of any other disposition. It may be reconciled with the criterion of conceiving/imagining as follows: we do not have adequate ideas of the intrinsic natures of material objects. Such ideas could only be acquired by an inquiry penetrating into these intrinsic natures. If we had such ideas, we could not conceive/imagine their objects to be otherwise, on pain of the ideas being inadequate.

While Hume's point about necessities owing to the intrinsic natures of things can be reconciled with a conceiving/imagining-based epistemology, it cannot be as easily reconciled with projectivism. If independent necessities are there to be detected, they cannot be merely due to projection. Given the problems with reading this passage in light of the projectivist picture, there is the salient alternative that Hume also offers us a very different picture. According to that picture, certain necessities are due to the independent natures of material things. These independent natures cannot be easily accessed by our pre-scientific ways of reasoning. However, Hume suggests they might be accessed by 'anatomizing nature'. What does he mean by the metaphor of an anatomy of nature? I suggest that the answer can be found in Hume's endorsement of experimental science:

> …the essence of the mind being equally unknown to us with that of external bodies, it must be equally impossible to form any notion of its powers and qualities otherwise than from careful and exact experiments… (THN Introduction 8, 21)

Here Hume talks about how to find out the powers and properties of the mind, but the suggestion can be generalized: just as scientifically conducted experiment provides our only way of systematically inquiring into the powers of our own mind, the same goes for the world of independent things. In the quote about anatomizing nature, Hume seems reluctant about guaranteeing access to the intrinsic nature of things, but we may say that, *as far* as there is any option of anatomizing nature, it would have to proceed by scientific experiment.

In sum, Hume raises the prospect of a view according to which the best scientific theory may ultimately settle modal truths by use of experimental data. The view may be called *skeptical realism* because of Hume's reluctance to admit that we are in a position to know modalities understood according to it. It opens up a different direction for naturalism to take: first, an *epistemic naturalism* according to which our best empirical science reveals modal truth; second, a *metaphysical naturalism* according to which modalities are grounded in the intrinsic powers, or to put it in more recent terms, *dispositional essences* of actual things.

Both aspects of Hume's skeptical realism have been pursued in more recent literature on modality. A skepticism about our access to modal truth has been expressed by authors like Robert Nozick (2001, 122): special evolutionarily instilled epistemic capacities would be required to get access to modal truth. Yet since there is no evolutionary advantage of detecting modal truths, there is no reason to assume that our natural capacities provide access to them. The idea that best science (including best logics) is the arbiter of modal truth has been recently revived in the anti-exceptionalist position of Timothy Williamson (2007, ch. 5, 2016) and Bob Fischer's (2016) theory-based modal epistemology. The suggestion of grounding modal truth in the dispositional essences of actual things has been pursued in essentialist (Fine, 2005, 7) and in disposition-based theories of modality (Vetter, 2013).

2.3. Reconciling Projectivism and Skeptical Realism

I have sketched two different approaches to modality in Hume's work, projectivism and skeptical realism. Both seem irreconcilable with each other. Either possibility and necessity are grounded in

[2] I admit that modern characterizations of determinism are more limited: the state of the world at one point in time and the laws determine any other state. I shall put aside this refinement.

our projective tendencies, or they are grounded in the intrinsic natures of independent objects. This duality raises the question whether, assuming the two readings are plausible interpretations, there is an irresolvable inconsistency in Hume's views of modality, or whether there are ways of reconciling the two views contrary to appearances.

I shall make two alternative suggestions. The first is to subordinate skeptical realism to projectivism, the second is to subordinate projectivism to skeptical realism. I begin with the first.[3] The priority of projectivism is supported by deep and general commitments of Hume's like the following:

> …when we say we desire to know the ultimate and operating principle, as something, which resides in the external object, we either contradict ourselves, or talk without a meaning. (THN 1.4.7.5, 267)

In this quote, Hume does not only deny that we can know an external operating principle. Rather he claims that we commit a logical mistake or talk nonsense when we express our desire to know such a principle. However, such an external operating principle seems to be what is intended when Hume talks of the intrinsic natures of things that ground their causal powers.

In this vein, Winkler denies that we can grasp the notion of causation as a mind-independent relation among objects:

> Hume's theory of ideas is relevant to all of our conceptions. And it is a consequence of the theory (joined to his emphasis on our projective tendencies) that we have no conception at all of causation as it is in objects. (Winkler 1991, 573, see Ehli 2024, 179)

Hume seems to deny our very capacity of even thinking about the epistemic possibility that fuels skeptical realism. We cannot even think of external necessities that are grounded in independent natures. The question becomes how to reconcile these claims with the claims supporting skeptical realism.

One suggestion how to answer this question is the following:

> His 'sceptical conclusion' is that we cannot grasp in re necessity, not that there is no necessity. (Kail 2003, 43)

Kail here combines the claim that we cannot grasp necessity as residing in independent objects with an openness for the epistemic possibility that there might be such a necessity. However, this combination seems unstable. One wonders how Hume can entertain the "sceptical conclusion" given that we cannot even understand the idea of *in re necessity*. After all, in talking of such a necessity, we contradict ourselves or talk without a meaning.

Here is my suggestion: we may not be able to grasp *in re necessity*, but we may approach such a grasp by extending the projectivist outlook. What Hume might have in mind in the passages on external necessities may still be projections from what we conceive, but projections that are made from a different position, the position of someone who has performed the scientific 'anatomy of nature'. This position may be inaccessible to us, but we may still be in a position to speculate about what would be projected from such a position. As a consequence, we would not have to endorse necessities as they are grounded in independent objects but only necessities as projected from a different viewpoint with different resources.

This proposal makes modality thoroughly relative to different projective tendencies. Its advantage is that it only requires a minimal revision of the most immediate reading of the passages supporting skeptical realism. Instead of telling us how things might be fundamentally, they tell us only how they appear to someone who has done a certain kind of inquiry.

Coming to the second alternative of subordinating projectivism to skeptical realism, I take inspiration from Trendelenburg's argument against Kant. Kant had argued that the human intellect imposes spatiotemporality on objects as they appear to us. We cannot know these things as they are in themselves. As things are in themselves, they are not spatiotemporally structured. Trendelenburg argued that, since we according to Kant have no access to things as they are in themselves, we cannot tell whether they are spatiotemporally structured. As far as we know, we cannot exclude that they might be so structured (Trendelenburg 1862, 165-166).[4]

[3] The more recent literature on Hume tends to dismiss skeptical realism (Ehli 2024, 169, ann. 4). Yet this does not mean that the passages that seem to point in the direction of skeptical realism cannot serve as a source of inspiration to later naturalists.

[4] "…the mere fact that necessity is mental, or that 'necessity lies in the mind,' is not sufficient to account for why necessity could not belong to the objects." (Boehm 2021, 826)

In a similar vein, we may reinterpret Hume's conceivability-based view not as projecting modal truths but as representing the way we cannot but make modal judgment. We cannot but judge that things could be the way we conceive them, and we cannot but judge that things could not be the way we find inconceivable. Yet this does not mean that the resulting judgments are true. They *might* be false because the fundamental natures of things are so as to render impossible what we find possible and possible what we find impossible. [5]

The resulting reading differs from the one presented in section (2.1.) in avoiding the claim that our conceiving-based judgments are true in virtue of being projected. The weaker claim is that they are unavoidable. This reading goes against the face value of statements like the following: "The plain consequence is, that whatever *appears* impossible and contradictory upon the comparison of these ideas, must be *really* impossible and contradictory…" (THN 1.2.2.1, 29). Moreover, the reading goes against the passages in which Hume insists that we contradict ourselves or talk without a meaning when we talk of modality as it resides in things. The question becomes how we could be in a position to even think of modality of a quality as it is in objects. One reaction would be to be more cautious in formulating skeptical realism: we cannot but project modal claims; whether these projections are reconcilable with the fundamental natures of things we cannot tell.[6]

3. Outlook

Hume is often seen as one of the founding fathers of the modern project of naturalism. Given his interest in modality, it does not come as a surprise that many varieties of naturalism about modality can take inspiration from the alternatives to be found in his work. While it remains to be discussed in how far Hume's own naturalistic stance on modality is coherent, the many motives to be found in his work on modality will continue to inspire further inquiry on modality.

I indicate some lines of inquiry along which the Humean picture can be differentiated. One such line concerns the distinction of several kinds of modality. It is nowadays standard to distinguish epistemic, deontic, and alethic modality. There is also a near-standard distinction of more or less restricted alethic modalities: the most comprehensive alethic possibilities are 'metaphysical' possibilities, those are limited by the laws of nature are nomological ones, those that are restricted by additional contextual restrictions like what is technically feasible may be called circumstantial ones. Another distinction is that of primary and secondary possibility. The former concerns what we can conceive a priori, the latter what we can conceive upon having adequate empirical information. A corresponding distinction on the side of conceiving is that of primary and secondary conceivability (Chalmers 2002, section 3). I have already mentioned ideal conceivability. Ideal conceivability concerns what an ideal subject with unlimited cognitive resources can conceive.

While I do not think that Hume drew clear distinctions corresponding to different kinds of modality, they can be integrated into a picture inspired by Hume. Take the projectivist version (2.1.): a conceiving subject may be thought of as more or less ideal. We get primary metaphysical possibility by considering what can be conceived by disregarding certain discoveries about the identities of things, for instance the identity of Cicero and Tully. We get secondary metaphysical possibility by keeping fixed the identities of things. We get more restricted alethic possibilities by imposing certain constraints, for instance the laws of nature, facts about technical resources and so on. We get epistemic possibility by adding the constraint on our conceiving that the known facts have to be held fixed.

In the skeptical realist version (2.2.), information about different modalities can be accounted for by different stances towards the scientific results on the dispositional essences of things. Epistemic

[5] A similar suggestion has been made by Dohrn (2010, section 3.3): we may limit the range of our conceivability-based modal judgments to claims about everyday objects like mountains and valleys, whereas skeptical realism concerns claims about the natures of more fundamental objects. However, we have to take into account the constraints that can be expected to be imposed by the natures of more fundamental objects on everyday objects.

[6] Ehli (2024) suggests that Hume "…grants that there may be qualities 'with which we are utterly unacquainted' in bodies, and allows that we may use 'power' or 'efficacy' to signify those. However, if we use those terms to signify something of which we have an idea—necessity as mental determination—and attribute that necessity to causally related objects, then 'obscurity and error begin then to take place, and we are led astray by a false philosophy" (T 1.3.14.27; 168).' Ehli's distinction allows us to use 'power' and 'efficacy' to signify qualities in objects which are inaccessible to us. In this vein, we may talk of powers residing in objects as long as we do not claim to have any further grasp of them.

and Chalmersian primary possibility (a priori epistemic possibility) can be obtained by disregarding what has to be figured out by empirical research in the first place. They roughly correspond to Hume's talk of what might be the case given that we do not know the dispositional essences of things. Secondary metaphysical possibility can be obtained by considering what is compatible with the dispositional essences and identities of things. More restricted possibilities can be obtained by adding further constraints.

In sum, there are many ways of differentiating the naturalist image inspired by Hume's original work.

References

Blackburn, S. (1993). *Essays in Quasi-Realism*, Oxford: Oxford UP.

Boehm, M. (2021). Hume's 'projectivism' explained. *Synthese, 199*, 815–833.

Chalmers, D. (2002). Does Conceivability Entail Possibility? In T. S. Gendler & J. Hawthorne (Eds.), *Conceivability and Possibility* (pp. 71–125). Oxford: Clarendon Press.

Chudnoff, E. (2011). The Nature of Intuitive Justification. *Philosophical Studies*, *153*, 313–333.

Dohrn, D. (2010). Hume on Knowledge of Metaphysical Modalities, *Logical Analysis and History of Philosophy, 13*, 38–59.

Ehli, B. (2024). Hume on Modal Projection. *Mind, 133,* https://doi.org/10.1093/mind/fzad020

Fine, K. (2005). *Modality and Tense. Philosophical Papers*. Oxford: Clarendon Press.

Fischer, R.W. (2016). A theory-based epistemology of modality. *Canadian Journal of Philosophy*, *46*, 228–247.

Garrett, D. (2006). Hume's naturalistic theory of representation. *Synthese*, *152*, 301–319.

Garrett, D. (2007). Reasons to act and believe: naturalism and rational justification in Hume's philosophical project. *Philosophical Studies, 132*, 1–16.

Hume, D. (1739–40). *A Treatise of Human Nature*, L.A. Selby-Bigge (Ed.), 1896. Oxford: Clarendon Press.

Hume, D. (1748). *Enquiries concerning Human Understanding and concerning the Principles of Morals*. L. A. Selby-Bigge (Ed.), 3rd ed., 1975. Oxford: Clarendon Press.

Hume, D. (1757/1779). *The Natural History of Religion* [1757] *and Dialogues concerning Natural Religion* [1779]. Colver/Price (Ed.) 1976. Stanford: Stanford UP.

Kail, P. (2003). Conceivability and Modality in Hume. *Hume Studies, 11*, 43–61.

Kail, P. (2007). *Projection and Realism in Hume's Philosophy*. Oxford: Oxford University Press.

Nozick, R. (2001). *Invariances. The Structure of the Objective World*. Cambridge, Mass.: Harvard University Press.

Sidelle, A. (1989). *Necessity, Essence, and Individuation*. Ithaca, NY: Cornell University Press.

Strawson, G. (1989). *The Secret Connexion: Causation, Realism, and David Hume*, Oxford: Clarendon Press.

Thomasson, A. (2021). How can we come to know metaphysical modal truths? *Synthese, 198*, S2077–S2106.

Trendelenburg, F. A. (1862). *Logische Untersuchungen*, Leipzig: S. Hirzel.

Vetter, B. (2013). 'Can' without Possible Worlds: Semantics for Anti-Humeans. *Philosophers Imprint*, *13*, 1–27.

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

Williamson, T. (2016). Modal science. Canadian Journal of P*hilosophy, 46*, 453–492.

Winkler, K. (1991). The New Hume. *The Philosophical Review, 100*, 541–579.

Wright, J. P. (1983). *The Sceptical Realism of David Hume,* Manchester: Manchester University Press.

Yablo, S. (1993). Is Conceivability a Guide to Possibility? *Philosophy and Phenomenological Research, 53*, 1–42.

NORMS OF BELIEF AND NON-PROPOSITIONAL PRIMAL BELIEFS

Madelaine Angelova-Elchinova

Sofia University, Bulgaria
email: madelaine.angelova@gmail.com
ORCID ID: https://orcid.org/0000-0002-5961-0327

Abstract:
Traditional normative theories of belief in epistemology presume that belief-forming includes a reflective component and a mental agency component. Beliefs are regarded as conscious doxastic attitudes with propositional contents. Let's call this view the Transcendental View about Belief (TVB). First, I argue that reputed norms of belief as the truth norm, the knowledge norm and the rationality norm all incorporate TVB. Further, I argue that the empirical evidence concerning belief-forming collected in the last two decades by Rüdiger Seitz, Hans-Ferdinant Angel, Raymond Paloutzian and their colleagues suggests that primal beliefs do not correspond to the description provided by TVB. Primal beliefs are byproducts of perceptive and affective type 1 processing. As such, they are non-propositional multilayer relational states of a cognitive system S that do not presuppose conscious awareness. Thus, I maintain that TVB is normatively inadequate, because it is descriptively false. I presume that a good normative account of beliefs should be able to accommodate primal beliefs and provide a norm of primal belief. I conclude by attempting to outline such a positive account.

Key words: belief-forming, non-propositionalism, norms of belief, epistemic normativity, functional environmentalism.

According to the traditional picture in epistemology, all beliefs are characterised as a) reflective, b) conscious doxastic attitudes towards a proposition (thus as c) having propositional contents. We can call such a widespread view the Transcendental View about Belief (TVB)[1]. TVB presumes propositionalism about *doxastic* attitudes. According to strong forms of propositionalism, all intentional attitudes are propositional (Grzankowski and Montague, 2018, 1). In what follows, I suggest that even if TVB is to remain neutral about strong forms of propositionalism (or to reject them), it would still face an important challenge: to show that all beliefs, disbeliefs and suspensions are propositional.

In the following paper, I argue that TVB is false and that [primal[2]] beliefs are neither a) nor b) nor c). Instead, I show that they should be regarded as multilayer relational states of a cognitive system S with non-propositional representational content. I propose a theoretical framework of primal beliefs built on the empirical data provided by Rüdiger Seitz and Hans-Ferdinant Angel's research (2016, 2020, 2023). The reason for selecting the research of Seitz, Angel, Paloutzian and their colleagues in particular is that – to the best of my knowledge – they are if not the only, then one of the very few teams of neuroscientists and cognitive scientists that has, for more than 10 years now, persistently studied belief-forming specifically. According to the view I am defending, primal beliefs are byproducts of perceptive and affective type 1 processing. As such, they do not presuppose conscious awareness.

There exist other views that propose an argument for the possibility of non-propositional beliefs.

[1] Following Hilary Kornblith's criticism of the transcendental view about knowledge (Kornblith, 2011).
[2] As it is going to become apparent later in my analysis, all beliefs *that* are grounded on primal beliefs.

Following Brentano, Uriah Kriegel is trying to show that beliefs-in are non-propositional in the sense that they are *cognitive objectual attitudes* (Kriegel, 2018, 198). Kriegel also suggests that beliefs-in may turn out to be more fundamental (Ibid, 199). Unlike Kriegel, I neither take non-propositional beliefs to be existential, nor objectual.

I am much more sympathetic to a proposal forwarded by Fred Sommers (2005, 2009) which he calls *mondialism*. To clarify, I agree with Sommers that *de mundo* beliefs are non-propositional beliefs about the world and, in a sense, I even agree that such beliefs are more primitive (Sommers, 2009, 269-270). Unlike him I do not presuppose that such beliefs are taken as factive. Primal beliefs are environmental representations that may be more or less accurate. More importantly, I argue that Sommers was wrong to construe *de mundo* and *de dicto* beliefs as "two very different kind of beliefs" (Ibid, 268). Primal beliefs are more primitive in the etiological sense[3].

Part 1 of the paper is focused on metaepistemology. I examine the three desiderata that an adequate normative account of belief must meet. I draw two important distinctions: one that considers differences in belief-forming, belief maintenance and belief revision and a second between descriptive and normative claims about belief. I conclude by examining how (if at all) normative claims are dependent on descriptive.

Part 2 provides an overview of epistemic accounts about the normativity of belief. Truth norms, Rationality norms and Knowledge norms of belief are presented and examined together with some objections. I argue that even if all other challenges that each account faces are to be resolved, they would still be confronted by one and the same problem: as subspecies of TVB, all these accounts presuppose a mental agency component and a reflective component. Thus, they have no means to account for the normativity of non-propositional beliefs. The challenge stands independently on how exactly one is defining mental agency or reflection, because primal beliefs would simply lack such characteristics. In addition, I allude to an additional challenge about the normative virtue of second-order beliefs that has been presumed by TVB.

Part 3 provides an argument in favour of non-propositionalism about belief. I examine empirical data of primal belief-forming and conclude that a normative account of beliefs should be applicable to primal beliefs. Further, I argue that if this is the case, then we are in request for a very different type of account that rejects the Transcendental View altogether.

In the final part 4 I provide some clues about how a positive account of primal beliefs should look. I conclude by suggesting that such an account should have a proper function component as well as an environmental component and call this proposal *Functional Environmentalism*.

1. Metaepistemological Considerations

A good normative account of belief-forming[4] should satisfy three simple conditions:

1.1. Desiderata

Consider *Jurassica* – a hypothetical environment inhabited by the magnificent creatures of the Jurassic period. Imagine a group of Xiaosauruses (small herbivore dinosaurs) hiding from a large predatorial dinosaur. Hearing the footsteps of the predator, one of them formed a belief that it is time to find a good hiding place. Based on its testimony the other members of the herd[5] gained knowledge of the upcoming danger. All of them also knew quite well that if they were to make a sound while hiding, they were most certainly going to alert the predator to their location.

Epistemic phenomena like forming beliefs, gathering evidence, accommodating testimony, making predictions and especially knowledge possession are not restrictively human phenomena. It is quite probable that there are interesting normative questions concerning specifically human practices[6]. However, I consider the ones related to belief-forming to belong to a broader category. Some bad

[3] I realise that this is a radical claim. Later, when I discuss belief-forming, maintenance and revision, I will provide an argument to motivate it and examine possible objections to it.

[4] A distinction between belief-forming, maintenance and revision would be proposed later on. For the purposes of the current paper, I am only interested in normative accounts of belief-forming. I consider it an open question if there is a place for (distinct) normative accounts of belief maintenance and revision.

[5] This speculative story is nevertheless based on evidence for herd-living being presented amongst herbivore dinosaurs in the Jurassic (see Pol et al. 2021).

[6] To tell if that is the case, we should dive deeper into the metaphysical question about the nature of human and non-human animal belief.

cat beliefs will also be considered bad human beliefs and maybe *for the same reasons*.

Thus, the first condition of the desiderata is that we should look for a normative account of belief-forming that explains *shared mechanisms of human and non-human animal belief*. By 'shared mechanisms', I simply mean the ways in which a belief is being acquired and I presuppose that belief-forming mechanisms can function more or less optimally. And while I do think that the same *type* of mechanisms is responsible for human and non-human animal belief-forming (as it is going to become clear in 4.), of course that does not require any exact similarities between *token* mechanisms[7].

The second condition of the desiderata is a *moderate form of content naturalism*. It is inspired by Nina Emery's examination of normative theories in her *Naturalism Beyond the Limits of Science* (2023) and defined as follows:

> **Content naturalism** [states that] we should not accept metaphysical theories that conflict with the content of our best scientific theories (Emery, 2023: 23).

It is important to stress that I am not necessarily endorsing a radical version of content naturalism[8], nor do I need to for the purpose of this paper. For now, I am simply assuming that Emery may be right and that there may be metaphysical content that goes beyond the empirical data. What matters for my account of belief-forming is that a good metaphysical explanation of how beliefs come to be (and what they are) should not contradict our best scientific theories about the nature of beliefs and should take into account the empirical data that we possess regarding this question. In other words: I hold that while the metaphysics of belief[9] should surely not be decided once and for all by the neuropsychology and neurophysiology of belief, it should nevertheless proceed in an empirically informed way.

The final condition of the desiderata is that a normative account of belief-forming should *presume doxastic involuntarism*: we do not choose which beliefs we form[10]. A similar consideration about normative accounts of belief has been motivated recently in great detail by Matthew Chrisman (2022) and Neil Levy (2021). Albeit very different from one another, both of their accounts of belief are developed in the framework of the dual-process theory of cognition. As related to type 1 processing, belief-forming cannot be the product of direct doxastic control and is heavily influenced by one's epistemic environment (see e.g. Levy, 2021: 29 and Chrisman, 2022: 7, 21). I agree with all of that but unlike Chrisman and Levy my account of belief is individual, not social[11]. I also accept the dual-process framework of cognition but I prefer to talk about two types of information-processing and not about two different systems *per se*[12].

Apart from the desiderata, one also needs to be aware of two distinctions I will presuppose to be correct in this paper. The first one differentiates between:

1.2. Stages of Believing

My argument is premised on the assumption that we can talk about three stages of believing: belief-forming, belief maintenance and belief revision. As I already pointed out, in this paper I am only going to be interested in belief-forming or belief acquisition, but for the sake of clarity I will draw an explicit distinction between the three stages.

Belief-forming: I sustain that every propositional belief we hold was initially a non-propositional, primal belief. Thus, when we form beliefs, we form primal beliefs. Animals and babies are just as capable of forming such beliefs as they

[7] E.g. think of a toad's visual system and a human's visual system. The similarity concerns the modus of delivering information about the environment. The actual hard-wiring is pretty different, but maybe we can introduce the same accuracy conditions regarding toad perception and human perception.

[8] Radical versions of content naturalism would insist on excluding any metaphysical content.

[9] One can wonder why a content-naturalism condition on the metaphysics of belief should concern the normative account of belief. Should not the metaphysics of belief depend on the normative account of belief and not the other way around? I would like to thank David deBruijn for suggesting this objection and I will address it in more detail in 1.3.

[10] That leaves open the possibility that doxastic involuntarism is false when it comes to belief-revision.

[11] I am much more interested in the role of perception for belief forming, rather than on agency, social cooperation or communication, thus sticking to traditional epistemology. I will not have the chance to examine the relations between perception and belief here, but I hope to do this in a forthcoming broader research project.

[12] When one endorses a dual-process theory of cognition, that invites the question "Two types of what have been distinguished?" See Evans, 2009 for a unified theory suggestion; see Bellini-Leite, 2018 for reasons to abandon the 'two systems' assumptions and what follows from it. Here, I go with Bellini-Leite, but nothing in my analysis will depend on that.

are by-products of perceptive and affective processing. Casting primal beliefs in this light is entirely compatible with beliefs being products of type 1 information processing and allows us to maintain doxastic involuntarism. Therefore, I consider that one of the problems of normative theories about belief in epistemology is that when they turn to belief acquisition, they actually examine a much later stage of believing where a belief has been updated multiple times and is way more refined[13].

Belief maintenance: In this paper, I neither propose nor defend a full-blown account of belief-maintenance. When I talk about a stage of 'maintaining' one's belief, I only assume that a) maintained beliefs were initially formed as primal, non-propositional ones[14] and b) during their maintenance they do not need to be actively presented in cognition. In this sense, belief maintenance involves nothing more than sustaining 'tacit' beliefs in Lycan's sense because it does not require any active reflection (Lycan, 1988). I depart from Chrisman's terminology where "maintenance" is used to refer to the "*belief*-system maintenance (as opposed to some other mental activity)" and such exercise involves agency ("believers are in charge of their beliefs") by having an active and holistic nature (Chrisman, 2022: 69–70).

Very loosely, I think of the process of belief-maintenance as largely automatic and involving the sustaining of a representational content in memory that is available for retrieval[15].

Belief revision: The final stage or – more precisely – stages that a belief may go through[16] is revision. I think that when epistemologists talk about beliefs and norms of belief they are primarily concerned with that stage of the process of believing. Revisions involve both type 1 as well as type 2 processing. During the stage of revision, a belief is necessarily occurrent and it is being updated based on new evidence. This presupposes that this is the stage when one is becoming conscious of her belief and its content becomes propositional.

The most important question is whether we can acquire a new belief during the stage of revision. In other words, can we say that this is the stage in which one acquires (forms) her propositional beliefs? I think that a positive answer to that question has been implicitly presumed by normative accounts in epistemology. Specific examples of normative accounts my challenge applies to are found in 2. However, the claim is that it actually applies to almost every available account on the market. Furthermore, I am going to show that if TVB holds, beliefs *are* actually only the newly acquired contents during the revision process. I argue against that by holding that *no new beliefs are acquired* during the revision stage and that norms of belief should not be focused on propositional beliefs. While non-propositional representational contents do become propositional and conscious in this stage, nothing new emerges in the process.

But what about updating based on new evidence? Surely, it should change the contents in a certain way and one may form new propositional beliefs in the process. I start by answering the first part of the objection: new evidence serves as a trigger for turning a dormant sustained belief into an occurrent propositional one. While it has been 'translated' into a conceptual content, it is still the same belief carrying the exact same information. New evidence can make our credence in it stronger or weaker[17], we can even drop the belief altogether. Turning to the second part of the objection, it is possible that one can form new second-order beliefs that may count as new propositional beliefs. While considering new evidence about Bobby cheating on his Epistemology exam, I may acquire the belief that I firmly believe that Bobby cheated. In other words, during revision I may form new beliefs about my own level of credence, for example. What I consider unlikely is that it is possible that I form any new beliefs about the external environment solely by reflecting on the contents of already existing primal beliefs.

Here, I am obliged to face a strong objection that I will call gaining knowledge through reflec-

[13] More about that will be said in 2. where I offer a detailed examination of norms of belief.

[14] It is interesting to see if during the stage of maintenance beliefs are still non-propositional. I make no claim one way or the other, but I think it sounds plausible that they may be.

[15] That is also compatible with J. Adam Carter's notion of "dispositional belief" understood as representational content stored in memory that is not necessarily (occurrently) consciously endorsed (Carter, 2022: 2,7). In Carter's case this representational content is propositional, while on my account it can turn out to be non-propositional.

[16] I say 'may' because I neither consider that it is necessary that every primal belief is maintained afterwards nor that it is revised.

[17] I am aware that such an account departs significantly from the traditional doxastic attitude framework by presupposing that the belief is not the attitude (towards a proposition) itself.

tion[18] or K through R. A classic example of K through R is found in reasoning in detective stories. Let's imagine a scenario where a detective is trying to find out who Mary's killer is. The suspects are Mary's husband and her landlord. Further, let's also presume that Mary was killed from a great distance with a blowgun using poisonous darts. From this point on, a supporter of the traditional view may construe the case as follows: initially, the detective believed that the husband is responsible for the murder because he had a very strong motive (unlike the other suspect who seemed to have none). By doing a fairly simple test, the detective found out that whomever killed Mary should have been not only a good shooter, but also should have had extremely well-functioning lungs in order to blow the darts so far away and with such precision. Next, by presuming that the detective knows that the husband is a smoker and has a bad case of asthma, the defender of the traditional view can argue that the detective can infer that Mary's husband could not have possibly killed her, gaining not only a new belief but also K through R in the process.

While K through R challenges are indeed important, the possibility of gaining *knowledge* through reflection remains open even if I am right about everything that I claim in this paper. More importantly, I think that the particular detective case is unable to show that one can *form new beliefs* merely by reflection. At the very beginning of the examination of the case, the detective simply has a non-propositional belief about the possible killers. I say "non-propositional", as I refer to the very first moment when the primal belief is formed (before any type 2 processing has taken place). After revising this belief based on the evidence concerning the husband's motive for murder, the detective becomes much more certain that the husband is responsible for the crime. When obtaining the evidence about the lung condition of the murderer, he simply revised the belief that the husband is guilty once more, this time dropping the belief altogether. The non-propositional representational environmental component – the initial content of the primal belief – has not changed throughout the entire process. It is the very same primal belief that has been revised over and over again by shifting the focus on a different relational aspect or rejecting its presence altogether.

[18] I am especially indebted to Jack Lyons for making it stronger by using the Sherlock Holmes example. I will adapt the case in point to fit any detective story out there.

A final distinction I draw is between descriptive claims about the nature of beliefs and normative claims about beliefs. We can answer the question of what beliefs are descriptively if we look into empirical data about belief-forming. It is possible that we may have to go beyond empirical data and engage in metaphysical considerations in order to get the complete picture. Normative claims about belief are not traditionally concerned with what beliefs in fact are, but with what one should believe or disbelieve and why.

1.3. The Descriptive/Normative Dependence Relation

A defender of TVB can proceed by suggesting that normative claims about belief should determine the nature of belief proper. If that is about right, the following normative dependence relation should hold:

NDR: Norms about a type determine if something should be considered a token of said type. The norm should be transcendental in the sense that it precedes any realisation of the norm and that it is a priori justified. If TVB and NDR hold, it could be objected to my project that empirical data about belief-forming should be regarded as irrelevant when dealing with normative questions about belief. I would like to try to show why NDR is false.

Let's take a simple example and suppose for the sake of the argument that there was indeed a normative condition preceding any pizza-token that something is a pizza if and only if it is a round pastry with a sauce on it baked at a very high temperature. Any pastry that satisfies the norm is therefore a pizza. But we should not forget that constitutive (or definitional) norms for which NDR provides the most general model allow us to make evaluative assessments based on if something satisfies the constitutive norm or not[19]. Thus, tokens that do satisfy a type-norm can also be evaluated as *good* tokens. In our example, pastry tokens that satisfy the transcendental norms of pizza-making are not just pizzas, they are *good* pizzas. But if something does not satisfy the norm it is simply not a pizza, so it turns out that each and every pizza is a good pizza. In other words: if NDR is true and normative claims about

[19] I would like to thank an anonymous reviewer for stressing the need to distinguish between constitutive and evaluative norms of belief. I am also grateful to Erwan Lamy for showing me the way to draw the relation between the two types of norms.

pizzas or beliefs determine what pizzas or beliefs are, then bad pizzas or beliefs should be impossible (because normatively impermissible). They will simply fail to be tokens of the relevant type.

Can a defender of NDR simply bite the bullet by saying that if a pastry goes against the transcendental norms of pizza-making it is simply not a pizza? Such a response is not only possible, but in the pizza case it may sound plausible as well. However, the case is different if NDR is applied to beliefs. To say that bad beliefs should not be regarded as a token of the belief type is a non-starter. This result does not square well with ascribing epistemic blame for bad belief-forming and for sustaining bad beliefs. We still want a sexist to be regarded as (really) believing that women are incapable of exceeding academically in order to hold them responsible for having such beliefs. If it turns out that a sexist does not have *any belief* of the sort, because they fail the normative condition, that would not allow us to blame people for having bad beliefs[20].

Therefore, I conclude that NDR is false and I proceed by suggesting that the opposite descriptive dependence relation holds:

DDR: Normative claims about beliefs are constitutively dependent on descriptive claims about the nature of belief that we endorse.

According to DDR, if our descriptive views about the nature of beliefs are wrong, then it is very probable that our normative claims will also turn out to be wrong[21]. Even if a normative theory of belief grounded on a wrong metaphysical view turns out true, that would be merely felicitous. Let's turn back to the pizza example and imagine that someone has baked a really good pizza without knowing what a pizza is[22]. It is by sheer luck that Smith followed all the steps for making a good pizza, because he had no idea what he was making in the first place.

Turning to beliefs, when I say that norms of belief are constitutively dependent on the nature of beliefs, I mean that different descriptions of belief presuppose a different normative theory of belief. For example, if believing is *a performance*, that would give us one set of norms about how to perform skilfully; if beliefs are *mental states* with propositional contents, that would probably give us a truth norm and so on and so forth. Moving forward, my argument in this paper suggests that TVB and standard normative accounts of belief associated with it are misguided because they are wrong about what beliefs are. They are false as normative claims because they are descriptively false[23].

2. Transcendentalism about Belief

This part of the paper is devoted to providing a clear outline of the traditional normative accounts of belief in epistemology and to propose an overview of existing norms of belief developed in this framework. The label "Transcendental View about Beliefs" (TVB) and parts of my criticism have been inspired and influenced by a project of Hilary Kornblith aiming to show that Transcendental Views about Knowledge (TVK) are false[24]. Transcendentalism is related (in a variety of ways) to the Kantian approach to human cognition (Kornblith, 2011: 116) and tells a story about the importance of second-order beliefs and the forming of beliefs in the realm of reasons:

> Human beings differ from other animals in a number of fundamental ways. First, human beings are language users, while other animals are not. And second, leaving aside, for a moment, the question of whether nonhuman ani-

[20] Maybe we can still blame them for failing the normative condition in the first place even if we presume that they lack any beliefs about their female colleagues. We can say for example that the sexist *should have formed* a belief and they *failed to do so*. That may work, but it is not what we usually consider to be going on in these cases.

[21] Note that unlike NDR, DDR does not tell you what positive epistemic norms of belief you should adopt; it only gives you some clues about what kind of epistemic norms would be incompatible with the nature of belief.

[22] We can imagine Smith just putting something together in the kitchen without having any idea what exactly he is doing and while thinking that pizza is actually a type of cake.

[23] It is not a precedent to argue against a normative claim about belief from a descriptive one. To give just one example, in *Belief is Weak* John Hawthorne, Daniel Rothschild and Levi Spectre argue against the assumption of entitlement equality and, consequently, against the knowledge norm of belief because they do not "correspond to our basic concept of belief" (2016: 1395) or that there is a "radical mismatch" between such normative conditions and the literal meaning of belief (2016: 1397).

[24] Kornblith himself refers to transcendentalist accounts of knowledge in philosophy as "anti-naturalistic" and argues against them by raising the claim that knowledge is a natural kind (2011: 98). I take no stance in the debate about the nature of knowledge and remain silent on the question if knowledge is indeed a natural kind. My only focus in this paper will be anti-naturalism about belief-forming.

> mals have genuine beliefs, it is quite clear that non-human animals are incapable of reflecting on their mental states and forming beliefs about them. Thus, human beings not only form beliefs about the world around them; they form beliefs about their own beliefs. Nonhuman animals cannot do this. This ability to reflect on our own mental states marks a crucial difference between humans and other animals. (Kornblith, 2011: 99)

On TVB, believing is about what we do in our heads about evidence and how we do it. When forming a belief, one forms a doxastic attitude towards a proposition that is supported by one's evidence. Let's call this *a mental agency component*. As we can see, a fundamental characteristic of transcendentalist accounts is that "rationality" is about the second-order level. The mental agency component is grounded on a *reflective component* as TVB presupposes that second-order beliefs are fundamental filters for the forming of first-order beliefs and that second-order beliefs improve first-order beliefs[25]. Thus, TVB can be summarised as follows:

TVB: Beliefs are a) reflective, b) conscious doxastic attitudes towards a proposition and c) the content of any belief is thus propositional.

In what follows, I examine some influential norms of belief and show that they incorporate TVB at their heart, starting with:

2.1. The Truth Norm of Belief

One can sustain that beliefs should be aiming at truth. There are at least two construals of the truth norm. They both fall under the TVB premise, but there are some differences that are worth mentioning. I will start with what I call the weak construal that constitutes maybe the first official construal of a belief norm that (to the best of my knowledge) started this whole debate about norms of belief in epistemology.

Truth Norm (Weak): One must: believe that *p* iff one takes *p* to be true (Williams, 1973)

First, it is worth noting that on the Weak interpretation, the Truth Norm does not require that one only forms true beliefs. It only requires that one only forms such beliefs that one "takes" to be true, i.e. that the contents of said beliefs appear to be true or seem to be true[26]. Thus, the Weak Truth Norm requires that every time one forms a belief, one has already formed a second-order belief that the first-order belief is true. I think it is obvious that the Weak construal requires us to accept TVB as its premise as it suggests that the reflective component is present in belief-forming.

Second, on a different interpretation of Weak, taking *p* to be true involves taking *p* as a practical reason for action. Thus, I believe that *p* for epistemic reasons, but at the moment when I start to take *p* as true, my belief becomes a reason for action. Even if this interpretation can be twisted around a little bit so it does not *prima facie* require that when forming the belief that *p* one does it in accordance with the reflective component[27], it would still require that belief-forming involves a mental agency component. Therefore, as the mental agency component is based on the reflective component, it remains a mystery how one can shape belief-forming as involving one and not the other.

Maybe it should already be obvious that any stronger version of the Truth Norm would also fall under TVB, but the stronger construal also invites a different problem.

Strong: One must: believe that *p* iff *p* is true. (Shah, 2003).

Note that Strong does not simply say that one should not believe what one has evidence to think is false. It goes way over and above by claiming that one should not form *any* false beliefs *ever*. When presenting this construal of the Truth Norm, Mona Simion, Christoph Kelp and Harmen Ghijsen refer to it as NFB, or ascribing negative status to any false belief, and argue that there is tension between NFB and radical sceptical scenarios (Simion, Kelp and

[25] One can argue that another reason to question TVB is that the reflective component takes it for granted that second-order processing improves first-order processing and that is not always the case (for an analysis of empirical data on metacognition see especially Kornblith 2012). It is an open *descriptive* question how second-order beliefs influence and change first-order beliefs. For starters, empirical evidence of ideological commitments influencing belief-forming as well as a large number of experiments devoted to studying cognitive penetration should be taken into account when accessing the pros and cons of second-order processing.

[26] Traditional internalists about evidence for example would consider that evidence consists of true propositions of one's own mental life (Littlejohn and Dutant, 2021: 802). On such a normative account of beliefs, a true belief that *p* and a false belief that *p* would be equally justified because equally supported by evidence.

[27] For example, an account like this can avoid stating that beliefs are always formed consciously.

Ghijsen, 2016). If we want to allow for justified false beliefs in cases of philosophical hallucination, Strong would be a non-starter.

Consequently, if we hold Strong then belief-forming should not only answer to the reflective component and the mental-agency component, it should be done in a way such that all beliefs are justified and all beliefs are knowledge. Thus, such a construal of the Truth Norm would not be much different from:

2.2. The Knowledge Norm of Belief

The knowledge norm of belief is primarily forwarded by knowledge-firsters in epistemology, but not all knowledge norms are alike. For example, in her *Resistance to Evidence*, Mona Simion explicitly breaks with the traditional construal of the knowledge norm and allows for justified false beliefs in a knowledge-first framework (Simion, 2024). At the same time, the norm of belief that she proposes is definitely *a* knowledge norm and it can be argued that it does not require that one accepts TVB[28]. When I address *the* knowledge norm of belief, I exclude Simion's proper functionalism from my criticism. As it will become clear in 4., I think that the right way to go about shaping a normative account of primal beliefs is pretty close to Simion's functionalism.

The knowledge norm was first introduced by Timothy Williamson as somewhat equivocal to the norm of assertion he upholds (Williamson, 2000: 208; 256-257)[29].

Knowledge Norm: One must: believe that p iff one knows that p (formulation proposed in Simion, Kelp and Ghijsen, 2016; norm sustained by Williamson, 2000; Littlejohn, 2013, 2017; Smithies, 2011; Gibbons, 2013).

As part of the knowledge-first program, evidence and belief are defined in terms of knowledge. Evidence is presumed to be factive and to obey the E=K principle (following Williamson, 2000).

It is very important to see the way supporters of the knowledge norm fall prey to TVB. Factivist accounts and especially versions of factivism that endorse the knowledge norm are all supporting an externalist theory of justification. Being *factive*, the conditions for a belief being a good belief (or justified belief) are objective as opposed to subjectivism about justification[30]. One can wonder if that would not take us further away from the reflective component and bring to the table a new way to talk about normativity of belief. As an externalist, I firmly believe that objectivism gave us a new and a better way to think about justification. The problem is that when it comes to belief-forming, the knowledge norm is not *that different* from other proposals.

What Williamson and Littlejohn's objectivism is giving us is the following:

> If (OC) is correct, it is possible for pairs of thinkers to process things in just the same way, draw the very same conclusions using the same reasoning after things appear precisely the same to them, and still form beliefs that differ in justificatory status. This suggests that a kind of rational perfection might fall short of ensuring that our beliefs are justified. (Littlejohn, 2017: 144)

The difference between objectivism and subjectivism seems to be that objectivism *adds* an objective, external condition and states that rational perfection is good, but it is not enough[31]. Thus, in a certain sense objectivism is simply more demanding by excluding cases of bad luck and nasty environment from the knowledge department (even if the believer does nothing wrong[32]). Littlejohn even proposes a transcendental argument to support his objectivism (2017: 146–147). Norms of belief are objective in the sense that when forming a belief one *has a duty* to consider one's evidence or what one knows (invites the reflective component of TVB) but

[28] Simion's normative account stems from the etiological theory of functions and identifies the epistemic function of our belief-forming system as producing good beliefs (i.e., knowledge) and the proper functioning of the system as determining the norm of belief (2024: 99–101). To be precise, on Simion's account, beliefs are described as moves in inquiry and, because the function of inquiry is to deliver knowledge, belief-forming *in* inquiry inherits this function and the associated norm (Ibid.: 101–102). I suspect that primal beliefs are not formed as part of the inquiry practice and are formed *outside* of any such practice.

[29] As already mentioned, that also raises the problem about Entitlement Equality (see Hawthorne, Rothschild & Spectre, 2016) that I will not be considering here.

[30] See Littlejohn, 2017: 143–146 for an outline of the opposition between subjectivism and objectivism about justification.

[31] Other forms of externalism about justification may still be quite different in this regard.

[32] In radical sceptical scenarios, for example, believers would be "excused" and would not be "blameworthy" but they still fall short of justification, i.e. their beliefs are still *bad* and they *should not have formed them*.

also that the right norm of belief requires identifying standards that should guide belief "even when not seen perfectly by an agent" (Ibid.). Even if my failure to recognize the evidence cannot be ascribed to a mistake on my part, but stems from a bad epistemic environment, my second-order assessment fails the norm.

To summarise, the knowledge norm makes one more prone to failure, because it adds an objective component to the norm of belief. Although it is an externalist norm, it requires that one endorses TVB.

2.3. Rationality Norm

Unlike the truth norm or the knowledge norm, rationality norms are maybe best depicted as a family of norms. However, I think that each supporter of a rationality norm can agree that something like the following normative condition should hold:

Rationality of Belief: One must: believe that p iff it is epistemically rational to believe that p (Hughes, 2019).

The difference between accounts of the rationality norm would probably stem from debates about what should be considered as 'rational'. Two examples of such disagreements are if it can be rational to sustain a belief solely on the basis of practical reasons (Comesaña, 2015; Rinard, 2019) and if it can be rational for any agent ever to be epistemically acratic (Horowitz, 2013; Christensen, 2024).

In both debates, appeals to the reflective component and the mental agency component of TVB are quite common and are usually part of the description of one's rationality[33]. For example, Sophie Horowitz's non-acratic constraint is supported by the concept of higher-order evidence or the evidence about what one's evidence supports[34] (Horowitz, 2013: 718–719). Comesaña's criticism of pragmatism is entirely based on the following assumption: even if one can (unbeknownst to them) believe for practical reasons, it is never rational (or maybe not even possible) to believe due to what one *takes* to be a practical reason (Comesaña, 2015: 190–191). Thus, the real question in the debate should be: If there is a second-order belief about one's evidence, is it possible to believe for practical reasons?

Debates surrounding the rationality norm provide the clearest indication that contemporary epistemology only cares about the normativity of beliefs as presented by TVB. In the remaining pages I aim to show that TVB is false and that we should reconsider what normative constraints we should impose on beliefs. What I propose is a naturalist account of belief. I argue that if empirical data can show that beliefs are not what epistemologists suppose they are, i.e. if TVB is descriptively false, then we should not presume that it is normatively right. If during the stage of belief-forming all beliefs are non-propositional, then none of the TVB inspired norms is able to account for the normativity of belief. I also aim to show that a naturalist account does not have to remain merely descriptive and can propose a viable normative constraint on beliefs that is more inclusive than the abovementioned manifestations of TVB.

3. Why Primal Beliefs are Non-propositional

In accordance with the content naturalism condition of the desiderata, in this part of the paper I examine the empirical data about belief-forming that is provided by neuropsychology and cognitive neuroscience. Surprisingly enough, the process of believing and the nature of belief have been explicitly examined by very few research teams. Notably however, beliefs have been the central focus of Rüdiger Seitz, Hans-Ferdinant Angel and Raymond Paloutzian's research for the last two decades. I suggest that empirical data provided by their teams is in conflict with TVB. The scope of this paper does not permit for an extensive analysis of the empirical results, but I will try to provide a naturalistic theory of primal beliefs and the process of belief-forming that is founded on the data provided by three of Seitz and Angel's studies (2016, 2020, 2023).

[33] It has been brought to my attention that one can conceive of a very minimalist concept of rationality that does not involve either a mental agency component or a reflective component. I have been asked if I think that a Rationality Norm construed in accordance with such a concept will be applicable to primal beliefs. Maybe that is possible. It remains to see if such a minimalist understanding would be able to deliver a viable epistemic norm of belief. An anonymous reviewer has also suggested that Gerd Gigerenzer's concept of ecological rationality (see Gigerenzer 2000: 57 for definition) is compatible with primal beliefs. That is true as Gigerenzer's concept is also an environmentalist one. However, it is not an *epistemic* concept of rationality and it does not go hand in hand with any epistemic norm of belief.

[34] Horwitz's criticism of level-splitting accounts that proclaim in support of the rationality of akrasia in certain cases is all about the second-order beliefs that detective should have about his epistemic state and his first-order belief (Ibid.: 725-726).

3.1. Primal Beliefs

Let's start by introducing a definition of what I am going to call 'primal beliefs' that is mostly influenced by Sachs and Hirsch (2008), Seitz and Angel (2020) and Seitz et al. (2023). I borrow the notion of primal belief from Seitz and Angel (2020) where primal or proto-beliefs are not explicitly defined. It is only pointed out that they are not conceptual or language-bound and correspond to expressions like "to believe that" and "to believe someone" (Seitz & Angel, 2020: 2,4). If a more precise and philosophically informed definition[35] is to be given, it should sound like this:

Primal Beliefs: Multilayer relational states of a cognitive system S that do not presuppose conscious awareness nor include propositional content, and are by-products of perceptive and affective processing.

Some clarifications are in order. Starting from the last part of the definition, in a 2016 study, Seitz, Paloutzian and Angel showed that perception and tacit emotional valuation are the basis for any primal belief (Seitz, Paloutzian and Angel, 2016). Accordingly, belief formation can be regarded *without referring to the notions of truth, knowledge and rationality* (emphasis added) (see Angel and Seitz, 2016).

Later, their results also point out that action or prediction based on primal belief predates awareness (e.g., Seitz, Angel and Paloutzian, 2023: fig.2). Primal beliefs are formed during perceptual exposure to concrete objects and they can be described as "probabilistic perceptual representations" of, say, the objects's shape or heaviness or location (Seitz & Angel, 2020: 2).

If Seitz, Angel and Paloutzian's assessment of primal beliefs is about right, such beliefs are non-propositional representational states whose acquisition does not require any conscious awareness. In other words, empirical data about belief-forming and primal beliefs brings bad news about TVB. To remind the reader, according to TVB beliefs are a) reflective, b) conscious doxastic attitudes towards a proposition and c) the content of any belief is thus propositional. One thing should be clear by now: not *all* beliefs are like that. It remains to be seen if *any* beliefs *are* like that.

[35] Empirical data can be interpreted in a variety of ways. When it comes to matters of the mind, precision of analysis and interpretation is just as much (maybe even more) a task for philosophy as it is for hard science.

3.2. Belief-forming

Going beyond the restrictions of TVB and formulating an empirically-informed and descriptively-accurate normative account of belief-forming which would be able to accommodate primal beliefs is an important philosophical task. First, if (following Seitz and Angel's distinction) empirical beliefs *that* are generally found to be by-products of perceptive and affective processing, i.e. to be primal beliefs, then existing accounts of belief-forming and acquisition are unable to account for empirical beliefs *that*. Even if one is to disagree with my more radical claim that all revised beliefs are grounded on primal beliefs, one is bound to agree that many of our beliefs are empirical beliefs *that* and that we need an account which can be applied to such beliefs.

It is an open question how to proceed and what kind of definition we should give to belief-forming in this more inclusive sense that parts ways with TVB. I propose one that allows for a naturalistic approach to belief and answers the conditions stated in the desiderata. Moreover, my proposal is motivated by the role of perception and the environmental component, which – if Seitz and Angel are right[36] – are extremely important for accurate belief-forming:

Belief Forming: A unifying concept signifying all bodily acts and environmental relations involved in the forming of primal beliefs.

Thus, I consider belief-forming as an interaction between a cognitive system[37] and its environment. In this interaction, a non-propositional representation of one's external environment has been acquired. This representation can be more or less accurate. The belief-forming function of the cognitive system has evolved in a pre-linguistic fashion and acquired

[36] In one of their papers they sustain that "belief formation seems to reflect to the individual subject the meaning of his/her past and future interactions with the environment" and that the function of belief is to aid problem-solving and manoeuvring in one's environment (Seitz and Angel, 2020: 2, 5). As it is going to become apparent in 4., the environmental component will be crucial for the positive suggestion I have to offer as my normative account would involve a proper function condition, as well as an environmental condition.

[37] Another interesting challenge would be to pinpoint how far down the evolutionary tree this interaction can be stated to go. Would it be that every organism capable of perceptive processing can form beliefs? I admit that this is an interesting question that I have to turn to someday. However, for now I can just say that it seems obvious that primates, crows, dogs, cats, octopuses and a variety of other life forms have primal beliefs.

representations can be stored in memory for retrieval, which is compatible with the claim that primal beliefs may be non-occurrent[38] (see Seitz, Angel and Paloutzian, 2023).

Belief-forming is shown to be guided by perceptive and subjective valuations of environmental input (Seitz & Angel, 2020: 2). Non-propositional representations of one's environment in primal beliefs can be influenced by our tacit evaluation. If Tony the human and Tony the hedgehog see a poisonous snake, the object of perception for both creatures would be the same external object. However, they may form very different non-propositional primal beliefs. Tony the human may have a belief with environmental content of a snake being present, but his representation is also influenced by the subjective valuation of what he perceives to be DANGEROUS[39]. Tony the hedgehog may have a belief with the exact same environmental content[40], but his representation will not be influenced by the same subjective valuation (i.e., a hedgehog has no evolutionary reasons to fear a poisonous snake).

Depicted like that, belief-forming can go a long way in explaining interesting philosophical phenomena as cognitive penetration or Tamar Gendler's "aliefs". Unfortunately, it will not be possible to examine these matters here, but I still think it important that an account of primal beliefs can re-open other philosophical debates by suggesting a new solution to the problems they raise.

Finally, and on a more speculative note, primal beliefs can be examined as natural kinds. There are grounds to assume that there are basic units produced in the process of belief-forming. It has been insinuated that their formation can be observed and they are referred to as "babs" (Angel & Seitz, 2016: 5). The emergence of new basic units is accompanied by a change in behaviour. Belief formation, in this sense, is not 'guided' by anything apart from our relation and affective reactions to the external environment. Note that such reactions may be misguided; thus it would be a stretch to claim that primal beliefs are related (in any relevant way) to truth.

Seitz and Angel's studies also involve a third type of beliefs which they call "conceptual beliefs" or "belief in". They are supposedly much more influenced by our social environment and cultural evolution and are described as abstract beliefs whose contents are "narratives" (2020: 2). That may be closer to what Matthew Chrisman and Neil Levy are looking for in their social accounts of believing. Traditionally however, epistemology has always been concerned with beliefs *that*[41]. For the time being, I will continue to focus on this more traditional problematic.

What remains to be seen is, if we are to reject TVB and all norms that adopt it as their premise, what positive normative account of belief forming can be proposed? In the last part of this paper, I will attempt to provide some guidelines for how such an account should look. The proposed account is by no means complete and it is facing numerous challenges, but I hope that it lays the foundations for an interesting new direction into normative theories of belief.

4. Functional Environmentalism

If primal beliefs pre-date language, are often tacit and non-propositional and are products of type-1 information processing, it can be questioned whether we can even have epistemic norms about them. At any rate, none of the ones involving the reflective component and the mental agency component of TVB would work for primal beliefs. In this last part of my paper, I want to suggest that a normative approach to primal beliefs can be justified.

I would like to go back to proper functionalism for a moment. An evolutionary explanation would allow us to see that belief-forming is a function of a cognitive system that has been acquired to serve a certain goal. Thus, I agree that primal beliefs should *aim* at *something*. On the other hand, if Angel and Seitz (2016) are correct, the aim of belief-forming should be assessed without referring to notions of

[38] See 1.2. and compare to my assumptions about belief maintenance.

[39] By this I don't mean that Tony the human needs to apply the concept DANGEROUS to any environmental content nor that his representation needs to have any conceptual, non-sensory content.

[40] That assumption is highly problematic because of the differences between hedgehog and human perception, but for the sake of the argument, let's grant that it is possible under certain conditions.

[41] There are various ways to account for the difference between beliefs *that* and beliefs *in*, but an example may serve best. Martha's belief *in* God can also be referred to as "identity belief" or "identity-defining belief" and is sometimes examined in virtue of features defining social cognition and group belonging (e.g., Williams 2022). The traditional target of philosophical analysis, however, remains Martha's belief *that* God exists. The last one can correspond or fail to correspond to an actual state of affairs, and in virtue of that, its propositional content may turn out to be true or false.

truth, knowledge and rationality (see 3.1.). Being non-propositional, beliefs could not possibly be aiming at truth. At the end of the day, truth is (depending on your preferred view) either a property of propositions or sentences. Representational content, especially non-conceptual content, is not true or false. It is simply more or less accurate[42].

Could it be that beliefs aim at knowledge? While I agree with a lot of what Mona Simion has to say in her *Resistance to Evidence*, I think that it would be too ambitious to say that the function of a belief-forming system is to produce knowledge and that it is questionable if primal beliefs amount to knowledge in any non-minimal sense[43]. That is not to say that knowledge is not valuable or that epistemology should not raise normative claims involving the concept of knowledge. However, norms of belief and especially primal belief should simply involve a different notion[44].

Nevertheless, I think that Simion is right that every function comes with an associated norm:

> [T]hese [associated norms] prescribe the right way to proceed in order to reliably enough fulfil the function in question under normal conditions. Because its function contributes to the explanation of its very existence, the trait in question ought to perform in a way that is associated with likely enough function fulfilment. (Simion, 2024: 101)

Studies have shown that the hard-wiring and proper function of a cognitive system are crucial factors determining if the system is going to do better or worse in belief-forming. For example, delusion or social malfunction manifested in the holding of inaccurate beliefs can be traced to a breakdown in neural processes that correspond to and are constitutive of belief-forming (see Pechey & Halligan, 2011). Thus, I agree that a belief system's proper function should play a key role in a normative account.

What do I mean by saying that a cognitive system can do "better or worse"? As already stated, primal beliefs are non-propositional and a probability criterion will not be applicable to this case. Therefore, I propose that we impose accuracy conditions. A cognitive system functions well in forming primal beliefs when acquired non-propositional representations are accurate enough, i.e. when the system can reliably form predictions for effective action. Which actions are effective? Those that promote the adaptiveness of the system to its environment. Belief-forming is about being informed about one's environment or "being tuned" to it[45]. In other words, the function of primal beliefs is to be tuned to signals from the surrounding environment.

I think that the right normative approach to belief-forming (i.e. to primal belief) would combine the proper function component with an environmental component. That would probably also require a lot to be said about the metaphysics and epistemology of perception. This is a future project that needs a lot of work, but for now I propose that the following norm of belief that I contend is applicable to primal beliefs:

Accuracy Norm: Beliefs are aiming at accurate representation of one's environment.

Accurate representation requires proper function as a necessary condition for fulfilment. However, proper function is not enough. Even if one's cognitive system functions perfectly, in cases where one interacts with a bad epistemic environment[46], this interaction may produce inaccurate beliefs. There-

[42] I would like to thank Mitch Green for our long discussion on accuracy conditions that has inspired the last part of this paper.

[43] It remains possible that they can if knowledge is understood in some specific sense, e.g., as non-propositional, tacit and not requiring truth, but simply an accurate representation of one's environment. As I am not sure if any existing normative account would sign under this picture of knowledge, for now I will presume that the belief-forming function has not been acquired to produce knowledge in any familiar sense.

[44] Maybe it is the case that belief revisions aim at knowledge. That remains to be explored.

[45] Another reason to think that perceptual systems were selected for delivering information that would allow life-forms to improve interactions with their environment (see Godfrey-Smith (2017) for extensive speculation). As by-products of perceptive and affective processing, primal beliefs were not selected for some higher-order rational aim as assumed by TVB, but for a much simpler goal – surviving and adapting to one's environment.

[46] The analysis of primal beliefs would have a lot to do with examining epistemic environments and defining the features of a bad vs. good epistemic environment. I am extremely sympathetic to the Environmental epistemology project (Amico-Korby, Harell & Danks, 2024) and to attempts to account for beliefs in epistemically polluted environments (Levy, 2021, 2023). However, I should underline that while all these projects refer mainly to one's social environment, I use "environment" in a wider sense.

fore, I think that the right way to go is to adopt Functional Environmentalism.

Radical sceptical scenarios and arguments from hallucination would present an especially interesting problem for my approach. Representations of simulated environments are – in an important sense – inaccurate. Being positioned in such an environment would render the system incapable of reliably making predictions for effective action.

However, we do not need radical scepticism to enter the scene in order to examine another interesting and much more trivial challenge. It can be argued that cases of gaslighting are such that the belief-forming system is functioning properly and it is even possible that the primal beliefs are initially accurate, but an environmental factor (in this case – the Gaslighter) brings about a second-order belief that leads to a revision and dropping of the primal belief. Cases like this show that environmental conditions can distort even perfectly tuned cognitive systems that tend to reliably produce accurate beliefs. They also show that – on occasion – second-order beliefs can do more harm than good.

Finally, everything I have said by now applies to belief-forming and primal beliefs. But one can object that this is not what we should care about in epistemology. The challenge can continue by assuming that even if I am right about epistemologists not providing norms about primal beliefs, they should not be expected to do so. Epistemic normativity may be said to apply to updated, revised beliefs and to concern rational belief-revisions (where "rational" is not understood in a minimal sense – see footnote 32). If that is the case, there may be a tension between an accurate primal belief and a seemingly rational revision that distorts its initial accuracy as suggested by the Gaslighter example. Also, it remains a valid point that the descriptive account of beliefs provided by TVB is not compatible with mounting empirical data about beliefs and belief-forming. To satisfy a modest content naturalism requirement[47], the transcendentalist should account for these discrepancies and should provide reasons why we should discard said empirical evidence.

Acknowledgements

My research was funded by the "Sofia University Marking Momentum for Innovation and Technological Transfer" (SUMMIT) project and the Centre for Advanced Study (CAS) Sofia and I would like to express my gratitude for that. I am also very grateful to Mitch Green, Giorgia Foti, Jeremy Pober, Jack Lyons and David deBruijn for the long, provocative and productive conversations concerning different aspects of this paper. The final result would not be the same without the insightful comments of my two anonymous reviewers, as well as without the comments and questions of the audiences of Naturalism in Philosophy International Conference, the Metaphysics course at the Inter University of Dubrovnik and the Empirical Epistemology panel at the XXV World Congress of Philosophy. Last, but not least, I would like to thank Mona Simion for hosting me as a guest researcher at the COGITO centre where I had the chance to meet a lot of talented colleagues and to receive detailed feedback on a draft of this paper.

References

Amico-Korby, D., Harrell, M. & Danks, D. (2024). Environmental epistemology. *Synthese, 203*(81), 1–24.

Angel, H.-F. & Seitz, R. J. (2016). Process of believing as fundamental brain function: The concept of credition. *SFU Research Bulletin, 3,* 1–20.

Bellini-Leite, S. C. (2018). Dual Process Theory: Systems, Types, Minds, Modes, Kinds or Metaphors? A Critical Review. *Rev.Phil.Psych. 9*, 213–225.

Carter, J. A. (2022). On some intracranialist dogmas in epistemology. *AJPH, 1*, 44.

Chrisman, M. (2022). *Belief, Agency, and Knowledge: Essays on Epistemic Normativity*. Oxford: Oxford University Press.

Christensen, D. (2024). Epistemic akrasia: No apology required. *Noûs, 58*(1), 54–76.

Comesaña, Juan (2015). Can We Believe for Practical Reasons? *Philosophical Issues*, *25*(1), 189–207.

Emery, N. (2023). *Naturalism Beyond the Limits of Science: How Scientific Methodology Can and Should Shape Philosophical Theorizing.* New York, US: Oxford University Press.

Evans, J. St. B. T. (2009.) How many dual-process theories do we need? One, two, or many?, In Jonathan Evans, and Keith Frankish (Eds.), *In two minds: Dual processes and beyond* (pp. 33–54). Oxford: Oxford University Press.

[47] Which I think Nina Emery has successfully showed that most philosophers attempt to satisfy (Emery, 2023: 10–44).

Gibbons, J. (2013). *Norm of Belief*. Oxford: Oxford University Press.

Gigerenzer, G. (2000). *Adaptive thinking: Rationality in the real world*. Oxford: Oxford University Press.

Godfrey-Smith, P. (2017). *Other Minds: The Octopus, the Sea, and the Deep Origins of Consciousness*. London: William Collins.

Grzankowski, A. & Montague, M. (2018). Non-propositional intentionality: an introduction. In Alex Grzankowski & Michelle Montague (Eds.), *Non-Propositional Intentionality*. Oxford, United Kingdom: Oxford University Press.

Hawthorne, John, Rothschild, Daniel, & Spectre, Levi. (2016). Belief is weak. *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition*, *173*(5), 1393–1404.

Horowitz, S. (2013). Epistemic Akrasia. *Noûs, 48*(4), 718-744.

Hughes, N. (2019). Uniqueness, Rationality, and the Norm of Belief. *Erkenntnis, 84*(1), 57–75.

Kornblith, H. (2011). Transcendental Philosophy. In Joel Smith & Peter Sullivan (eds.), *Transcendental Philosophy and Naturalism* (pp. 96–119). Oxford University Press.

Kornblith, H. (2012). *On Reflection*. Oxford: Oxford University Press.

Kriegel, U. (2018). Beliefs-That and Beliefs-In. Which Reductive Analysis? In A. Grzankowski & M. Montague (Eds.) *Non-propositional Intentionality* (pp. 192–213). Oxford: Oxford University Press.

Levy, N. (2021). *Bad Beliefs: Why They Happen to Good People*. Oxford University Press.

Levy, N. (2023). It's Our Epistemic Environment, Not Our Attitude Toward Truth, That Matters. *Critical Review*, *35*(1–2), 94–111.

Littlejohn, C. (2013). No evidence is false. *Acta Analytica*, *28*(2), 145–159.

Littlejohn, C. (2017). Objectivism and Subjectivism in Epistemology. In Veli Mitova (Ed.), *The Factive Turn in Epistemology.* New York: Cambridge University Press.

Littlejohn, C. & Dutant, J. (2021). Even if it might not be true, evidence cannot be false. *Philosophical Studies*, *179*(3), 801–827.

Lycan, W. G. (1988). *Judgement and Justification.* New York: Cambridge University Press.

Pechey, R., & Halligan, P. (2011). The prevalence of delusion-like beliefs relative to sociocultural beliefs in the general population. *Psychopathology, 44*, 106–115.

Pol, D., Mancuso, A. C., Smith, R. M. H. et al. (2021). Earliest evidence of herd-living and age segregation amongst dinosaurs. *Sci Rep, 11*, 20023.

Rinard, S. (2019). Believing for Practical Reasons. *Noûs,* (4), 763–784.

Sachs, O., & Hirsch, J. (2008). A neurology of belief. *Annals of Neurology*, *63*, 129–130.

Seitz, R. J. & Angel, H.-F. (2020). Belief Formation – A Driving Force for Brain Evolution. *Brain and Cognition, 120*(105548).

Seitz, Rüdiger J., Paloutzian, Raymond & Angel, Hans-Ferdinant. (2016). Processes of believing: Where do they come from? What are they good for? *F1000Res*.

Seitz, R. J., Paloutzian, R., & Angel, H-F. (2023). Bridging the Gap Between Believing and Memory Functions. Eur J Psychol, *19*(1), 113–124.

Shah, N. (2003). How truth governs belief. *Philosophical Review, 112*(4), 447–482.

Simion, M. (2024). Resistance to Evidence. Cambridge: Cambridge University Press.

Simion, M, Kelp, C., & Ghijsen, H. (2016). Norms of Belief. *Philosophical Issues*, *26*(1): 374–392.

Smithies, D. (2011). The normative role of knowledge. *Noûs, 46*, 265–288.

Sommers, F. (2005). Belief De Mundo. *American Philosophical Quarterly, 42*(2), 117–124.

Sommers, F. (2009). Dissonant beliefs. *Analysis,* 69 (2):267–274.

Williams, B. (1970). Deciding to believe, reprinted in his *Problems of the Self*. Cambridge: Cambridge University Press, (1973), 136–151.

Williams, D. (2022). Identity-Defining Beliefs on Social Media. *Philosophical Topics, 50*(2), 41–64.

Williamson, T. (2000). *Knowledge and its limits*. New York (NY): Oxford University Press.

NATURALISM AND FORMAL LOGIC

Pablo Caballero

Universidad de Sevilla, Spain
e-mail: pabcabfer@gmail.com
ORCID ID: https://orcid.org/0000-0002-4187-8413

Abstract:
We offer an overview of the relation between formal logic and naturalism. In Section 2, examples are considered of empirically-motivated logics. In Section 3, an empirically-motivated interpretation of alethic modalities is outlined. Through both issues, we attempt to support the thesis that formal logic can be regarded as a quasi-empirical, or already naturalized discipline.

Key words: naturalism, empiricism, classical logic, bivalence, quantum logics, alethic modalities, syntactic predicates.

1. Introduction

We offer an overview of the connections between naturalism and formal logic. To this end, different examples are considered of developments in the field of logic in the last century which are directly connected to empirical research. Firstly, Section 2 provides several examples of empirically-motivated logics, (non-classical) logic systems which show an empirical motivation and connections with scientific research. Secondly, in Section 3, an empiricist interpretation of alethic modalities is outlined on the basis of the ideas proposed in (van Fraassen, 1980) and, above all, in (Carnap, 1935) and (Stebbing, 1935). We thus attempt to support the thesis that formal logic can be regarded as a *quasi-empirical* and *already naturalized* discipline.

2. Empirically-motivated logics

The reflection on the connection between formal logic and empirical sciences (now involved in the more general philosophical debate on naturalism — see, e.g., Maddy, 2024) emerged early in the 20th century and led to the construction of different types of non-classical logics that can be regarded as empirically-motivated.[1] The aim of this section is to briefly describe several examples of such logics so as to motivate the idea that *formal* logic (not equivalent to *classical* logic — a particular case of the former) can be regarded as an already naturalized discipline often susceptible to empirical interpretation.

Quantum mechanics is frequently seen as the paradigmatic example of a field of knowledge whose discoveries have led to the development of alternative logic systems. For instance, and as a consequence of his rejection of the distinction between synthetic and analytical statements, Quine argued in *Two dogmas of empiricism* that

> [r]evision even of the logical law of the excluded middle has been proposed as a means of simplifying quantum mechanics; and what difference is there in principle between such a shift and the shift whereby Kepler superseded Ptolemy, or Einstein Newton, or Darwin Aristotle? (Quine, 1951, p. 43).

Thus, insofar as the aforementioned distinction is abandoned, the revision of certain classically-valid logical laws (e.g., the excluded middle[2]) in the light

[1] As will be seen, many of them are inspired by the research in the field of quantum physics.
[2] The law of excluded middle fails, for example, in Reichenbach's quantum logic (Reichenbach, 1944) and in

of empirical evidence will not be essentially different to the revision of theories of the kind mentioned in the above fragment. (For a detailed and up-to-date study of Quine's position, see (Maddy, 2024)). It will now be seen that several of the empirically-motivated logics described below are inspired by the research in the field of quantum physics.

Among the so-called *quantum logics* is Reichenbach's three-valued logic (R3), originally developed in (Reichenbach, 1944) as a means to solve the causal anomaly that arises in the standard interpretation of the double-slit experiment. The main distinctive feature of R3 is that (together with the classical values 1 and 0, 'truth' and 'falsehood') it adopts a third truth-value, i, interpreted as 'indeterminate' or 'neither verified nor falsified' (Reichenbach, 1944, pp. 42, 145), which is assigned to propositions describing *interphenomena*, physical phenomena that cannot be understood in causal terms (cf. Reichenbach, 1944, p. 117). Among other things, the addition of *i* allows for invalidating the classically-valid patterns of inference that lead to the aforementioned anomaly. The technical details of the solution are given in (Reichenbach, 1944, §33), and recent studies of Reichenbach's project can be found in (Estrada González & Cano Jorge, 2021) and (Caballero & Valencia, 2024).

From a different point of view (but having quantum physics in mind), Łukasiewicz regards the *principle of bivalence* of classical logic as a presupposition underlying *determinism* (Łukasiewicz, 1922/1923, p. 128), a conception which, for him, was undermined precisely by the development of atomic physics in the early 20th century (Łukasiewicz, 1922/1923, p. 110). Furthermore, he sees his revision of classical logic and his development of a three-valued logic in (Łukasiewicz, 1920) as essentially connected to empirical investigation:

> Whether that new system of logic [Łukasiewicz's three-valued logic] has any practical importance will be seen only when logical phenomena, especially those in the deductive sciences, are thoroughly examined, and when the consequences of the indeterministic philosophy, which is the metaphysical substratum of the new logic, can be compared with empirical data (Łukasiewicz, 1920, p. 88).

According to Łukasiewicz, then, empirical factors have to be considered both for the questioning of the classical principle of bivalence (essential to determinism) and to provide a justification of the relevance of his three-valued system, the "metaphysical substratum" of which is "indeterministic philosophy" — and which is, in turn, motivated by quantum physics (see Łukasiewicz, 1922/1923, p. 110, quoted above).

On the other hand, the discussion on the validity of the laws of distribution in quantum physics emerged in (Birkhoff and Von Neumann, 1936) and also in (Fevrier, 1937), and is later addressed in detail in (Putnam, 1968). Following the latter, the inference from $A \wedge (B_1 \vee \dots \vee B_n)$ to $(A \wedge B_1) \vee \dots \vee (A \wedge B_n)$ by applying the distribution of $\wedge$ over $\vee$ may fail in the domain of quantum physics, for instance, when A states the momentum of a particle and $B_1 \dots B_n$ its possible positions in a certain system[3] (see also Maddy, 2024: 14 on this issue).

Finally, other non-classical systems (apart from quantum logics) have been developed that ultimately pose the question whether classical bivalent logic should be regarded as a limit case in a variety of logics, many of which are empirically-motivated. Consider, for example, Graham Priest's three-valued logic, LP (Priest, 1979), which includes the truth-values 'true' (1), 'false' (0) and 'both true and false' or 'paradoxical' (b), and regards 1 and b as designated values.[4] The main feature of LP is that it invalidates the classically-valid principle of explosion (the inference of any formula from a contradiction, $A \wedge \neg A \vDash B$).[5] In doing so, LP attempts to account for *non-trivially inconsistent* theories (Priest, 1979, p. 226). Although this concept is not discussed here, examples of such theories are provided, for instance,

Łukasiewicz's three-valued logic (Łukasiewicz, 1920), both mentioned below.

[3] The conclusion of the inference, according to Putnam, determines both the momentum and position of the particle beyond the constraints imposed by the uncertainty principle.

[4] The designated values of a logic are those truth-values preserved in logical consequence. In classical logic, only the truth-value *true* is designated.

[5] The truth-tables for LP are exactly those of K3, Kleene's three-valued paracomplete logic (Kleene, 1952, p. 332 ff.), although the paradoxical value is taken as a designated one in LP (whereas only 1 is designated in K3). If we let $v(A) = b$ and $v(B) = 0$, then $v(A \wedge \neg A) = b$, so $A \wedge \neg A \nvDash B$, given that the value of the premise is designated and that of the conclusion is not. The principle of explosion is thus invalidated in LP. On the other hand, note that other classical laws (like the law of excluded middle) hold in LP.

in (Priest, 2009, pp. 75-76), including some from the field of natural sciences.

3. A naturalistic approach to alethic modalities

3.1. Alethic modalities as syntactical predicates

After considering several examples of empirically-motivated logics, we now attempt to provide a naturalistic approach to alethic modalities (AM hereafter), such as *necessity* and *possibility*.[6] It has been often assumed within the field of formal logic that *possible worlds semantics* provides an adequate interpretation of AM. As is well-known, a Kripke model for modal logic is a structure

$$M = \langle W, R, V \rangle$$

in which the non-empty domain W is interpreted as a set of *possible worlds*, R is a binary relation on W (the accessibility relation between worlds), and V maps atomic formulas onto subsets of W. If R is assumed to be universal, the semantics for necessity ($\Box$) and possibility ($\Diamond$) can be defined as follows (where 'iff' stands for 'if, and only if'):

- $M \vDash \Box\varphi$ iff $M, w \vDash \varphi$ for all $w \in W$
- $M \vDash \Diamond\varphi$ iff $M, w \vDash \varphi$ for some $w \in W$

Intuitively, a formula is necessary iff it holds in every possible world, and possible iff it holds in at least one possible world. Although *modal realism* concerning the nature of possible worlds is not examined here, we note that some of the main arguments in favor of this theory can be found in (Lewis, 1986) and (Divers, 2002), whereas detailed criticisms are given in (Bueno and Shalkowski, 2014) and (Bueno, 2022).

Alternative interpretations of AM have been proposed from an empiricist point of view that intend to avoid the metaphysical commitments involved in modal realism and the possible worlds interpretation. Following van Fraassen, justice can be done to modal language from the aforementioned point of view "if we can explicate the use and structure of modal language without concluding that anyone who does use it is committed to some sort of metaphysical beliefs such as that alternative possible worlds are real" (van Fraassen, 1980, p. 198). Furthermore, he clarifies the use of AM as follows:

> Guided by the scientific theories we accept, we freely use modal locutions in our language. Some are easily explicated: if I say that it is impossible [...] to melt gold at room temperature, this is because no counterpart to such events can be found in any model of the science I accept. But our language is much subtler and richer than that; its modal locutions reflect the fact that in the models of our theories we see structures that correspond to alternative courses of events, not all of which could be jointly actualized. [...] It does not require belief that all significant aspects of the models have corresponding counterparts in reality. This applies to many aspects discussed by philosophers of science: space-time, elementary particles, fields, and, finally, alternative possible states and courses of events. The locus of possibility is the model, not a reality behind the phenomena (van Fraassen, 1980, pp. 201–202).

Three main theses can be extracted from the above fragment:

1. Modal locutions make sense within a certain *theory* or *model* (e.g., a scientific theory).
2. Alethic modalities are *logically* or *linguistically* meaningful (but not, say, *ontologically* meaningful) insofar as they express a relation between a certain event or proposition with a currently accepted theory or model.
3. It is not necessary that all elements of modal locutions have a real counterpart (in particular, this is the case for "alternative possible states and courses of events").
   The interpretation of AM outlined in (Carnap, 1935) and (Stebbing, 1935) provides a very interesting starting point for the development of a formal system for AM which conforms to the above requirements.

From a point of view close to that of van Fraassen, Carnap noted in 1935 that the logical character of AM had not yet been conceived with complete clarity (Carnap, 1935, pp. 72-73), and proposed to regard AM as *syntactical properties* of propositions (formalizable as modal operators). He explains his conception in the following way, where the prefix 'P' implies reference to a physical language within which a sentence has sense (see Stebbing, 1935, p. 506 on this):

> Modality sentences are in fact veiled syntactical sentences [...]. Impossibility is a quality to which there is a parallel syntactical quality,

[6] Other AM may be defined, e.g., 'impossibility' or 'contingency' (see Section 3.2).

> namely contradictoriness [...]. Sometimes, however, the idea of impossibility has the sense not of *logical*, but of *physical* or *real* impossibility. In this latter case the parallel syntactical term is not 'contradictory' but the corresponding general term, namely 'contravalid'. Take for instance the sentence "The state of a particular solid iron ball swimming on the water is physically impossible". The translation is: "The sentence 'This solid iron ball is swimming on the water' is contravalid" [...], that is, incompatible with the system of physical laws [...]. The sentence "An iron ball is heavier than a wooden ball of equal size" [...] is P-valid, that is, logically deducible from the system of physical laws (Carnap, 1935, pp. 73–74).

The idea that AM are properties of propositions which express a connection between such propositions and a given theory (e.g., a system of physical laws) is common to both van Fraassen's and Carnap's proposals. On the other hand, crucial to our study of Carnap's approach is Susan Stebbing's 1935 review of (Carnap, 1935), in which the characterization of AM as "syntactical predicates" is introduced. In Stebbing's own words, Carnap maintains that a P-valid sentence is deducible from the system of physical laws, whilst a P-contravalid sentence is incompatible with this system. The prefixing of P entails reference to a P-language within which the sentence has sense, and from which, indeed, it derives its sense. Thus modality terms are syntactical predicates (Stebbing, 1935, p. 506).

Carnap and Stebbing's interpretation of AM seems to conform to the three main theses extracted from van Fraassen's arguments, and provides several hints on how an empirically-motivated logical system for AM should be constructed. In Section 3.2 we outline the main characteristics of such a system.

3.2. Outline of an alternative formal system for alethic modalities

Let L be a first-order language[7] in which the concept of a well-formed formula (WFF hereafter) is defined by the following rule:

$$\mathrm{WFF(L)} \coloneqq P^n(t_1 \ldots t_n) \mid t_1 = t_2 \mid \neg\varphi \mid \varphi \wedge \psi \mid \forall x\varphi$$

where P^n is an n-ary predicate and every term t_i is either a variable, a constant, or a functorial term. Let us now define a new language L_N with the following syntax:

$$\mathrm{WFF}(L_N) \coloneqq \varphi \mid N\varphi \mid \neg\Phi \mid \Phi \wedge \Psi$$

where φ is any WFF(L), Φ and Ψ are $\mathrm{WFF}(L_N)$ and Nφ is read as "Necessarily φ" or "It is necessary that φ".[8] Finally, let us use L to define the following three sets for formulas (following Carnap, 1935, pp. 43-50):

1. a set S_1 of valid formulas, rules of inference and axioms of arithmetic,
2. a set S_2 of laws (e.g., of physics), and
3. a set S_3 of observational statements (i.e., particular empirical statements).

Note that, in 1, by *valid* we do not mean *analytical*, but (as in all formal logics) true under any interpretation in the semantics assumed, which may or may not be classical.[9] For the sake of simplicity, we will assume that our formal language is interpreted in bivalent classical semantics, but, for instance, a many-valued semantics could have been assumed. We now call

$$S_1 \cup S_2 \cup S_3$$

a *theory* and hereafter denote it by $\mathcal{T}$. As long as the logical framework assumed is classical, the requirement that the theory $\mathcal{T}$ be consistent could also be added. Hereafter we assume that $\mathcal{T}$ is indeed consistent.

It is now possible to define AM as syntactical predicates by providing that they are interpreted in terms of the syntactical relations between a given WFF(L) and a certain theory $\mathcal{T}$. For this, introduce AM according to the conditions shown in Table 1, where ⊢ expresses the relation of deduction from the premisses (the theory $\mathcal{T}$, in this case) to the conclusion (in a given calculus suitable for the logic assumed). Many of the expected properties of AM still

[7] We choose a first-order language as our starting point, given that propositional language is unable to formalize, e.g., scientific laws in a satisfactory manner.

[8] Note that, given the syntax of L_N, the iteration of AM is not allowed, so the characteristic axioms of modal systems such as S4 or S5 ($\Box\varphi \rightarrow \Box\Box\varphi$ and $\Diamond\varphi \rightarrow \Box\Diamond\varphi$, respectively) are not expressible within this language.

[9] The distinction between *valid, contingent,* and *contradictory* propositions is essential to all formal logics, not only to classical logic. Note that accepting such a distinction does not further involve accepting a distinction between analytical and empirical statements, nor accepting any other ontological commitments related to it. Obviously, the distinction between analytical and empirical statements is contradictory with the point of view expressed in the first part of this paper, in which formal logic is treated as an already naturalized or quasi-empirical discipline.

Table 1. The syntactical interpretation of AM

Modality	Operator	Definition	Syntactical interpretation
Necessity	Nφ	—	$\mathcal{T} \vdash \varphi$
Possibility	Pφ	¬N¬φ	$\mathcal{T} \nvdash \neg\varphi$
Impossibility	Iφ	N¬φ	$\mathcal{T} \vdash \neg\varphi$
Contingency	Cφ	¬Nφ ∧ ¬N¬φ	$\mathcal{T} \nvdash \varphi$ and $\mathcal{T} \nvdash \neg\varphi$

hold in this interpretation. We list some of them below (since the proofs are simple in all cases, we omit them):

$\vdash N\top$	$\vdash N(\varphi \to \psi) \to (N\,\varphi \to N\psi)$
$\vdash I\bot$	$\vdash (N\,\varphi \wedge N\psi) \leftrightarrow N(\varphi \wedge \psi)$
$\vdash N\,\varphi \leftrightarrow \neg P\neg\,\varphi$	$\vdash P(\varphi \wedge \psi) \to (P\,\varphi \wedge P\psi)$
$\vdash N\,\varphi \to P\,\varphi$	$\vdash (N\,\varphi \vee N\psi) \to N(\varphi \vee \psi)$
$\vdash N\,\varphi \to \varphi$	$\vdash (P\,\varphi \vee P\psi) \to P(\varphi \vee \psi)$
$\vdash \varphi \to P\,\varphi$	$\vdash \neg C\,\varphi \leftrightarrow (N\,\varphi \vee I\,\varphi)$

$\top$ stands for any valid WFF(L), $\bot$ stands for any contradictory WFF(L), and φ and ψ stand for any two WFF(L).

Note that, although AM become theory-dependent notions in the approach we have just outlined — as in that of Carnap (1935) and Stebbing (1935) — , the formalism presented here remains neutral with respect to both (1) the semantics and (2) the theory $\mathcal{T}$ chosen in a particular case.

Concerning (1), remember that the semantics assumed for the construction of the theory $\mathcal{T}$ is here assumed to be classical, but it could have been chosen to be non-classical: if it were, e.g., paracomplete, certain classical tautologies like the law of excluded middle would not be necessary.

Concerning (2), let us consider, for instance, Newtonian and relativistic physics. Let F_1 and F_2 be inertial frames of reference, and let E_1 and E_2 be two events which are simultaneous from the point of view of F_1. Then, in Newtonian physics, the proposition "E_1 and E_2 are simultaneous in F_2" is necessarily true (since simultaneity is assumed to have an absolute significance, it follows that the events in question must be simultaneous in both frames of reference), while it is not in relativistic physics (the proposition in question may be false, indeed). Whether or not a given modality applies to a proposition thus completely depends on the theory $\mathcal{T}$ chosen. This implies two things: firstly, it makes no sense to state that a proposition is possible, necessary, etc. *without* referring to a theory; secondly, the characterization of AM presented here involves a *relativization* of AM, the application of which depends on the chosen theory.

Conclusions

Throughout this paper, an overview of the connection between logic and naturalism has been offered. In Section 2, several examples were provided of empirically-motivated non-classical logics developed in the 20th century, most of which show a clear connection with quantum physics. In Section 3, an alternative and empirically-motivated interpretation of AM was proposed which is based on their consideration as syntactical and theory-dependent properties of propositions that get their meaning within the framework of a theory. We consider that both the variety of logics considered in Section 2 and the proposal developed in Section 3 allow us to support the idea that *formal* logic (not equivalent to *classical* logic) can be regarded as an already naturalized discipline susceptible to empirically-committed interpretations.

References

Birkhoff, G. & von Neumann, J. (1936). The Logic of Quantum Mechanics. *Annals of Mathematics*, *37*(4), 823-843, https://doi.org/10.2307/1968621

Bueno, O. (2022, December 14–16). *Models and modal epistemology* [Key-note conference]. 10th Conference of the Spanish Society of Analytical Philosophy (SEFA 2022), Santiago de Compostela, Spain.

Bueno, O. & Shalkowski, S. A. (2014). Modalism and theoretical virtues: toward an epistemology of modality. *Philosophical Studies*, *172*(3), 671–689, https://doi.org/10.1007/s11098-014-0327-7

Caballero, P. & Valencia, P. (2024). A sound and complete tableaux calculus for Reichenbach's quantum mechanics logic. *Journal of Philosophical Logic*, *53*, 223–245, https://doi.org/10.1007/s10992-023-09730-7

Carnap, R. (1935). *Philosophy and logical syntax*. London: Keagan Paul, Trench, Trubner & Co. Ltd.

Divers, J. (2002). *Possible Worlds*. London: Routledge.

Estrada González, L. & Cano Jorge, F. (2021). Revisiting Reichenbach's logic. *Synthese*, *199*, 11821–11845, https://doi.org/10.1007/s11229-021-03313-2

Fevrier, P. (1937). Les relations d'incertitude d'Heisenberg et la logique. *Travaux du IXe Congrès International de Philosophie*, *6*, 88–94, https://doi.org/10.5840/wcp919376147

Kleene, S. C. (1952). *Introduction to Metamathematics*. Amsterdam: North Holland Publishing Company.

Lewis, D. (1981). *On the Plurality of Worlds*. Oxford: Blackwell.

Łukasiewicz, J. (1920). On three-valued logic. English translation in: (Łukasiewicz, 1970).

Łukasiewicz, J. (1922/1923). On determinism. English translation in: (Łukasiewicz, 1970).

Łukasiewicz, J. (1970). *Selected Works*. (L. Borrowski Ed.). Amsterdam: North Holland Publishing Company.

Maddy, P. (2024). Varieties of naturalism in logic. In: Ferrari, F. et al. (Eds.). (2024). *Oxford Handbook of Philosophy of Logic*. Oxford: Oxford University Press. https://www.academia.edu/98346545/Varieties_of_naturalism_in_logic?email_work_card=thumbnail

Priest, G. (1979). The logic of paradox. *Journal of Philosophical Logic*, *8*(1), 219–241, https://doi.org/10.1007/BF00258428

Priest, G. (2008). *An Introduction to Non-Classical Logic*. Cambridge: Cambridge University Press.

Putnam, H. (1968). Is Logic Empirical? In: Cohen, R. S., Wartofsky, M. W. (Eds). *Boston Studies in the Philosophy of Science*, *5*. Dordrecht: D. Reidel Publishing Company.

Quine, W. V. O. (1951). Two dogmas of empiricism. In: Quine, W. V. O. (1961). *From a Logical Point of View* (pp. 20–46). Cambridge: Harvard University Press.

Reichenbach, H. (1944). *Philosophic Foundations of Quantum Mechanics*. Berkeley and Los Angeles: University of California Press.

Stebbing, L. S. (1935). Review of Logische Syntax der Sprache (1934), Die Aufgabe der Wissenchatslogik (1934), Philosophy and Logical Syntax (1935) and The Unity of Science (1934), by Rudolf Carnap. *Mind*, *44*(176), 499–511.

van Fraassen, B. C. (1980). *The scientific image*. Oxford: Oxford University Press.

I BET ON CÉZANNE'S CARDS: AN AESTHETIC APPROACH OF "STILL-LIFE". PHILOSOPHICAL UNDERTAKINGS OF LUCK, SOCIABILITY AND HUMAN CONDITION IN IMPRESSIONIST ART

Oana Serban

University of Bucharest, Romania
email: oana.serban@filosofie.unibuc.ro
ORCID ID: https://orcid.org/0000-0002-6699-2482

Abstract:
This article tackles the possibility of reframing Cézanne's reception of *still life* as a *meta-painting* in which individuals occasionally become objectified, gathering objects with other ontological dignities, such as things and tools, in a Heideggerian manner. One of the rare situations in which individuals are presented like objects throughout a *still life* is represented by the series of *Card Players*. I will argue that the still life form is reflected in objectual painting in which the insertion of individuals is performed in order to recover a Bergsonian *élan vital*, of which Cézanne retains, by engaging Impressionist techniques, only an instance: the circumstance of a meditative, decisive, ludic action, taken during the game. Once humanised, *still life* has, in Cézanne's version, the capacity to express the human condition through two inherent determinations, luck and sociability, reflecting what Arendt would call "the human corner"–something that anyone needs in order to preserve natural instincts, to be "happy around things" and to begin a civilisational process, supported by the double realm of arbitrariness and normativity that lies at the heart of a game.

Key words: still-life, Impressionism, meta-painting, Cézanne, élan vital, luck, sociability.

Introduction. Premises for a phenomenological undertaking of Cezanne's Impressionism: perception, luck and *élan vital* in the *Card Players*

Even though still life was a familiar form in painting, it progressively became a trademark of Cézanne's paintings in particular. More than 200 sketches and fine artworks combine traditional and modern themes, in Provençal colours and habits, in pursuit of an impressionist outlook. Tools and objects – in Heidegger's terms (Heidegger 1977) – are depicted in multiple forms of still life, presented in order to satisfy the exigencies of perspective *studies*. Some examples, such as *Apples, Peaches, Pears, and Grapes* (1879-80), *Still Life with Plate of Cherries* (1885-87), *Still Life with a Basket* (1890-95), *Le vase paillé* (1895) or *Still Life with Skull* (*Nature morte au crâne*) (1895–1900) reveal that Cézanne was an object addict. All these examples implement variable objects, but the significance of each object is diluted once it comes into contact with another. Tools, reflected by such arbitrary objects, emerge into things, and so meaning paintings that concretise the so-called representational perspective, which is of interest here. This article tackles the unconventional traits of *still life* once *individuals*, not *tools*, are captured in an impressionist painting, a situation which is rarely encountered in Cézanne's art. Implicitly, we shall see that unlike the traditional perspectives on still life, reduced to representing the impression that the artist lives tracking the fall of light on objects that repeat in his paintings – e.g. a skull, a pipe, fruits or flagons – the variations of *Card Players* challenge the spectator to follow impressions that human agents create in both handling such objects while they are engaged in a game of cards and reacting to their interlocutors. Impressionism frames, in this particular case, three types of impressions: how human agents are impressed by objects, how human agents are impressed by their interlocutors and how the artist and spectator are impressed by card players depicted while reacting to both objects and humans. We confront, in these circumstances, a chain of impressions, one embedded in the other, that evokes the character of meta-painting in the *Card Players* in a very original manner.

I will argue that on the one hand, still life prescribes, under these circumstances, a specific human condition pleading for privileging the arbitrary liberty of an individual and its contingencies, revealed through luck as an impressionist vital energy. Reactions of card players are not just a matter of strategy, they are also influenced by uncalculated or unpredictable forces, belonging to contingencies. Technique is not enough: luck enlightens the virtues of players. More than half of the variations of *the Card Players* depict witnesses of the game with various impressions and stances toward this playful activity. The fact that one does not have the privilege of knowing a priori how a game will end makes the spectators behave attentively, impressed by a cluster of elements: the decisions of the card players, grimaces, gestures of choosing one card in particular, which, by the moment it reaches the table, remains an unknown variable, a truth still not enlightened. Luck – although rooted in a very naïve perspective on life – divides the table between losers and winners. Would this imminent ingredient be the one responsible for inducing a variety of impressions, based on astonishment, wonder, disbelief, enthusiasm, both in the players and spectators? On the other hand, once the human factor is inserted into the composition of the still life, still life begins to reflect "an extension of our being as masters of nature, as artisans and tool users", which coincides with the process of "humanizing of culture through the discovery of nature's all-inclusiveness and man's power of transforming his environment" (Schapiro 1952, 14). Consequently, I do argue in favour of accepting the insertion of the human subject into still life paintings as a trick to challenge the spectator reflecting on the shift from nature to culture, on the possibility of turning the human subject into a stimulus for our aesthetic experience grounded on impressions–a stimulus of aesthetic dignity equal to any other object included in the still life, (albeit ontologically, it has more dignity than they have). It seems to me that Cézanne masters the still life very differently than other painters. Still life concerns, in these terms, how objects and subjects cross a particular moment of their existence, an instance, which is temporally irrecuperable, although the painting succeeds in freezing it and converting it into a static representation worthy of being further explored. Life really is transformed on multiple levels from the life of objects to the life of beings mutually interacting or manipulating objects. Still life without human subjects is more a matter of placement and context encouraging reflection on things; on the human subject who is inserted into the painting; on the teleology of objects, which becomes obvious from a certain point of view. We shall see that painting mirrors the emancipation of objects, as Heidegger would call them, transitioning from things to tools, giving meaning to entities useful in satisfying human needs. And yet, in these variations of still life, the life of objects is not the only one at stake: what counts is the impression that a human subject has in interacting with objects that reveal the luck of the card players during a game. Cards are there not only to perform the game itself, but to raise awareness of this subtle entanglement of thinking and feeling, strategy and luck.

This twofold argument is supported by a case study on Cézanne's *Card Players* series as a still life composition that challenges our eye in quite unprecedented impressionist manners. However, unlike Schapiro, who restricts himself to appreciating Cézanne's mastery of still life with human actors, without further investigating the implications of such a challenge to the status of the painting as an object, I will drive my analysis towards a more provocative path, that of concluding the existence of a meta-still life in Cézanne's *Card Players*, as the individuals manipulating objects "move our eye"[1]

[1] According to Merleau-Ponty, the Impressionist style of Cézanne is particular because somehow in it the mind

from one layer of primary objectuality – reflected by the table of cards – to a meta-situation that collates impressions, emotions and a certain temporality of waiting for the players' reaction – depicted by the game itself. Such a shift occasions new insights of Cézanne's impressionism as a source for mastering meta-paintings that engage a particular phenomenology, which has at its heart the crossing of multiple subjectivities, and thus, temporalities.

Before moving further, I would like to clarify two major aspects that will shed light on my intentions and on the fundamental terms invested in this critical endeavour that, at first glimpse, is rooted in a phenomenological and hermeneutical analysis. First, the matter of luck might seem at least irrelevant, if not unsustainable in terms of a powerful philosophical literature supporting the viability of such a topic. And yet, it seems to me that *the élan vital* is not simply related to a mechanistic view of Cartesian inspiration, further prolonged by the Spinozist perspective on the *conatus* as perseverance into existence, and later consecrated in Bergson's terms. The vital impetus, as Bergson coined the primary understanding of *élan vital*, concerns a creative evolution that human beings transgress, along which natural life develops by its organic means, while consciousness proposes an intuitive perception on the self and inner time. However, this intimate will to live is not exclusively limited to organic factors. My argument is that the consciousness related to the so-called *élan vital* also embeds reflections on contingencies provided by unpredictable events and luck. Although Bergson never discussed luck directly, I suggest that being willing to persevere through existence is also determined by success and a positive reflection on our life. Being worthy of existence, willing to continue, in a Camusian perspective, finding reasons to transgress life, although it often seems despicable and pessimistic attitudes might arise, in a Nietzschean tone, brings luck to the heart of the major incentives of *élan vital*. In political philosophy, this would not be something odd. It was Machiavelli who considered that a good life and the conservation of power depends on the exercise of virtue (*virtù*), which is the only thing left into our hands and thus, predictable, and the use of *virtù* to overcome obstacles of our destiny (*fortuna*). Luck was related in the equation of political realism with this unpredictable layer of our existence occasioned by *fortuna*. But in phenomenology luck is still a tricky term to be invested into canonical topics, such as the Bergsonian *élan vital*. Cezanne's *Card Players* reflects the creative life and the inner time revealed by the consciousness of human agents engaged into a playful, normative and ultimately, cultural activity. But any card game brings resolutions equally based on luck. Some authors consider luck a cognitive illusion (Hales & Johnson 2015), whereas others assume that luck is a matter of training and control (McKinnon 2015). Implicitly, one might assume that "better players tend to stay in the game longer and thus tend to have more opportunities for luck to rear its head" (McKinnon 2015, 113). My assumption is that luck could be one of the many pillars from which the psychological, intentional and creative insights related to the *élan vital* might emerge. Consequently, I would prevent readers of this analysis from expecting to find a typical Bergsonian critical undertaking of luck, because there is none. However, I will tackle the challenge of bringing completion to the traditional perspectives on *élan vital*, mirrored by experiences of inner and outer time, contextualised, in this case, as a card game engaging different actors and witnesses, by adding a philosophical perspective on luck, sociability and the human condition. Secondly, my intention is to engage the artistic variations of *Card Players* as phenomenological stances that allow us to reflect on four major concepts of our human existence: the vital energy that people invest into cultural and playful activities; the nature of impressions that card players and witnesses of card games develop either by being attentive to their competitors or by following the dynamics of the game provided by decision-making, card exchanges, calculations and luck; the role of tools, understood from a typical Heideggerian perspective, as entities determined by a "readiness-to-hand" character, a category in which we might find cards, but also glasses, bottles, hats, pipes, chairs and other objects depicted in Cezanne's paintings; and ultimately, life itself, which is not only a matter of interval, but also a matter of moment or instant. Cezanne's *Card*

follows the eye in painting. (See Merleau-Ponty 1964, 159–190).

Players represents, in my opinion, an Impressionist depiction of life by engaging different regimes of temporality. For card players, who are directly engaged in their game, time is reduced to instances, to crucial moments when they have to calculate and cope with random, contingent factors that enrich their experience and pave their way to success. For witnesses observing a card game, the relationship with time becomes processual – it seems a long-lasting spectacle that people watch disinterestedly, but what seems to be interesting is their perception on cards falling, shifting, on player's looks, gestures and movements, on objects proving their essence as they are engaged in actions revealing their teleology and functionality. This is particularly why Cezanne's *Card Players* provides a different impressionist approach to the form of still life worthy of being philosophically explored.

Objectual painting: *still-life*

One of the canonical critiques of Cézanne's still life is that of Mayer Schapiro, a voice which throughout the history of the arts has been quite virulent when it comes to artistic accuracy and at the periphery of philosophical precision, as we have seen throughout his polemics with Heidegger[2] regarding Van Gogh's *Shoes*. In what concerns Cézanne, his observations are rather analytical, non-precipitated and, due to this, largely sustainable. Schapiro addressed Cézanne's still life with prudence, assuming that it invites contemplation, not assimilation. Still life is passive, almost non-performative.

"We are invited to look, but not to enter or traverse the space." (Schapiro 1952, 14)

A still-life is the result of a perspectival isolation, a cut-off that becomes, even though it explicitly defines an arbitrary part of a totality, "inaccessible and intraversable" (Schapiro 1952, 14). It stands for a vision, not for an objectual entity, and it uncovers someone's intimacy, related to the non-human essence of things. Schapiro observes that traditionally, still life paintings were concentrated on disparate objects reflected as static instruments, often related with a profession and thus, becoming "emblems of moral ideas, like vanity" (Schapiro 1952, 14). They used to be represented as predisposed to being manipulated, handled in accordance with their ontological inferiority, as tools or things that serve human creativity, needs or interests. Nonetheless, Cézanne removes this instrumentality as an inherent and obvious characteristic of objects. Fruits or skulls have "nothing of the formality of a human purpose, they are scattered in a random fashion and the tablecloth that often accompanies them lies rumpled in accidental folds" (Schapiro 1952, 14). None of these elements aspire to become an *objet d'art* in Cézanne paintings. In fact, the painting itself, as a whole, could hardly become such. Still life is about an approximated life of an individual who is surrounded by objects that seduce one's eye, not one's mastering. Therefore, an object painted by Cézanne does not communicate with the spectator. The same is true of the individuals painted by the Provençal artist. No transfer is permitted from a painted individual to a living individual. This is true even in the case of autoportraits:

> Even in representing himself, Cézanne often assumes an attitude of extreme detachment, immobilized and distant in contemplating his own mirror image. The suspended palette in his hand is a significant barrier between the observer and the artist-subject. (Schapiro 1952, 15)

One could say that Cézanne tried all the forms of representational reduction: he eliminated objects, individuals, even himself, from what a true impressionist painting could be. There is only one cluster of entities that could not be dismantled by Cézanne's impressionist introspection: social groups that gain, as contexts of the still life, suspensions of the human condition. By this, Cézanne's paintings become quite metaphysical. But his impressionist sociology, so to say, confronts multiple scenarios that should be attentively dismissed from the current analysis, in order to accomplish the objectives advanced at the beginning. It is not my aim here to prospect mythical or religious relics of Cézanne's artworks, although nudes – as a primitive expression of the human con-

[2] To be consulted M. Schapiro 1968; Schapiro 1994, 135–142, and Schapiro 1994, 143–151.

dition – offer such hermeneutical puzzles from this standpoint. My preference is rather rooted in social groups – constituted by adherence to ludic predispositions, beliefs and habits. The *Card Players* series offers such social assembly: it explores entertainment, and yet, it introduces competition. It feeds enthusiasm and yet it occasions deception. Cézanne's option was for a series of paintings that neglect "these familiar aspects of the card game" (Schapiro 1952, 16). The characters freeze in meditation, and each of them has an ideal "poker face". There is nothing that could be extracted from their silent attitude, or, as Schapiro rightfully noticed, "the players all concentrate on their cards without show of feeling" (Schapiro 1952, 16). We would have expected to transgress the fever of gambling, perhaps the ovations of the public, the hysteria of a winner. But Cézanne's card game is an open situation that conserves its logical advantage of depicting a world in which any possibility is still valid.

> They are grouped in a symmetry natural to the game; and the shifting relation between rules, possibility, and chance, which is the objective root of card playing, is intimated only in the silent thought of the men. Cézanne might have observed in an actual game just such a moment of uniform concentration; but it is hardly characteristic of the peasants of his Provence – their play is convivial and loud. In selecting this intellectual phase of the game – a kind of collective solitaire – he created a model of his own activity as an artist. (Schapiro 1952, 16)

In this series of paintings everybody has the chance to win. This is not optimistic or pessimistic, it is just logical. And yet, as Schapiro argues, you cannot get to someone's desire to win. I would rather call Cézanne a *situationist* – not in a political sense, but in a factual one. Interested in contexts that he can reproduce instance by instance, the Provençal artist succeeds in framing sequences of human–object interactions that produce exactly what the sequence spectator–artwork should: an aesthetic contemplation. This is the main reason for which Cézanne's painting has never been processual – no one should really care about the resolution of the card game, as long as each player counts for himself in a certain situation. This monadology – a chain of possible worlds harmoniously communicating throughout logical scenarios – is raised in Cézanne's paintings as a logical structure of this world that creates multiple impressions to someone only because it enjoys multiple situations.

"Cézanne preserves a characteristic meditativeness and detachment from desire. His tendency coincides with what philosophers have defined as the aesthetic attitude in general: the experience of the qualities of things without regard to their use or cause or consequence." (Schapiro 1952, 16)

Consequently, the quality of an emotion, the expectancy of an object to be engaged by an individual, the essence of a temporality that involves balance, teasing, waiting, represent some of the aesthetic stimuli for impressionist acting. Still life is, as it emerges from a card game, a set of these behaviours, attitudes and values that confront the free will of decisive thinking, the luckiness[3] behind arbitrary human options, or even the cohesion of a player's community that performs a particular model of sociability. Schapiro's remarks on the *Card Players* are developed around the idea that symmetry makes possible the still life. Norms and choices, standards and options reflect the double genealogy of any game. At the end, something that is fictional (the

[3] At a first glimpse, the philosophy of luck reflects a struggle between partisans of determinism and adherents of pure and absolute freedom. Nonetheless, from epistemic luck in theories of knowledge to moral luck in ethics, luck is considered through a philosophical framework. In what concerns this particular case of the *Card Players*, one must say that such a situation offers the context of a controllable opportunity: luck is provoked. You can buy a ticket lottery or you can play cards – regardless of what you choose, these options involve the possibility of luck. Needless to say, in the former case you enter into a contingent situation – you might win or not, the chances of winning are statistically lower than the chances of losing, whereas the latter case expresses a situation relative to intervening and environmental luck – you depend not only on your choices and strategy but also on the luck of other players. To be consulted on this topic: Pritchard 2015. An equal critical inquiry is reflected in Hales' arguments (2016, 490–508). Hales distinguished probability, modal and control views on luck, arguing that all three theories are irreparably defective since "these are the problems of lucky necessities, skilful luck, and diachronic luck" and this is why "a serious revaluation of the role of luck in philosophy is called for" (Hales 2016, 490).

experience of the game) and something that is real (the social interaction of the players) maintains the same duality as a principle of artistic inspiration. Nonetheless, Schapiro restricts himself to appreciating this inner constitution of the tableau without speculating on anything related to the final product: the painting representing a still life, which becomes, in my opinion, a meta-painting. The *Card Players* address the world of the game itself. Each character is a referent for an object – a set of cards, and for another individual – a competitor. The atmosphere of the game comes out of the affective allure of each character. The image is static because each of the players is unmovable, focused, concentrated, imperturbable. The interaction between tools and things, cards and game, with human beings, reflects, in an Impressionist discourse such as Cézanne's, a meta-painting. Impressionism is a still life by itself: it objectifies emotions, it concretises perceptions, it dictates what the eyes see. But, are there any consequences of addressing Impressionism as the still life inspired by the vital energy of human beings interacting with objects of different ontological dignities – from tools to things and other individuals?

Human(ised) still-life: *The Card Players* and the Impressionist human condition

Cézanne's The *Card Players* represents, at a first glimpse, a painting series inspired by Provençal lifestyles. Leisure and entertainment are transgressing the atmosphere of a card game that colonised modern art history from Caravaggio and Chardin. But unlike his predecessors, Cézanne aspired to design a scene of his own: beyond the Impressionist trademark, the authorship of these paintings evokes his interest in situations imposing tension, observation, judgement, that are slightly contingent. Before discussing each canvas by itself, some general remarks should be considered in order to understand why their assembly performs a still life devoted to the human condition. As the paintings are compositionally framed as frozen instants – from this standpoint, Cézanne's artworks are oriented toward depicting events rather than processes. In fact, their hermeneutical backgrounds reflect prolonged critical attitudes against authority, law, censorship. According to Padiyar, game players inspired different generations of art, from Picasso to Duchamp, from Surrealists to Cold War contemporaries, from defiants of Soviet and Fascist experiences to partisans of Fluxus, configuring a reactionary attitude in front of hypernormativity, control and dictatorship. This escape from the real world occasioned by the game, which remains, as Gadamer would say, a normative construction but voluntarily and ludically consumed, fed Cézanne's appetite for art seeking a playful dimension. Therefore, if art was not playful enough, the subject of card players succeeded in compensating for such a shortcoming. Implicitly, Cézanne demonstrated through such topics "a playful impulse in modernism" (Padiyar 2013, 12), that turned the entire era back to the play instinct rooted in human culture. We are *homo ludens* (Huizinga 2016), and more than that, we are, as Cézanne framed it, simultaneously transgressed by "*Ludus*, games that are subordinated to the disciplining framework of rules, *Alea* (games of chance – like the lottery), and *Ilinx*, games that pursue or end with vertigo, disorder and physical disorientation" (Caillois 2001, 12). Padiyar advanced an interesting argument at the edge of these remarks: the series of *Card Players* reflects a disjunction between *Ludus* – "but in the manner of *Ilinx*" (Padiyar 2013, 14), that describes the world of game (both normative and related to objects placed at the proper time in the proper order), and the aesthetic of a disordered process, that affects the options of players, from a hypothetical fatigue to a possible dizziness that any smoker suffers from. The *Alea* element is recycled by the technique of painting: Impressionism is, at its best, the emergent result of rational choice and accidental manifestation. Colours used to "fall like rain from the painter's brush" (Iversen 2013, 158), so impressionism is an opportunist painting, so to say. Then, many critics remarked, such as Schapiro did, that the characters of these series are not Provençal peasants – they inspire nothing loud, primitive, uneducated. The explanation for this static, meditative and disciplined atmosphere inspired by The *Card Players* lies in the power of the game to produce what Huizinga would call as "a civilising process" (Huizinga 2016). People subject themselves to certain rules and behave accordingly. The *Card Players* are, in a Foucauldian sense, individuals that have internalised the norms of game performativity and consequently react as normalised subjects. It would have been eloquent for the hermeneutical reconstruction of such a civilising process to know the order in which the paintings of this series have been created. T. S. Clark attempted a reconstruction, by following a hypothetical order, more or less justified. In his opinion, the first canvas was the midsize painting in The Metropolitan Museum. Now, deposited in the Barnes collection, this painting is the largest of the series ("4 feet 5 inches high and six feet wide"), and remarked for its "heaviness" and solidity ("solid as a rock"). According to Clark,

> Depth and solidity in Cézanne are not, even in a curtained interior, purchased at the price of loss of immediacy (or not when things go well): they come out of the flutter of sensation: this is the visual point. When in due course we face the question of actual massiveness and obstruction in the Card Players, particularly that of the fabric on the wall at right, the balance of qualities in Cézanne's brushwork – the decisiveness, the brevity, the raciness however grave the colour – will make everything hard to decide on. (Clark 2013, 15)

Right after, the series is continued by canvases depicting two players, the smallest group from Cézanne's paintings: the effect of scale is more modest, and yet "monumental" (Clark 2013, 16). This series has the capacity to drive the spectator's eye from tangible to intangible contents, from the position of cards, pipes, legs, shoulders, hats, to the emotions of indecisiveness and strategic thinking that any card player engages in. In fact, spaces are used in Cézanne's paintings to dispense objects and individuals from certain moments, and to question them by themselves. It is not the card player that is important here, but the human condition that he represents.

> (…) all five Card Players (…) surely meant as an image – a tremendous synecdoche – of pictorial compactness and horror vacui. (Clark 2013, 16)

Clark remarks on Schapiro's understanding of a dialectical "to and from between the Platers", entertained by "their opposites of colour and drawing", a dialectic that contributes to the solidity of each sketch. Nonetheless, there is also an affected symmetry, suspended, by what Clark observed as traces of miniaturisation, marooning, tiny disorders that reflect the stroke of genius: "the way the right-hand player is tipped and swivelled just enough for symmetry to be broken" (Clark 2013, 17). As Clark describes the five Cézannes, he reaches a particular semantics of *Card Players*, a series of paintings reflecting separate individuals who perform togetherness, in a closed space. The room and the game are blended in "a stiffed reciprocity" (Clark 2013, 17). They both depict enclosed realities that the spectator has to pursue, seduced by objects and individuals that fit into this cut-off reality. Cézanne is, as we all know, hypervisual. He performs "phenomenal events" (Clark 2013, 19) of the eye[4]. So, on what grounds do the spaces that host the card game represent an ideal in order to allow the sense of living, sociability and conviviality that Cézanne attempted to unveil here? In order to answer such a question, Clark engages a challenging parallel between the series of *Card Players* and Cézanne's *Paul Alexis Reading to Emile Zola*, from 1872. The table divides the inner world of that room inhabited by two intellectuals that are quite static. The curtain seems to be a portal for any spectator that perceives from backstage the interaction between poetry and novel. The legs of the two characters develop their own geometry and colours support an opposition not only between characters, but also between cultural genres represented by them. However, Clark considers that this impressionist territory inspires the *Card Players* in terms of "picturing room and sociality" (Clark 2013, 20). Rooms are not important here, but interiors: Cézanne needs a dark space – a black wall – to support a world of colour, so this is why backgrounds are habitually "interrupted, unstable, perforated". Zola and Alexis are placed in a "Ur-form" that is recycled in a sketch inserted by Cézanne in a letter sent to Zola in January 1859, read by Clark in

[4] Marion considered that regarding the theory of genius in Modernity, such a model is represented by Cézanne and his capacity to engage a mutuality between painting and spectator, from seeing to receiving ("*voir*" and "*recevoir*") throughout the aesthetic experience senses, values and emotions. Therefore, I consider that this double visibility is what might support a vitalism that transgresses Cézanne's paintings. To be consulted *Thomas-Fogiel* 2013/4, 527–548, and Marion 2004.

a psychoanalytical key. Two days before Cézanne's twentieth birthday, he created this drawing that organises a spiritual world around a table, gathering members of a mystical, "ceremonious cannibalism", around a skull.

> The two men in the door-way are Dante and Virgil, and the figure at the right of the table is The Father – inviting his children and grandchildren to make a meal. Badt was struck by the similarities between the cannibal feast and the Barnes Card Players – the same rigid symmetry of the father and son, plus their mirror-image legs underneath the table, the same transfixed children, even the same picture on the wall. (Clark 2013, 27)

The fact that Cézanne and Zola are Dante and Virgil is fantastic and invites both cultural eulogiums and personal confessions. But there are two elements that are lacking in Clark's analysis and that I find relevant for reconsidering the success of The *Card Players* and the potential of this series to prescribe a still life of the human condition. First of all, it is not the curtain or the door that should be identified as incentives or visual stimuli for impressionist aesthetic experience, but rather the table that has the power to gather individuals, to perform an *axis mundi* and to develop something in common between them, either material or spiritual.

Secondly, Clark notes that both spiritism sessions or cannibalistic ceremonials and card games are means that favour human interaction with death. I consider the opposite: they invite a vital energy that fits Impressionism. You cannot track what you do not know – that's the principle of Impressionism – and here, the eye must focus on what is familiar, meaning life, at its best. Sociability would not be possible in the absence of such vitalism. In the name of this empowered life, we would expect characters to move, to escape the geometry of their hands and knees, to shake their heads, to raise a brow: impression comes from an expected animation that is intimately and inexorably connected with life. This is particularly the reason for which I strongly consider that phenomenologically Cézanne can be apprehended only as long as Merleau-Ponty's experiential ambiguity addressed to the Provençal impressionist is connected with Bergson's *élan vital*. The table is in front and between things – it epitomises the painting's geometry and supports a distance: between spectator and painting, between characters themselves, between objects themselves, between characters and objects–in the end, between forms of life, embodied by such games. It creates proximities, vividness and humanness, that are unveiled by the eye, in its impressionist behaviour.

Card games are a metaphor for life on its path towards civilisation. On the other hand, this vitalism that seduced Cézanne's paintings is, from the beginning, loyal to stoic, "anti-sentimental forces", as Dombrowski argued (2013, 36). Contemporary to the *Card Players* of Cézanne was Meissonier's canvas, conceived quite mimetically, in my opinion. It reproduced curtains, windows, tables and an inner space that had corporal positions of characters performed within a very exact patented geometry. One would say that such a topic was fashionable because it reflected the routine of that time. Nonetheless, there is something deeper, related with a hygiene of life: card games teach you how to master your self-control, being-in-space and being-in-use (a distinction advanced by Dombrowski that I cherish much, although he fails, in his turn, to provide a comprehensive overview of Cézanne's vitalism emerging from this puzzle still life). The *Card Players*, regardless of their version and authorship, prophesise "the end of scientific perspective": there is a visual code that cards develop around the players, creating a relationship between eye and body, hand and leg, table and window, human and object. The solidity remarked by Clark feeds the three-dimensionality noted by Dombrowski, and they both are intermediated and expressed by the game of cards as a modern phenomenon. Both the topic and the representations are pathologically modern: the individual manipulates objects and invests them into routines and pleasures that subject themselves to a cultural synergy. While Dombrowski succeeds in recognising, step by step, the modern symptomatology of Cézanne's paintings – allegorical themes, post-perspectival spaces, the large-scale pathos of representations – he tremendously fails in understanding the *Card Players* as a still life. Rather, he considers such representations "as opposed to still-life and landscape", from the beginning taking up Clark's position of clarifying the particularities of the "Ur-Cezannian problem" that create a material metaphor for "the optic and the haptic" (Dombrowski 2013, 38). The main reason for which he lacks the ability to identify the card game as a still life is inspired by the relationship between the pictorial spaces and properties of painting, a view inherited from Novotny, whom he attentively quotes.

Dombrowski reiterates this distinction advanced in an essay from 1938 (so from a post-Cézannian period), that addressed "a tense cohabitation in Cézanne's works between the *Darstellungswert* (the value/order of the represented world) and the *Bildwert* (the flattened image order inherent in the medium of painting)" (Dombrowski 2013, 38). Card games facilitated both but, according to Dombrowski, "in card play, the depicted orders achieve a kind of mutuality and contemporaneity rare in other spheres of life" (Dombrowski 2013, 38). My argument is that card games are meant to reproduce such mutuality from real life. Cards reproduce a pictorial space that inspires the painting as such; they have their own narration, a fictional and symbolic world prescribing models of community – embraced by cards alike – and models of negotiation, exchange and conviviality – that a full package of cards depicts. If we accept the definition that a still life explores commonplace objects – either natural or manmade, then the *Card Players* performs one by double means: by cards and other tools placed in the chamber inhabited by players, and by individuals who are converted into "natural" objects. Impressionism has this greatest quality to dismantle animation into sequences, so that each representation of the series of *Card Players* is a static situation from a larger process, that of playing a game. Spaces are invested in this technique in order to "recover" what is "unstill": the geometry of legs suggests that one of the players was almost moving, as he was going to choose and throw a card; the distribution of cards reveals that the game is at its beginning and someone is about to perform a first movement. Therefore, *The Card Players* are a series of still life paintings because gaming is instrumentalised in order to apprehend a static situation prior to a movement. Technically, this series is not the sum of landscape, still life and portrait techniques. It is a still life by itself, because in these particular cases, Cézanne converts the individual into an object in order to perform a painting not about players, but about the human condition. Even though the context – the card game – is prominent, in fact, he de-objectified and decontextualized the painting. Portraits engaged in a card game are subjected to a phenomenological reduction; if you remove everything – cards, pipes, hats, individuals, curtains, tables, you remain with a general, sensitive topic: the human condition. As a complex still life, the *Card Players* expose a serial perspective on the ontology of subjects and objects: portraits are a function for such elements. Cards by themselves are not as important as we think at a first glimpse. They are rarely exposed directly in Cézanne's paintings, and when they are saved from a blurry representation, they seem to reflect warm colours – that anticipate their cold, dark counterparts in the opponent's hand. This is part of what Merleau-Ponty (2003) would recognize as *an experiential ambiguity*. You can conclude at most that cards expect to be paired, meaning that one gamer needs the other one just as one human needs another one in order to fulfil his or her condition:

> "Part of the attraction of card play for Cézanne certainly lays in such metaphoric potentialities of the game: a system of rules and regulations utterly self-contained and separate from workaday life ("a complete miniature society in itself") yet precisely because of the separation providing an essential metaphor for it (life is so like a rubber at Whist)." (Dombrowski 2013, 45)

Sociability and solidarity are moral drives of card-play in Cézanne's paintings, but there is also something that Dombrowksi remarked by recourse to a particular exegesis and that confirms my theory on the Provençal vitalism of such impressionist artworks: "A surplus of energy" that comes from any play for animals and human beings, that activates instincts of survival, of not getting lost. Here, Groos's psychology (1898) and Schiller's "play drive" as the expression of human freedom and education, developed in his *On the Aesthetic Education of Man* (1794), are more than eloquent to support the need of Cézanne to bring back the experience of playfulness to the *élan vital*. Moreover, in order to play something, you must be disposed to waste time, as well as "energy", "and often money" (Caillois 2001, 6), as Caillois would put it. The *Card Players* depict playfulness "as remoteness from life" (Dombrowski, 2013, 48), for which one has to accept the vitalism transgressed by the sequence of gaming.

It is the same energy that will destroy the temporary order of the micro-still life from the painting: someone will take a pipe, will move a card, will shake hands. The spatial properties of things will be confronted throughout a new proximity that the human subject will create by his actions. The table has its own tableau, but it is just a piece of a larger puzzle in which individuals become objects for our perception. Cards are either in players' hands or at the table, pipes are everywhere. They flood the space and float as in Magritte's tableau-object, when the initial painting of a pipe becomes object for a second painting that encapsulates the former and negates it by performing both experiential and semantic ambiguities – *Ceci n'est pas une pipe* (Foucault 1976, 6–21). Impressions are not emerging solely from objects by themselves, but also from their numerology – once there used to be five pipes, now they are only four. Here, Cézanne creates his meta-painting as well (asMagritte did later on). He paints cards that usually are printed:

> Transformation cards are ontologically different from painting – printed on paper, not painted on canvas – but that does not mean that a painter like Cézanne (who, as mentioned, relished copying in paint fashion plates and bottle labels) might not have been open to a more cross-fertilization between high and low than is usually assumed. (Dombrowski 2013, 60)

Cézanne takes a look at the social register of cards and makes it visible – especially in those representations where there are more than two players. Wherever the table is surrounded by people, cards become visible, as a moral subterfuge for the idea that one has to distinguish himself from the crowd, from the *hoi polloi*. In a crowd you need to gain a space that is invented by the significance of cards and their symbols or colours. At the end, forms of life and forms of art intertwine in Cézanne's series of paintings, performing different registers of visibility, cohesion and intimacy, that all together compose a still life.

One significant observation comes from Parkinson, who recalls Picabia's *Still Lives* from 1920, that engaged Cézanne's name as a creator of collage, a highly seductive technique to the Dada artistic world. But "the merrily wanking, cheerfully waving stuffed monkey" (Parkinson 2013, 91) depicting, as the discourse inserted in the tableau said, le "Portrait de Cézanne", "Portrait de Rembrandt", "Portrait de Renoir" as "Natures Mortes" was not accidentally connected with still life. The evolutionist argument – we have evolved, as modern artists, from impressionist monkeys to vertebrate avant-gardists – has a point: those names, those biographies, those human beings are still life expressions from the history of art, this "cemetery-museum" of our civilisation, as the Futurists recognised it later. Breton's reaction to Cézanne – that of repugnance – is well-known, but there would have been room for a fruitful debate on the relationship between the automatism of the eye engaged by Impressionism and the automatism of the consciousness engaged by Surrealism. For Surrealists, our inner emotions, dreams and discharged desires are deeply submerged by reason. They are invited to remain static in our subconsciousness, in order to be discovered and empowered by art. In short, all those Freudian undergrounds of our consciousness are a still life for our reason, and Surrealists pleaded for their liberation. They were *Cézannians* if we leave aside their critique of the so-called Catholic bourgeoisie of Provençal artists, who in fact, painted more peasants than aristocrats. However, Cézanne seemed to turn back to the history of the arts at the beginning of the 20th century:

> The conjecture gains some credence from Breton's estimation that around The Card Players floats a half tragic, half guignolesque menace in every point resembling the one pictured in the card game of Chaplin's film, A Dog's Life. (…) Given Breton's frequently expressed fascination with the curtain as a metaphor, we might assume that his remark about the other painting with a halo subject by

> Cézanne, The Young Man before a Skull, in which metaphysical unease falls on the painting through the pleats of the curtain, was the memory trigger that brought that not very menacing scene from A Dog's Life inaccurately back to his mind. (Parkinson 2013, 92)

Absolutely it was a *halo* around the *Card Players*, but it came from the situational still life that each painting in part performed. Plus, the halo works as an aura for his dubitative painting, as Merleau-Ponty would put it. This series of card players encapsulates both the archetype and the most modern expression of still life. History is brought to the centre, on the table, and then modern attitudes surround the table, experiencing entangled human attitudes and tools that together offer a new experiential phenomenology for the act of perception, a new world for ludic and quotidian actions, in short, a new aesthetics that confirmed the paternity of Modernity to Impressionism.

Padiyar was right when he saw in Arendt's *The Human Condition* (1958) a silent remark on Cézanne's manner of playing with things to evoke individuals. His still life was like Arendt's "human corner" where each of us reveals being "happy among small things, with the space of their own four walls, between chest and bed, table and chair, dog and cat flowerpot" (Padiyar 2013, 133), an intimacy that we cherish only after being suffocated by industrial and technical intrusions of modern society in our lives. Cézanne's paintings reflect the phenomenology of togetherness supported or disturbed by things that improve solidarity, acting and conviviality. We all live in a house of cards, but not like Italo Calvino's one. Human beings are those who raise the castle in Cézanne's painting. Their attention, their devotion, their indifference, their seductions are part of this human fortress, or corner, that anyone needs in order to preserve natural instincts, that would have made Nietzsche more than happy. This still life made up of passions, risks and luck varies in Cézanne's series of paintings that gains its universality – its *aura* or *halo* that Benjamin (1969) and Breton (1992) identified – because it brings to the same table adults and children, sinners and innocents, smokers and non-smokers. Childhood is part of this still life, so is adulthood: together they depict this human condition, an impressionist meditation on life, solidarity and sociability, a winning card for the modern painting that only Cézanne had in his hands.

Conclusions

Was Cézanne's art more fertile for philosophical reflection than that of Van Gogh's? This remains an open question that does not aim to create an epochal competition, nor to divide the realm of Impressionism between winners and losers of the phenomenological turn in painting. However, we might discover that as Cézanne remained at the heart of Merleau-Ponty's analysis (whereas Heidegger embraced Van Gogh's perspective), Impressionism has never been discussed properly as an aesthetic approach to the human condition. In fact, perception has been from the beginning the exclusive object of interest for philosophical reflection when it comes to Impressionism, the human condition being addressed in these terms implicitly and unsatisfactorily. One of the major outcomes of this incursion is that it provides a new philosophical framework for understanding the topic of the human condition through still life, a concept that extends its canonical properties by encapsulating objectified human actors. On the one hand, the role of this critical inquiry has been to open the philosophical discussion of Impressionism to a Bergsonian phenomenological approach, a narrative that competes with that advanced by Merleau-Ponty, overcoming the tradition of "experiential ambiguity" by imputing to the Cézannian approach an implicit *élan vital*. On the other hand, still life has rarely been considered as meta-painting: this hypothesis facilitates a better arbitration of animated and nonanimated subjects and confirms why Impressionism is not solely an artistic paradigm that tests the phenomenology of perception, but also an applied study on a historical preoccupation, that of defining and tracking the evolution of the human condition.

References

Arendt, H. (1958). *The Human Condition*. Chicago: The University of Chicago Press.

Benjamin, W. (1969). The Work of Art in the Age of Mechanical Reproduction.In H. Arendt (ed.), H. Zohn (transl.) *Illuminations.* New York: Schocken Books.

Bergson, H. (1998). *Creative evolution.* Mineola, New York: Dover Publications.

Breton, A. (1992). *Œuvres completes*, *tome II*, Édition de Marguerite Bonnet avec la collaboration d'Étienne-Alain Hubert et José Pierre. Paris: Gallimard.

Caillois, R. (2001). M*an, Play and Games*. Champaign, IL: University of Illinois Press.

Clark, T. J. (2013) "A House of Cards." In S. Padiyar (ed.) *Modernist Games. Cézanne and his Card Players* (pp. 15–34). London: The Courtauld Institute of Art.

Dombrowski, A. (2013). "The Cut and Shuffle. Card Playing in Cézanne's Card Players." In S. Padiyar (ed.) *Modernist Games. Cézanne and his Card Players* (pp. 35–68). London: The Courtauld Institute of Art.

Foucault, M. (1976). "Ceci n'est pas une pipe." In *October,* MIT Press, (1): 6–21.

Groos, K. (1898). *The Play of Animals,* translated by E. L. Baldwin. New York: Appleton.

Hales, S.D. (2016). "Why Every Theory of Luck is Wrong." In *Noûs,* 50: 490–508, https://doi.org/10.1111/nous.12076.

Hales, S. D. & Adrienne Johnson, J. (2015). "Luck Attributions and Cognitive Bias". In Duncan Pritchard & Lee John Whittington (eds.), *The Philosophy of Luck. Hoboken* (pp. 59–77). New Jersey: Wiley-Blackwell.

Huizinga, J. (2016). *Homo Ludens: A Study of the Play Element in Culture.* Kattering: Angelico Press.

Iversen, M. (2013). "Chance and Choice in Cézanne." In S. Padiyar (ed.) *Modernist Games. Cézanne and his Card Players* (pp. 151–162). London: The Courtauld Institute of Art.

Krell, D. F. (ed.) (1977). *Martin Heidegger: Basic Writings*. New York: Harper & Row.

Marion, J. L. (2004). Réduction et Donation. Paris: Puf.

Machiavelli, N. (2014). *The Prince.* London: Penguin Books.

McKinnon, Rachel (2015). You Make Your Own Luck. In Duncan Pritchard & Lee John Whittington (eds.), *The Philosophy of Luck* (pp. 107–126). Hoboken, New Jersey: Wiley-Blackwell.

Merleau-Ponty, M. (1964). ye and Mind. In James E. Edie (ed.) *The Primacy of Perception* (pp. 159–190), Trans. Carleton Dallery. Evanston, IL: Northwestern University Press.

Merleau-Ponty, M. (2003). *Phenomenology of Perception.* New York: Routledge.

Padiyar, S. (2013). "Building a World Between Men (Or Cézanne and with Arendt). In *Modernist Games. Cézanne and his Card Players*. London: The Courtauld Institute of Art, 123-150.

Padiyar, S. (2013). Introduction. In S. Padiyar (ed.) *Modernist Games. Cézanne and his Card Players* (pp. 11–14). London: The Courtauld Institute of Art.

Parkinson, G. (2013). Greengrocer, Bricklayer or Seer. Surrealist Cézanne. In S. Padiyar (ed.) *Modernist Games. Cézanne and his Card Players* (pp. 69–100). London: The Courtauld Institute of Art.

Pritchard, D., Whittington, L.J. (2015). *The Philosophy of Luck.* Oxford: Blackwell Publishing Ltd.

Schapiro, (M). (1994). The Still Life as a Personal Object – A Note on Heidegger and van Gogh (1968). In *Theory and Philosophy of Art: Style, Artist and Society* (pp. 135–142). New York: George Braziller Inc.

Schapiro, M. (1952). *Paul Cézanne.* New York: Harry N. Abrams.

Schapiro, M. (1994). Further Notes on Heidegger and van Gogh. In *Theory and Philosophy of Art: Style, Artist and Society* (pp. 143–151). New York: George Braziller Inc.

Thomas-Fogiel, I. (2013). La tournure empiriste de la phénoménologie française contemporaine. *Revue philosophique de la France et de l'étranger*, *138* (4), 527–548,
doi: 10.3917/rphi.134.0527,
URL: https://www.cairn.info/revue-philosophique-2013-4-page-527.htm

KANT'S HIGHEST GOOD AS A WIDE OBLIGATION AND ITS NORMATIVE GROUND

Neşe Aksoy

Institute of Philosophy and Sociology, Bulgaria
Bulgarian Academy of Sciences
email: aksoynesee@gmail.com
ORCID ID: https://orcid.org/ 0000-0002-3815-4350

Abstract:

In his Critical corpus, Kant makes two seemingly inconsistent claims concerning the highest good and its relation to the postulates of immortality and God. On the one hand, he argues that the highest good is a duty to be promoted that must therefore be possible by human powers ('ought implies can'). On the other hand, he asserts that the highest good is an "unconditioned object" of practical reason that can only be attained on the ground of the postulates of immortality and the aid of God. In order to resolve this apparent conflict in Kant's claims, I intend to argue in this paper that Kant implicitly classifies the highest good as a wide obligation (imperfect duty), which is to be promoted in the world by human agency but goes beyond the limits of human powers and must be assisted and complemented by the postulates of immortality and God. To do so, I argue that the highest good stems from the moral law since the moral law embodies the ground of absolute goodness and holiness that is never fully achievable by rational agents and thus can generate duties that are impossible to attain. In parallel, I argue that two of the components of the highest good, namely moral worthiness and distribution of happiness, are classified as wide obligations that the moral law commands us to pursue infinitely and with the aid of God. In sum, I try to resolve the initial conflict by suggesting that the highest good as an infinite progress towards the holiness demanded by the moral law signifies a wide obligation, which is not achievable merely by human agency but fundamentally requires the postulates of immortality and the existence of God as a complementing and assisting ground.

Key words: the highest good, ought implies can, immortality, holiness, the aid of God, the moral law.

1. Introduction

Kant's theory of the highest good (*summum bonum*) and its relation to the postulates of pure practical reason can be regarded as the culminating subject of his entire Critical philosophy. However, in his Critical corpus, as regards the connection between these two conceptions, Kant seems to be making two apparently conflicting claims. On the one hand, he argues that it is a duty to promote the highest good, and that it must therefore be possible in the world by human powers (*CPrR* 5:125).[1] On the other hand, he asserts that the highest good is "an unconditioned object" of practical reason that can never be fully attained in the sensible world but can only be accomplished in an eternal life and with the aid of God (*CPR* A811/B 839 and A 829/B 857; *CPrR* 5:108, 5:122 and 5:124; *Rel* 6:98). Since the first premise suggests that the highest good is a moral obligation to be performed with human efforts (complying with the 'ought implies can' principle), the second premise seems to be in apparent contradiction with it, as it entails that the highest good requires superhuman elements for its total realization.

[1] The abbreviations of Kant's works are as follows: *CPR*: Critique of Pure Reason; *G*: Groundwork of the Metaphysics of Morals; *CPrR*: Critique of Practical Reason; *CPJ*: Critique of the Power of Judgment; *LE*: Lecture on Ethics; *Rel*: Religion within the Boundaries of Religion; *MM*: Metaphysics of Morals; *CF*: Conflict of Faculties; *TP*: On the Common Saying; and *LPDR*: Lectures on the Philosophical Doctrine of Religion.

This apparent conflict gives rise to various interpretations that aim to resolve it in their own ways. Some commentators such as Silber (1959), Yovel (1980), Wike (2010) and Pasternack (2017), etc. offer a two-aspect reading of the highest good, which positions the highest good as a moral duty to be promoted in the world through human efforts, consistent with the principle that 'ought implies can', although its full realization is a matter of eternal life that requires the aid of God. According to this two-aspect reading, although the highest good is not realizable in the world in full, it is still a duty to promote it in as much as it lies within our power to do so. Another group of commentators, such as Villaran (2013), Zeldin (1971), Wood (1970), and Denis (2006), on the other hand, tries to resolve the above-mentioned conflict by rejecting the idea that the highest good is a moral duty that complies with the 'ought implies can' principle, maintaining instead that it is a special type of duty in kind. According to this account, the highest good oversteps the principle of the 'ought implies can' since it requires the postulates of immortality and God as a necessary ground for its real possibility.

In this paper, engaging these two lines of interpretation, I will argue that the highest good is not a strict and narrow duty (or perfect duty) that we are absolutely and categorically obligated to achieve. Instead, I will suggest that the highest good is a wide obligation (similar to imperfect duties) that we need to strive to realize, but may not fully accomplish at any point in this life, because it oversteps the limits of our human capabilities (extending beyond the 'ought implies can' principle). To do so, I will follow the following structure. In the first section, I will give a brief overview of the dominant approach in the literature, which aims to resolve the above-mentioned conflict by suggesting that the highest good is an imperfect duty that either satisfies the 'ought implies can' principle or oversteps it. In the second section, I will claim that two of the components of the highest good, namely worthiness to be happy (or virtue) and distribution of happiness, stand beyond human capacities. I will primarily point out that, although moral worthiness to being happy and distribution of happiness require moral striving in the first place, it is not possible to achieve either of them without the presupposition of eternity and the assistance of divine agency since rational agents are limited and imperfect by nature. However, as I will note, this idea can easily lead to the problem of heteronomy. Therefore, in an attempt to avoid falling prey to heteronomy, in the third section, I will propose that the highest good fully stems from the moral law in a normative sense since the moral law embodies the law that is absolute and demands holiness, which can thus give rise to duties that are not achievable by human efforts. On this basis, I will argue that the highest good is a wide obligation (or imperfect duty) that normatively stems from the moral law but, targeting the absolute holiness of the moral law, it extends beyond the boundaries of the 'ought implies can' principle and ultimately gets grounded on the postulates of immortality and God. In doing so, I will attempt to resolve the above-mentioned inconsistency by suggesting that the duty to the highest good does not conflict with the idea that the postulates of immortality and God are necessary for its realization, because the highest good is not a strict and narrow duty that lies within our power to realize, but rather is a wide obligation (or imperfect duty) that we ought to pursue to infinity and by appealing to the assistance of God as a complementary principle.

2. The Highest Good as a Duty to Promote and 'Ought Implies Can'

In numerous places such as the *Critique of Practical Reason*, *Critique of the Power of Judgment* and the *Religion within the Bounds of Bare Reason*, Kant claims that the highest good (*summum bonum*)[2] as the ultimate end of morality, which comprises the synthesis of virtue (*Tugend*) and happiness (*Glückseligkeit),* is a moral duty (or categorical imperative) to be promoted:

> If, therefore, the highest good is impossible in accordance with practical rules, then the moral law, which commands us to promote it, must be fantastic and directed to empty imaginary ends and must therefore in itself be false (*CPrR* 5: 114, 92);
> … we ought to strive to promote the highest good (which must therefore be possible) (*CPrR* 5:125, 101);
> … Yet the aim of promoting the final end of all rational beings (happiness, insofar as it is con-

[2] Kant identifies the *summum bonum* as the archetypal ideal that all ancient ethical systems were founded on (*LE* 27:248, 44). According to Kant, in ancient philosophy there were three forms of *summum bonum*: the Cynic ideal, the Epicurean ideal and the Stoic ideal (LE 27:248, 45). Kant writes that the *summum bonum* in all its versions embodied the ideal of moral perfection that constitutes the ultimate end of morality.

> sistent with duty) is still imposed precisely by the law of duty (*CPJ* 5:471, 335);
>
> Now, here we have a duty *sui generis*, not of human beings toward human beings but of the human race toward itself. For every species of rational beings is objectively – in the idea of reason – destined to a common end, namely the promotion of the highest good as a good common to all (*Rel* 6:98, 109).

In view of this, it is possible to put it simply that the highest good is commanded by the moral law, that is, it is grounded in the absolutely free and self-legislative power of human reason. Furthermore, in the second statement, Kant argues that rational agents are duty-bound to promote the highest good and that it must, therefore, be possible for it to be performed within the limits of human capabilities. So, Kant clearly points to the compatibility between the duty of the highest good and the principle of 'ought implies can.' [3] As he suggests, the highest good must be practically possible to the extent that human capacities can achieve it since it is a command that derives from moral law.

Although this idea that the highest good is an ordinary moral duty that satisfies the 'ought implies can' principle can be traced in Kant's work, it is apparently contradicted by other claims that Kant makes in his Critical corpus which suggest that the highest good does not necessarily conform to the principle of 'ought implies can.' For instance, in *Religion*, Kant explicitly claims that the highest good differs from other moral laws in kind and in principle because "it is the idea of working toward a whole [a system or unity of well-disposed human beings] of which we cannot know whether as a whole it is also in our power."[4] Hence, for Kant, the highest good as a duty to produce a unity of human beings who are morally worthy of happiness is not a common task to fulfill, for one can never be sure of the morality of others. Kant further adds that the highest good is different from other moral laws in a fundamental way since it requires the presupposition of another idea, namely, of a higher moral being "through whose universal organization the forces of single individuals, insufficient on their own, are united for a common effect" (*Rel* 6:98, 109). Though Kant here essentially refers to the highest good in a collective sense, his primary concern is to show that the highest good as a goal for human history is not achievable without the assistance of God. And again, in *Religion* 6:7n, Kant writes that the highest good is not a regular moral duty that analytically derives from moral law but rather is a synthetic *a priori* proposition that "exceeds the concept of duty that morality contains."[5] According to Kant, since the highest good demands the presupposition of a superhuman being or a moral lawgiver (or God) for the possibility of the synthesis of virtue and happiness, it exceeds the scope of ordinary moral duties that are merely grounded on moral law.[6]

That the highest good is controversially regarded by Kant as a duty that follows from moral law but extends beyond human capacities leads us to identify a dominant approach in the literature to interpret the status of the highest good as a special type of duty. Commentators such as Silber (1959), Yovel (1980), Pasternack (2017), Wood (1970) and Villaran (2013), etc. suggest that the highest good is not an ordinary moral duty but an imperfect duty that we are obligated to pursue although it transcends our powers. John Silber, as one of the prominent representatives of this approach, says that the highest good is a duty to be promoted to the extent that human powers are able (the immanent highest good)[7], and it gains a normative ground through the transcendent highest good that has a regulative impact on the promotion of the immanent highest good[8], suggesting that the highest good is an imperfect duty to be realized by human powers and by an appeal to the ground of immortality and God. Yirmiyahu Yovel, in his book *Kant and the Philosophy of History*, argues that the highest good is to be promoted to the extent that it is possible by human efforts with reference to the transcendent idea of God and immortality.[9] Yovel argues that it is only after human agents deserve moral worthiness on their own that they can appeal to God, which provides the ground for hope (Yovel 1980, 97). By this, Yovel points out that divine assistance will only "encourage and sanction human activity" rather than eliminate

[3] Generally speaking, the principle of ought implies can' entails that obligations and abilities mutually depend upon each other. Simply put, it holds that duties are determined by the possibility of their being performed by human powers. However, 'ought implies can' as the epistemic principle of Kant's theory of duty is interpreted in subtly different ways. For the purposes of this paper, see Kohl 2015, and Stern 2004.

[4] *Rel* 6:98, p. 109.

[5] *Rel* 6:7n, 35-36.

[6] For a detailed discussion on the extension of the highest good beyond the moral law, see Kleingeld 2016 and Zeldin 1971.

[7] Silber 1959, 478.

[8] Silber 1959, 491.

[9] Yovel 1980, 96-97.

their moral action (Yovel 1980, 97). By offering a wide variety of textual evidence to support his claim, Pasternack lays out several reasons to favor the idea that moral worthiness of being happy complies with the Kantian principle that 'ought implies can.' Pasternack writes that despite the need for divine assistance in the total realization of the highest good, moral worthiness is a duty that is consistent with the 'ought implies can' principle because divine aid is compatible with active moral agency (Pasternack 2017, 462-463). Alonso Villaran, on the other hand, relying on Stern's account,[10] argues that the highest good is a "fallible duty" that one can fall short of realizing but still would not be blameworthy for not being able to accomplish it (Villaran 2013, 37). According to Villaran, the highest good is not an ordinary moral duty (or an infallible duty) that we would be blameworthy for failing to do. Instead, he maintains that Kant has a softer view of 'ought implies can' with regard to the duty of the highest good, that is, we are not duty-bound because the highest good is achievable (it is within our capacities and capabilities), but we are duty-bound to promote the highest good although we would not be blameworthy for failing to realize it. According to Villaran's account, Stern's "soft sense of the highest good" offers a better way to understand the highest good as a fallible duty whose accomplishment lies beyond our control.

Having briefly outlined the main approach in the literature that aims to expound the status of the highest good in regard to the 'ought implies can' principle, it is worth noting that the highest good is widely regarded to be an imperfect duty that goes beyond human powers. In the next section, I will consider the nuanced details of these accounts by laying out the intersections between them, along with my primary goal of examining the highest good in reference to moral worthiness and the postulates of immortality and God.

[10] In his article "Does 'Ought' Imply 'Can'? And Did Kant Think It Does?", Stern argues that Kant has a weak sense of the 'ought implies can' principle rather than a strong version of it. According to Stern's view, Kant adopts the principle of 'ought implies can' in a weak sense, by arguing that the moral law is obligatory for us because we have the capacities and conditions to act upon it. But, for Stern, this is different from the claim that moral actions are not right (rather than just obligatory for us) if we cannot perform them, which is how the strong sense of the 'ought implies can' operates (Stern 2004, 57).

3. The Highest Good, Moral Worthiness and the Postulates of Immortality and the Existence of God

When I proceed to Kant's second claim, which leads to a conflict with the former one (as I have identified at the outset), it is vital to recall Kant's main assertions in the *Critique of Practical Reason* which follow as:

1. As pure practical reason it likewise seeks the unconditioned for the practically conditioned (which rests on inclinations and natural needs), not indeed as the determining ground of the will, but even when this is given (in the moral law), it seeks the unconditioned totality of the object of pure practical reason, under the name of the highest good" (*CPrR* 5:108, 87–88);
2. … since [morality or virtue] can be fully accomplished only in an eternity, [the highest good] led to the postulate of immortality (*CPrR* 5:124, 100);
3. … [proportionate distribution of happiness to morality] must lead to the supposition of the existence of a cause adequate to this effect, that is, it must postulate the existence of God (*CPrR* 5: 124, 100).

In view of this, it is key to recognize that Kant conceives of the highest good as necessarily grounded on the postulations of the immortality of the soul and the existence of God since: 1. morality is an unconditioned good that can be accomplished only through the postulation of immortality, and 2. the distribution of happiness in proportion to virtue can only be practiced through the aid of divine agency (postulation of the existence of God). That Kant evidently refers to superhuman elements in the realization of the highest good suggests that the highest good does not simply follow from moral law and comply with the 'ought implies can' principle; instead, it goes beyond these. In this section, dividing the highest good into two components, namely worthiness to be happy and distribution of happiness, I will try to show that each of the components of the highest good has a reference to the superhuman elements of the immortality of the soul and the existence of God, and thus, it seems that the highest good has no direct compliance with the 'ought implies can' principle in either sense.

3.1. Worthiness to Be Happy (Virtue or Moral Perfection)

As the prime requisite for the highest good, virtue (or worthiness to be happy) is denoted in the *Cri-*

tique of Practical Reason as the "supreme" (*supremum*) condition of the highest good since it is an "unconditioned good" in character which is not subordinate to any other condition (*CPrR* 5:110, 89). Defining it as an "unconditioned good", Kant further argues that virtue or worthiness to be happy requires "a complete conformity of dispositions with the moral law", which he calls as "holiness", and he contends that it is the ultimate goal of morality (*CPrR* 5:110, 89). In the *Metaphysics of Morals*, Kant similarly writes that we are duty-bound to be "holy", but it is not a "perfect" and "narrow" duty that we are supposed to accomplish; rather, it is a "wide" and "imperfect" duty, a duty that we need to strive to attain but not necessarily have to reach in this life (*MM* 6:447, 241). For Kant, the complete attainment of holiness is not achievable by finite human agency, but instead virtue or worthiness to be happy (to be holy) embodies an "endless striving"[11] or "continual progress"[12] that will go to infinity. In this regard, Kant writes that because endless progress is only possible "on the presupposition of the existence and personality of the same rational being continuing endlessly (immortality of the soul)", the highest good is practically possible only on the presupposition of the immortality of the soul (*CPrR* 5:122, 99).

Despite the explicitness of Kant's argument here that virtue or worthiness to be happy is not achievable in this world merely by finite human agency but rather requires the immortality of the soul, some commentators argue that the endless striving for virtue and worthiness to be happy is an "imperfect duty" that is still consistent with the 'ought implies can' principle. By offering a wide variety of textual evidence to support his claim, Pasternack (2017), for instance, lays out reasons to favor the idea that moral worthiness to be happy complies with the Kantian principle that 'ought implies can.' Pasternack writes that despite the need for divine assistance in the total realization of the highest good, moral worthiness is a duty that is consistent with the 'ought implies can' principle because divine aid is compatible with active moral agency. According to Pasternack, Kant is careful not to fall prey to the traditional and Augustinian view of grace, which espouses that God's help substitutes for our own work as "a foreign influence to which we must remain passive" (Rel 6:118, 124). In order to ground his argument, Pasternack refers to the "Conflict of Faculties" where Kant condemns the foreign influence over our moral conduct since "the effect of this power would not be our deed and could not be imputed to us" (Pasternack 2017, 462; *CF* 7:59). However, regarding the compatibility between divine aid and active moral agency in the realization of the highest good, Pasternack argues that the God-human relationship is akin to a family relationship: God provides some resources and guidance for the moral improvement of agents without diminishing their individual accomplishments (Pasternack 2017, 462–463).

In response to Pasternack, I would say that divine assistance is not incompatible with active human action in the realization of the highest good, i.e., that divine assistance is necessary after moral agents have deserved it, though this would not change the fact that moral worthiness still requires the ultimate supportive assistance (or grace) [13] of God. In other words, I argue that the highest good as the state of ultimate holiness is an impossible task within purely human powers because it ultimately requires us to fulfill moral law and be pleasing to God[14], which therefore requires us to assume the presence of God as the ultimate highest being. Moreover, I shall lay out other reasons to oppose the idea that earning moral perfection is compatible with the 'ought implies can' principle. In addition to his claims in the first and second *Critiques*[15], in the "Lectures on Ethics", Kant remarks that moral perfection is a duty but since human beings are imperfect by nature, divine assistance is inevitably needed for its possibility (*LE* 27:252, 47). Similarly, in *Religion*, after identifying the duty of the highest good as a "sui generis duty not of human beings toward human beings but of the human race towards itself"[16], Kant continues to argue that the promotion of the highest good as a good common to all (bringing about the ethical community) is not an ordinary moral duty that is within our power. Rather, Kant holds that the highest good is a duty that differs from all other duties in kind and in principle "for it is the idea of working toward a whole of which we cannot know whether as

[11] *CPrR* 5:122, 98.
[12] *MM* 6:447, 241.

[13] At this juncture, it is suggested to read Fremstedal (2011) who holds that Kant does not necessarily reject the possibility of the divine grace in the attainment of moral perfection but reformulates it such that the divine grace may be needed only after moral agents fulfill their moral duties. See Fremstedal 2011, 168.
[14] See *Rel* 6:47, 67-68 and *CF* 7:444, 268.
[15] In the first and second Critiques, Kant famously argues that moral worthiness can be attained with moral striving that is accompanied by the presupposition of the existence of God and a future life. See *CPR* A811/B 839 and A 829/B 857 and *CPrR* 5:122.
[16] *Rel* 6:97, 108

a whole it is also in our power" (*Rel* 6:98, 109). And finally, in the *Metaphysics of Morals,* Kant writes that moral perfection is not one of the regular duties, but it is an "imperfect and wide duty", which definitely implies that moral perfection is not achievable by human capabilities (*MM* 6:447, 241). In addition to that, Kant delineates moral worthiness as an imperfect duty in the *Metaphysics of Morals* by suggesting that two of its components, namely the promotion of "one's own perfection" and "the happiness of others" count as end-duties (duties of virtue) that are classified as wide obligations (*MM* 6:386-390).

Viewed in this way, it is apt to end this section by suggesting that moral worthiness to be happy (or virtue) is not possible merely within human abilities, but instead it is necessary to appeal to the postulates of immortality and God. One could therefore contend that without the postulations of immortality and God as a necessary ground, it would be incoherent to think of the possibility of earning moral worthiness to be happy or moral perfection since our striving for moral progress would always be incomplete toward the ultimate state of holiness.

3.2. Distribution of Happiness

In the second Critique, after defining the first component of the highest good, namely moral worthiness to be happy (or virtue), as the "supreme and unconditioned good", Kant nevertheless holds that virtue itself is not the complete good. He rather contends that it needs to be combined with happiness. As he writes, "to need happiness, to be also worthy of it, and yet not to participate in it cannot be consistent with the perfect volition of a rational being."[17] Similarly, in the *Groundwork*, Kant writes that happiness is "the natural necessity"[18] and "natural end"[19] for dependent human beings, which they all necessarily pursue. Also, in the *Groundwork*, Kant goes on to argue that the promotion of one's own happiness is an "indirect duty" (*G* 4:399, 12). And finally, in the *Lectures on Ethics* and the *Metaphysics of Morals,* Kant writes that the promotion of the happiness of others is one of the "imperfect" and "wide" duties that we have as part of moral worthiness (*LE* 29:618, 237-239 and *MM* 6:393–394, 196–197).

The fact that Kant incorporates happiness into the purview of the highest good causes certain problems, one of which arises regarding the proportionate synthesis of the two elements of the highest good. As Kant suggests, the synthesis of virtue and happiness cannot occur in an analytic way[20], but it can only be possible in a synthetic a priori manner. Defining happiness as "the state of a rational being in the world in the whole of whose existence everything goes according to his wish and will"[21], Kant claims that happiness rests upon the dynamics of nature in terms of its determining ground. So, Kant holds that happiness is not included in the purview of moral law, and therefore, moral law does not command any proportionate relation between virtue and happiness (*CPrR* 5:124, 100). The lack of any ground for the proportionate relation between virtue and happiness in rational human reason leads Kant to conclude that happiness can be distributed in proportion to morality only through the activity of a third agent, or God, who should be regarded heuristically[22] as a supreme cause that has the will and understanding for a just distribution of happiness among moral agents (*CPrR* 5:125, 101).

Given that the moral postulate of God is argued to play an integral role in the distribution of happiness, some commentators such as Villaran (2013), Denis (2005), Zeldin (1971) and Wood (1970) argue that the highest good is a special type of duty in kind that does not necessarily satisfy the 'ought implies can' principle since the need for divine assistance is inevitable in its distribution. For instance, against the argument of Beck, who holds that rewarding happiness cannot be a duty for us[23], Villaran aptly argues that rewarding our virtue is a "fallible duty", that is, we are duty-bound to synthesize our virtue with happiness, but since it is beyond our capabilities, we need to refer to God as the ultimate distributor of happiness (Villaran 2010, 33-34). Wood, on the other hand, argues that, on Kant's view, in order to hold the systematic connection between worthiness and happiness, the cooperation between "human volition

[17] *CPrR* 5: 110, 89–90.
[18] *G* 4:416, 26.
[19] *G* 4:431, 39.

[20] Here Kant criticizes the analytic relation between virtue and happiness as it is famously argued by the Epicurean and Stoic doctrines of the highest good. As Kant argues in the second Critique, the Epicurean view of the highest good consists merely of happiness, and thus Epicurus holds that virtue simply corresponds to happiness. On the other hand, the Stoic view of the highest good consists of morality, and thus the sense of morality is equal to happiness (*CPrR* 5:111, 90).
[21] *CPrR* 5:124, 100.
[22] On a basic level, Kant argues that the ideas of immortality and God are not objective truths to be known theoretically but are moral postulates that we need to assume heuristically (or for practical reasons), because this helps us to understand the order and relation of nature and morality in a systematic and unified manner.
[23] See Beck 1963, 244–245.

(which can pursue this end only imperfectly and weakly) and God's creation of a good world" is necessitated (Wood 1970, 140) According to Wood, in order to ensure that there is a proportionate relation between virtue and happiness, there must be an overlapping and compatibility between the purposiveness of human action and the purposiveness of God (Wood 1970, 140). In line with this approach in the literature, one could say that the second element of the highest good, namely distribution of happiness, is by no means an ordinary moral duty that can be conceived within the purview of the 'ought implies can' principle. However, this does not necessarily suggest that it is completely beyond the scope of our moral duty as it is ultimately grounded on the aid of God. Instead, I would suggest that the distribution of happiness is a wide obligation that we have to pursue by relying on our moral worthiness and hoping for the aid of God. This can be grounded on the fact that Kant conceives of the moral postulates of immortality and God as part of the distribution of happiness not as an objective duty but a morally commanded (conducive to moral purposes) duty to which we are obligated.[24] In this regard, as far as I see, it is apt to define the distribution of happiness as a special type of duty, namely a wide obligation, since it requires us to strive morally on our own but also demands us to presuppose the existence of God as the ultimate synthesizer of virtue and happiness. In sum, one could suggest that the distribution of happiness is a wide obligation (an imperfect duty) that we cannot realize on our own but should pursue to infinity by appealing to the assistance of God.

Thus far in our discussion, I have argued that neither moral worthiness to be happy (or virtue) nor distribution of happiness can be viewed as duties that simply satisfy the principle of 'ought implies can' since they both appeal to the superhuman elements in their possibility of realization. In this regard, I have categorized the highest good among wide obligations (imperfect duties) that we have to pursue to infinity and with the aid of God since it goes beyond our powers. In the remainder of the paper, I will attempt to give a more nuanced reading regarding the novelty of the Kantian highest good as a duty. My goal will be to show the normative ground of the highest good and the legitimacy of its necessary connection with the postulates of immortality and God.

[24] See *CPrR* 5:146.

4. The Highest Good as a Duty, Its Normative Ground and Extensions

An attempt to show that the highest good goes beyond human capabilities and it requires the postulates of immortality and God as a necessary ground is a thorny issue since it leads us into a vulnerable position against the problem of heteronomy, which Kant strictly argues against in the *Groundwork* and the Analytic of the second *Critique*. In this section, aiming to avoid the problem of heteronomy, I will propose that Kant's highest good is fully grounded on moral law in a normative sense despite the fact that it goes beyond the limits of human capabilities and ultimately appeals to the postulates of immortality and God as the moral ground of assistance and complementation. In this regard, I will attempt to show that the moral law provides a normative ground for the highest good because it ultimately embodies the ground of absolute and unconditioned goodness (or holiness) which can produce unattainable duties. Therefore, I will suggest that the highest good normatively stems from the moral law as an imperfect duty which cannot be fully performed with human capabilities but requires us to be assisted and complemented by the postulates of immortality and God.

In order to elucidate the legitimacy of the excessive scope of the highest good as a duty that goes beyond human capabilities and requires the postulates of immortality and God as a necessary ground, I should take a closer look at the unconditioned character of the moral law that commands categorically and absolutely, which can thus generate duties that are beyond the scope of the 'ought implies can' principle. In fact, in his Critical corpus, Kant states that the moral law embodies the ground of unconditioned and absolute goodness or the state of holiness[25] that suggests that it is impossible to be attained in full by

[25] Here I should note that the question of holiness is highly complicated with respect to the Kantian conception of the moral law. As Kant says in the Groundwork, the demand for holiness of the moral law can only apply to imperfect and finite human agents, since they lack perfection in their nature and seek the perfection and holiness of the moral law as a way to complete themselves (*G* 4: 414, 25). However, I would like to avoid any confusion by pointing out that for Kant, although the moral law gives rise to the requirement of holiness, this does not mean that the moral law is holy in itself, which would completely prevent moral agents from deviating from duty. Therefore, it is necessary to point out that the moral law is a duty from which arises the requirement that rational agents reach the state of holiness, which can be seen as a path to moral perfection that corresponds to an unending process for finite agents.

human powers. For example, in the *Groundwork* and the Analytic of the second Critique, Kant holds that the complete harmony between one's maxims and the laws of autonomy corresponds to the state of holiness or absolute goodness (*G* 4:439, 6 and *CPrR* 5: 32, 29). However, as Kant writes, moral agents are not capable of attaining holiness or absolute goodness that is dictated by the moral law because they are dependent by their nature (*G* 4:439, 46 and *CPrR* 5: 32, 29). Therefore, Kant holds that moral agents can never attain holiness but are merely obligated to pursue it. In the second Critique, Kant further writes that holiness as the complete conformity to law is a practical idea that serves as a model for finite moral agents (*CPrR* 5:33, 29–30). Kant, thus, defines moral virtue as an endless striving or approximation towards holiness since, as he writes, the holiness demanded by the moral law is never fully attainable but is always an ongoing process of developing one's maxims toward it. Moreover, in *Religion*, where Kant insists that there is a wide gap between the holiness of moral principles and the sensible human nature, he states that the purity of the moral law is unattainable by moral agents, but their moral striving to incorporate the purity into their maxims shows that they are on "the road of endless progress toward holiness" (*Rel* 6:47, 67–68).

Along similar lines, in the *Metaphysics of Morals*, Kant continues to write that the true value of morality as a state of holiness exceeds the value of any usefulness or empirical ends, and he thus proposes that it is the highest duty for rational agents to pursue this ideal state (*MM* 6:397, 199-200). This claim basically suggests that rational agents ought to overcome their limitations and strive to reach the state of holiness. And finally, in *Lectures on the Philosophical Doctrine of Religion*, defining holiness as "the absolute or unlimited moral perfection of the will", Kant argues that nobody can ever achieve the state of holiness since it is the absolute perfection that can only be achieved by God (*LPDR* 28:1076). But, as Kant writes, although moral agents who are fundamentally dependent and imperfect can never achieve the state of holiness, they can be virtuous. So, Kant suggests that although the state of holiness is not achievable by human capacities, it is still a duty for rational agents to pursue it by attempting to surpass their limitations.

Based on this textual evidence where Kant conceives of the moral law as an embodiment of the unconditioned holiness which can never be fully achieved by finite moral agents, it is possible to deduce a clear implication for the fact that the moral law is the sole determining ground for the derivation of the highest good despite the fact that the real possibility of the highest good oversteps human capabilities. In this regard, one could argue that since the moral law produces unconditional laws that target the state of holiness, the highest good as the synthesis of moral perfection and the distribution of happiness should derive from the moral law as an impossible goal but still constitute a legitimate duty to be pursued.

Having explicated that the moral law should serve as a normative ground for the highest good, it is now necessary to ground the legitimacy of the link that the highest good establishes between the moral law and the postulates of immortality and God. In what follows, I will try to indicate that the extension of the highest good over the ground of immortality and God is not heteronomous in character, because the ground of immortality and God is not a normative ground that the highest good stems from but is an assisting and complementing element that grounds its ultimate possibility.

In his three critiques, Kant openly lays out the necessary role of the postulates of immortality and God in the establishment and advancement of the highest good under the conception of "moral theology." In regard to the idea of moral theology, in the "Canon of the Pure Reason" of the first Critique, Kant writes that the moral laws would prove to be empty fabrications of the brain rather than being real moral commands unless they are related to the postulates of immortality and God because it is only the presupposition of the highest being, or God, that gives a purposive unity to the morality (*CPR* A812/B840, 680). Moreover, Kant argues that without the introduction of the postulates of immortality and God, which provide a systematic unity to morality, the moral concepts and precepts would lack a greater uniformity (*CPR* A 817/B845, 683).[26] On the basis of this idea, Kant goes on to define the moral laws as "divine commands" because, on Kant's view, we are internally obligated to fulfill them, and there is an inescapable link between the moral laws and the divine will since we can be in conformity with the divine will only if we accept the moral law as holy, and we can serve the

[26] Regarding the indispensable role of moral belief in God on our moral motivation, see Denis 2005, 39-42. Denis basically argues that although our moral commitment mostly aims at morality, we sometimes deviate from the demands of morality if we do not have the idea of happiness that will be granted by God. But she also notes that Kant never regards the idea of happiness as always sufficient for right moral action.

divine will only if we do our best for the world and ourselves and others (*CPR* A 819/B 847, 84; *CPrR* 5:130, 104; *Rel* 6:154, 153-154; *CPJ* 5:444, 310). In this regard, Palmquist argues that, on Kant's view, we must view our moral duties as if they were divine commands, that is, we must connect them through the moral law to the supreme being because otherwise we will never be able to fulfill their requirements (Palmquist 2000, 128).

Further, in "On the Common Saying", *Metaphysics of Morals* and his "Lectures on Philosophical Doctrine of Theology" Kant continues to argue in a similar fashion, by suggesting that the moral arguments are indispensably linked to morality. In "On the Common Saying", Kant holds that the attainment of the highest good in full demands that we appeal to the existence of God and a future life for practical purposes (*TP* 65). Similarly, in the *Metaphysics of Morals*, where Kant attempts to exclude religion from the purview of the pure philosophical morality[27], he writes that although the postulation of immortality and God is not an objective duty, it is a subjective need that strengthens our moral incentive (*MM* 6:487, 276). And finally, in the "Lectures on the Philosophical Doctrine of Theology", Kant explicitly writes that God must be the ultimate ruler of reason and moral laws, and thus, one must be a scoundrel or visionary if he denies his existence (*LPDR* 28:1072, 406-407). In this regard, Kant suggests that the postulates of immortality and God refer to the subjective moral beliefs which are *conditio sine qua non* for morality since, as he writes, if we deny the existence of immortality and God, we could lose our moral worth altogether. With this regard, Wood and Palmquist argue that Kant's moral arguments for immortality and God save morality from "absurdum practicum"[28] or the "abyss of meaningless."[29]

From all that I have unveiled about Kant's view of moral theology, it becomes pertinent to say that Kant directly and bluntly connects the moral laws with divine will through the intermediary of the conception of the highest good. It is profoundly clear that Kant conceives morality without moral theology as a meaningless and absurd enterprise. Therefore, it becomes possible to say that although Kant is highly critical and agnostic about the grace of God in a traditional sense[30], he makes room for divine assistance in the realization of the highest good in a legitimate manner since, on his account, moral agents are incapable of the full realization of the highest good and they need an assisting element to complement their imperfections through the presupposition of immortality and the existence of a God. And in parallel, our claim that the moral law is not limited by the principle of the 'ought implies can' and that it can generate duties that are not attainable in scope allow us to argue that the highest good is a wide obligation (imperfect duty) that we cannot attain at any point in this life merely by our own powers, but rather, we have to pursue it endlessly as a moral necessity with the accompaniment of the moral postulates of immortality and God as the ultimate ground of assistance and complementation.

5. Conclusion

One can conclude this paper by shedding some light on how the suggested argument is attempting to resolve the conflict caused by Kant's two contradictory statements about the relationship between the highest good and the postulates of immortality and God.

In the face of Kant's two controversial claims that suggested that the highest good is a possible duty in the world although it requires the postulates of eternity and the existence of a God for its real possibility, I have argued that the highest good is not an ordinary moral duty that is consistent with the principle of the 'ought implies can' but is a wide obligation (imperfect duty) that we have to pursue by overcoming our limits and by appealing to immortality and the aid of God. In this regard, I have argued that the highest good should derive from the moral law since the moral law embodies the ground of holiness or absolute goodness, which can produce duties that are impossible in scope. On this basis, I have argued that the highest good which normatively derives from the moral law can extend beyond the 'ought implies can' principle and appeal to the postulates of immortality and God in a legitimate manner since the ground of immortality and God does not constitute a normative

[27] In the concluding remark of the *Metaphysics of Morals* Kant argues that ethics has nothing to do with man's relations with God but only to do with the moral relations between men and men (MM 6:488-491: 276-279).
[28] Wood 1970, 25-34.
[29] Palmquist 2000, 75.

[30] In the *Conflict of Faculties* and *Religion* Kant argues that grace is incomprehensible and unknowable to us and it does not have an objective ground in morality (*Rel* 6:191, 182-183 and *CF* 7:43, 268). At this point, it is also suggested to read Pasternack (2017) where he argues that Kant opposes the Augustinian idea of grace and proposes a reformulation of divine aid that is compatible with human agency (See Pasternack 2017, 461-464).

basis for the highest good but an assisting and complementing ground that helps to advance our moral perfection by aiding us to reach the state of holiness. In a nutshell, one can summarize the claim of this paper by saying that in Kant's Critical corpus, the highest good is not presented as an ordinary moral duty to be strictly performed by rational agents, but rather presented as a wide obligation that normatively derives from the moral law but goes beyond the 'ought implies can' principle, appealing to the postulates of immortality and God as a ground of assistance and complementation.

References

Beck, L. W. (1963). *A commentary on Kant's critique of practical reason*. University of Chicago press.

Caswell, M. (2006). Kant's Conception of the Highest Good, the Gesinnung, and the Theory of Radical Evil. *Kant-Studien, 97*, 184–209.

Denis, L. (2011). Autonomy and the Highest Good. Kantian Review, *10* (2006), 33–59.

Fremstedal, R. (2011). The Concept of the Highest Good in Kierkegaard and Kant. *International Journal for Philosophy of Religion*, *69*(3), 155–171.

Gregor, M., Anchor, R., & Kant, I. (1996). The Conflict of the Faculties (1798) In: Wood A. W., di Giovanni G. (eds.) *Religion and Rational Theology* (pp. 233–328) The Cambridge Edition of the Works of Immanuel Kant. Cambridge University Press.

Kant, I. & Korsgaard, C. (1998). *Groundwork of the Metaphysics of Morals* (Cambridge Texts in the History of Philosophy) (M. Gregor, Ed.). Cambridge: Cambridge University Press.

Kant, I. (1998). *Critique of pure reason*. (Eds.) Guyer, P. et al. Cambridge, New York, Cambridge University Press.

Kant, I. (1997). *Critique of practical reason*. (Eds.) Gregor, M. J, & Reath, A. Cambridge: Cambridge University Press.

Kant, I. (2000). *Critique of the Power of Judgment*. Translated by Paul Guyer and Eric Matthews. Cambridge: Cambridge University Press.

Kant, I., & Adams, R. (1998). *Religion within the Boundaries of Mere Reason: And Other Writings* (Cambridge Texts in the History of Philosophy) (A. Wood & G. Di Giovanni, eds.). Cambridge: Cambridge University Press.

Kant I. & Gregor M. J. (1991). *The metaphysics of morals*. Cambridge University Press.

Kant, I. (1997). "Lectures on Ethics" (Heath & J. Schneewind, Eds.). Cambridge: Cambridge University Press.

Kant, I. (1991). "On the Common Saying: "This may be true in history but does not apply in practice". In *Political Writings*, (Ed.) H. Reiss and (trans.) H. Nisbet. Cambridge: Cambridge University Press.

Kant, I. (1996) "Lectures on the philosophical doctrine of religion" (1817). In *Religion and Rational Theology*, (Eds.) Allen Wood and G. Die Giovanni. Cambridge: Cambridge University Press.

Kleingeld, (2016). Kant on 'Good', the good, and the duty to promote the highest good. In T. Höwing (Ed.) *The highest good in Kant's philosophy* (p. 33–50). Berlin: De Gruyter.

Kohl, M. (2015). Kant and "Ought Implies Can." *The Philosophical Quarterly (1950)*, *65*(261), 690–710.

Palmquist, S. R. (2000). Kant's Critical Religion: Volume Two of Kant's "System of Perspectives" (1st ed.), Ashgate.

Pasternack, L. (2017). Restoring Kant's Conception of the Highest Good, *Journal of the History of Philosophy*, *55*(3), 435–468.

Reath, A. (1988). Two Conceptions of the Highest Good in Kant. *Journal of the History of Philosophy, 26*, 593–619.

Showler, A. & Wike, V. (2010). Kant's Concept of the Highest Good and the Archetype-Ectype Distinction. *Journal of Value Inquiry, 44*, 521–533.

Silber, J. R. (1959) Kant's Conception of the Highest Good as Immanent and Transcendent. *Philosophical Review*, *68*(4), 469–492.

Stern, R. (2004) Does 'Ought' Imply 'Can'? And Did Kant Think It Does? *Utilitas, 16*, 42–61.

Villarán, A. (2013) Overcoming the Problem of Impossibility in Kant's Idea of the Highest Good. *Journal of Philosophical Research, 38*, 27–41.

Wood, A. (1970). *Kant's Moral Religion*. Ithaca: Cornell University Press.

Yovel, Y. (1980). *Kant and the Philosophy of History*. Princeton: Princeton University Press.

Zeldin, M. (1971) The Summun bonum the moral law and the existence of god. *Kant Studien: Philosophische Zeitschrift Der Kant Gesellschaft, 62*, 43–54.

SEARCHING FOR IDENTITY: IS J.S. MILL ALSO A NORMATIVE CULTURAL RELATIVIST?

Olanipekun Olusola Victor

Department of Philosophy
Obafemi Awolowo University, Ile-Ife, Nigeria
email: Olanipekuno@oauife.edu.ng
ORCID ID: https://orcid.org/ 0000-0003-4324-9503

Abstract:
The article explores and attempts the possibility of reading normative cultural relativism into J. S. Mill's *On Liberty*. This article examines J. S. Mill's *On Liberty*. It argues that even though Mill is often described as a moral universalist as evident in the methodology employed in the postulation of his ethical utilitarianism, one could also argue that his position maintained in *On Liberty* qualifies him to be a normative cultural relativist. Basically, the article contends that beyond the universalist reading of Mill's works, he is qualified to be called a normative cultural relativist. The article interrogates and interprets the position maintained by J. S. Mill's in the *On Liberty* that commits him to normative cultural relativism. The paper concludes that anyone that maintains such a position as Mill need not necessarily reject normative cultural relativism.

Key words: the highest good, ought implies can, immortality, holiness, the aid of God, the moral law.

Introduction

John Stuart Mill is one of the major influential 19th century British philosophers. Some of his intellectual contributions to moral philosophy as well as political philosophy are found in his works, such as *On Liberty* (1859) and *Utilitarianism (*1861). Like Jeremy Bentham's version, Mill's *Utilitarianism* is also regarded as a classic. Similarly, *On Liberty*, by J. S. Mill is also one of the classics available in ethics as well as social and political philosophy. The focus of this paper is Mill's *On Liberty*. A careful examination of *On Liberty* reveals that the major part of the book is: (i) a reaction to the existing prevailing restriction of liberty and constraints of the time, (ii) a response to the moral limits of the state, and (iii) concentrated on the liberty and freedom that people in the society should enjoy as against what the customs and the political setting of the time offered.

The paper is divided into four main parts. The first part examines universalism in Mill. The second part focuses on the normative cultural relativist main claims. The third part discusses Mill's view in the *On Liberty*. Meanwhile, the fourth section examines how J. S. Mill's view could be interpreted in order to reveal the elements of normative cultural relativism. Arising from the view that Mill defended in the *On Liberty,* the paper concludes by showing how and why I think he could be a normative cultural relativist of a certain sort. Building on Jeremy Waldron's ideas, I examine how Mill's work celebrates diversity and why we are therefore entitled to think that the spirit of *On Liberty* is also hospitable to cultural diversity and by extension, normative cultural relativism.[1]

Now, it is essential to begin a paper of this nature by considering universalism in Mill's works in

[1] Jeremy Waldron, "Mill and Multiculturalism", *Mill's On Liberty: A Critical Guide*, ed. C. L. Ten, New York: 2008, p.165

order to provide a background for his relativist views.

Universalism in J.S. Mill

As a moral theorist, Mill rests his account of utilitarianism on the principle of utility, which he considered to be the highest moral principle. Mill is regarded as a universalist because he employed a universalist methodology in establishing his principle of utility, which is grounded on an impersonal or impartial point of view that can be applied to anyone. The fact that such a principle is open to everyone makes it universal. In Mill's view, the greatest happiness principle is the foundation upon which this principle of utility, as well as the notion of moral rightness, rests. Mill's formulation of the greatest happiness principle goes thus:

The utilitarian creed which accepts, as the foundation of morals, utility or Greatest Happiness Principle, holds that actions are right in proportion as they tend to promote happiness, wrong as they tend to produce the reverse of happiness. By happiness is the intended pleasure, and the absence of pain; By unhappiness, pain, and the privation of pleasure.[2]

The above view is often interpreted to be universal because the happiness in question is the happiness of all moral agents that are involved in a particular moral universe. That is, it is impartial and impersonal because it is about the promotion of everyone's well-being. Following Mill's view, we must be as strictly impartial as a disinterested and benevolent spectator.[3] James Rachels' view buttresses Mill's argument because for Rachels, every important moral theory must include the idea of impartiality.[4] The above view is also supported by Paul Gomberg who maintains that impartiality is regarded as one of the essential requirements of universalism in ethics.[5]

To further demonstrate the inherent universalism of his moral philosophy, Mill argues that the equal claim of everybody to happiness involves an equal claim to all the means of happiness.[6] The above view suggests universalism because the emphasis here is on the equal claim of all to happiness and not the claim of some people to happiness. Thus, universalism could be demonstrated at the level of the equal claim of every moral agent in a particular moral universe to happiness.

Another universalist foundation of moral rightness or wrongness is found in Mill's *Utilitarianism*. According to Mill, this firm foundation is that of the social feelings of mankind, the desire to be in unity with our fellow creatures, which is already a powerful principle in human nature.[7] The above view arises from the unique understanding of our human nature with the universal character of how human beings cannot fully realize their potential outside the community of others. Thus, given the experience of humanity as a species, Platonov noted how Mill reveals moral values and rules not only in the aspect of locality and variability, but also in the aspect of fundamental skills of interaction of people with each other.[8]

Also, for Mill, "utility would enjoin that laws and social arrangement should place the happiness… or the interest of every individual, as nearly as possible in harmony with the interest of the whole… so that a direct impulse to promote the general good may be in every individual."[9] It was the above view that influenced Richard Evanoff who contends that Mill's utilitarian creed rests on a universal conception of human happiness.'[10] However, this view may be contested in the sense that it is doubtful whether the universal conception of human happiness is realizable in practice. Contesting the above view, though, is not my focus in this section; instead, it is the demonstration of universalism in Mill.

A similar view was extensively discussed and expressed by R.S. Platonov in his "Moral Universality in J.S. Mill's Utilitarianism." According to Platonov, "moral universalism in J.S. Mill's philosophy is an important element of his methodology in

[2] J.S. Mill, *Utilitarianism*, Chicago: The University of Chicago Press, 1906, p.9
[3] Rachels, J., *The Elements of Moral Philosophy*, p.159. Also, J.S. Mill, *Utilitarianism*, p.68
[4] Rachels, J., *The Element of Moral Philosophy*, p.12
[5] Paul Gomberg, "Universalism and Optimism" in *Ethics*, vol.104, No.3(1994), 536–557
[6] Mill, J.S., *Utilitarianism*, p.93
[7] Mill, J.S., *Utilitarianism*, p.46
[8] *Roman S. Platonov*, "*Moral Universality in the Ethics of Classical Utilitarianism*" in *Антиномии*, vol. 20, No.4,(2020), 45–64
[9] J.S. Mill, *Collected Works of J.S. Mill*, ed. J.M. Robson, London: Routledge, 1991.
[10] Richard J. Evanoff, "Universalist, Relativist, And Constructivist Approaches to Intercultural Ethics" in *International Journal of Intercultural Relations*, vol. 28 (2004), 442

ethics...and the basis of maximizing utility."[11] The essence of the above quotation is to present Mill as a universalist due to the methodology he adopted in the conception of his version of utilitarianism.

Now, it is important to announce at this point that my main concern in this paper is neither to defend Mill's universalism nor his methodology, but the bulk of the work is to consider the element of normative cultural relativism in his work. As a matter of necessity, it is imperative to first of all clarify the common core of normative cultural relativism.

Normative Cultural Relativists' Main Claims

The following represent the claims defended by the normative cultural relativists:

I. There is no single morality. Different societies have different moral codes.[12]
II. The moral code of a society determines what is right within that society; that is, if the moral code of a society says that a certain action is right, then that action *is* right, at least within that society.[13]
III. There is no objective standard that can be used to judge one society's code as better than another's. There are no moral truths that hold for all people at all times.[14]
IV. There is no supreme moral principle. The moral code of our own society has no special status.[15]
V. It is arrogant for us to judge other cultures. We should always be tolerant of them.[16]
VI. The above views formed the main claims of the normative cultural relativists.

Basically, to be a normative cultural relativist, one needs to accept the above claims. Having stated the main claims of the normative cultural relativists, it is essential at this point to consider Mill's arguments in the *On Liberty* to compare and contrast, and to also make some deductions.

Mill's Arguments and Position in *On Liberty* Considered

In the chapter one of *On Liberty*, Mill introduced the readers to the book. Mill devoted chapter two to the liberty of thought and discussion. Chapter three was about the liberty of individuality as one of the elements of well-being. Mill devoted chapter four to the limits to the authority of society over the individual. Meanwhile, chapter five of the book was basically about applications. The article examines certain aspects of this particular work of Mill.

The article gives special attention to chapter three of *On Liberty* where Mill made the following claims. According to Mill, "to give any fair play to the nature of each, it is essential that different persons should be allowed to lead different lives. In proportion as this latitude has been exercised in any age, has that age been noteworthy to posterity."[17]

Now, below are some possible deductions as well as claims that are derivatives or identified in chapter three of the *On Liberty* by J.S. Mill:

a) Different persons should be allowed to lead different lives.[18]
b) People have diversities of taste, that is reason enough for not attempting to shape them all after one model.[19]
c) Different persons require different conditions for their spiritual and moral development.[20]
d) The same things which are helps to one person towards the cultivation of his higher nature are hindrances to another.[21]
e) There are differences among human beings in their sources of pleasure, their susceptibilities of

[11] For details, see Platonov R.S., "Moral Universality in J.S. Mill's Utilitarianism" in *Russian Journal of Philosophical Sciences.* Vol. 62, no. 11, (2019), 84–95.
[12] Gilbert Harman and J.J. Thomson, *Moral Relativism and Moral Objectivism*, Cambridge: Blackwell Publishers, 1996, p.5
[13] J. L. Mackie, *Ethics: Inventing Right and Wrong*, London: Penguin Book, 1977, p.36
[14] Gilbert Harman, "Precis of Moral Relativism and Moral Objectivism", *Philosophy and Phenomenological Research*, vol. LVIII, No.1:(1998):161
[15] Alasdair MacIntyre, *After Virtue: A Study in Moral Theory*, Notre Dame: University of Notre Dame Press, 2007, p. xviii
[16] James Rachels, *The Elements of Moral Philosophy*, New York: McGraw-Hill, 2012, p.16
[17] J.S. Mill, *On Liberty*, Ontario: Batoche Book Limited, 1859, p.59
[18] J.S. Mill, *On Liberty*, Ontario: Batoche Book Limited, 1859, p.59
[19] J.S. Mill, *On Liberty*, Ontario: Batoche Book Limited, 1859, p.63
[20] J.S. Mill, *On Liberty*, Ontario: Batoche Book Limited, 1859, p.63
[21] J.S. Mill, *On Liberty*, Ontario: Batoche Book Limited, 1859, p.63

pain, and the operation on them of different physical and moral agencies, that unless there is a corresponding diversity in their modes of life, they neither obtain their fair share of happiness, nor grow up to the mental, moral, and aesthetic stature of which their nature is capable.[22]

f) The principle of tolerance ought to be encouraged due to diversity of taste.[23]
g) People should be allowed to be different without depreciatory remark as if they have committed some grave moral delinquency.[24] That is, people should be permitted to make different moral choices without being accused of doing what is morally bad.

Arising from the above, the next equally important question is, how do we read normative cultural relativism into this view? The answer to this question shall be the focus of the next section.

Reading Normative Cultural Relativism into J.S. Mill's *On Liberty*

Having considered Mill's view above, here is the analysis as well as possible interpretations. The first question that arises is, could anyone hold the view that Mill advocates above without being sympathetic to normative cultural relativism? I supply a negative answer. I argue that J.S. Mill's views above could serve as arguments for normative cultural relativism. The reason is simply because one way to summarize the above views by Mill is to maintain that 'one should not pass judgment on others with different values.' That is, one need not hold the same values or moral norms as others. Incidentally, the same view is one of the core views defended by the normative cultural relativists.

Essentially, it has been argued by scholars such as Wilson Fred that Mill's arguments in chapter three of the *On Liberty* are basically about liberalism with the focus on liberty, freedom, and the rights of individual persons against social and cultural restrictions.[25] But when carefully examined, one will discover that the same argument could be marshaled and interpreted to suit the view defended by the normative cultural relativists at a certain level of discussion. Let us now attempt to interpret and juxtapose the above views by moving from being general to being more specific.

From the claim ***a*** above by Mill, the view that different persons should be allowed to lead different lives arises out of the fact that different societies have different values that ought to be respected. Interestingly, this is a restatement of the normative cultural relativist claim that the standard of moral evaluation in each society ought to be respected. This arises from the claim **I** of the normative cultural relativists above that different societies have different moral codes. It may be argued by the critics that Mill's claim is directed toward an individual person's liberty, and not to a society. But a careful examination reveals that society is an aggregation of the individuals living together in a particular community. Thus, this paper contends that the same view may be applied. In fact, in Tyndal's view, an appreciation for cultural diversity can be built into an appreciation for individual diversity.[26] Thus, when individual persons are free to pursue or lead different lives because of differences in values, it means that their societies or cultures permit such.

The fact that different persons ought to be allowed to lead different lives as argued by Mill, and that different societies have different moral codes as defended by normative cultural relativism means that issues such as homosexuality, child-marriage and polygamy will be permitted by some societies despite the fact that other societies frown on these practices. This notion of differences is responsible for the reason why there is no consensus on certain moral matters such as the examples given above.

Also, from the claim ***b*** above by Mill, the view that people have diversities of taste, and that this is reason enough for not attempting to shape them all after one model, could be interpreted as an anti-universalist statement as well as a denial of objectivity. The above view appears to be an anti-paternalistic claim. Of course, anti-paternalism is not

[22] J.S. Mill, *On Liberty*, Ontario: Batoche Book Limited, 1859, p.63
[23] J.S. Mill, *On Liberty*, Ontario: Batoche Book Limited, 1859, p.64
[24] J.S. Mill, *On Liberty*, Ontario: Batoche Book Limited, 1859, p.64
[25] Wilson Fred, "John Stuart Mill", *The Stanford Encyclopedia of Philosophy*, (Spring 2016 Edition) Edward N. Zalta.
[26] Tyndal Jason, "Culture and Diversity in John Stuart Mill's Civic Nation" in *Utilitas* Vol. 25, No. 1, (2013), 106

the same as cultural relativism. But one can be an anti-paternalist and still embrace cultural relativism. By implication, emphasizing diversity of taste and value is an allusion to normative cultural relativism. As a matter of fact, Jeremy Waldron and Tyndal Jason also attest to the fact that "the justification for a Millian multiculturalism is traced to the value placed on cultural diversity"[27] and this in turn paves way for diversity of cultural and moral values. Just like Mill, the normative cultural relativists also acknowledge and recognize differences.

Interestingly, the above claim by Mill is a restatement of the normative cultural relativist claim **III** that there is no objective standard that can be used to judge one society's code to be better than another's. For normative cultural relativists, there are no moral truths or values that hold for all people at all times due to the diversities in values. Thus, the fact that monogamy, polygamy and polyandry exist as equally legitimate marital norms corroborates the view that people have diversities of taste and need not be forced to embrace any universal standard of marital practices so long as there is no objective standard that can be used to judge the moral code of one society to be better than that of another. For instance, what argument would the societies that practice monogamy give to support the claim that monogamy should be the universal standard of marital practice? In the same spirit, what argument would societies that allow polygamy give to support the claim that polygamy should be the universal standard of marital practice? In an attempt to respond to these questions, some possible arguments to justify monogamy could be; one, 'the quality of life argument'[28], and two, 'the religious argument'[29]. However, these arguments are not sufficient and could easily be countered by the example of polygamous families enjoying quality of life and also by religions that support or justify polygamy.

Similarly, from ***c*** above, the claim that different people require different conditions for their spiritual and moral development according to Mill is, incidentally, a restatement of the normative cultural relativist's claim that each culture occupies an equally legitimate moral space because of their different social conditions. This view may be given the same interpretation as the normative cultural relativists' claim **IV** that the moral code of our own society has no special status; it is but one among many.

The implication of the claim that different people require different conditions for their spiritual and moral development is that individual people/cultures ought to have different standards of moral and spiritual development, and that those standards ought to be respected. For instance, X's situation may require the condition that will make him/her adopt the tenets of Christianity for moral and spiritual development. Meanwhile, Y's situation may require the condition that will make him/her adopt the tenets of Islam for moral and spiritual development. Incidentally, the normative cultural relativists need not reject the above view so long as that is the practice approved by the society under consideration.

Furthermore, from claim ***d*** above, the view that the same things which are helps to one person towards the cultivation of his higher nature are hindrances to another by Mill is an attempt to echo his view that each person and society should pursue whatever means of cultivation of higher nature are most suitable.[30] Interestingly, the above view by Mill is a restatement of the normative cultural relativist claim **II** above that the moral code of a society determines what is right within that society. That is, if the moral code of a society says that a certain action is right, then that action *is* right, at least within that society. By implication, what is right for one person or society may be wrong for another. This explains why individual persons or societies are expected to pursue a kind of value that is suitable for each of them so that the view that the same things which are helps to one person towards the cultivation of his higher nature will not become hindrances to another. For instance, the picture below painted by James Rachels captures the above view. According to Rachels,

[27] Jeremy Waldron, "Mill and Multiculturalism", *Mill's On Liberty: A Critical Guide*, ed. C. L. Ten, New York: 2008, p.183. Also, see Tyndal Jason, "Culture and Diversity in John Stuart Mill's Civic Nation" in *Utilitas* Vol. 25, No. 1, (2013), 96

[28] The quality of life argument defended by the supporters of monogamy states that people with one wife usually have a smaller family size, and often have a limited number of people to care for as compared to people with many wives. Thus, their family tends to enjoy a better quality of life. However, when carefully examined, this does not follow necessarily. There are situations where the reverse is the case.

[29] The religious argument in defense of monogamy states that some religions frown on the practice of polygamy.

[30] J.S. Mill, *On Liberty*, Ontario: Batoche Book Limited, 1859, p.63

Darius, a king of ancient Persia, was intrigued by the variety of cultures he met in his travels. He had found, for example, that the Callatians, who lived in India, ate the bodies of their dead fathers. The Greeks, of course, did not do that — the Greeks practiced cremation and regarded the funeral pyre as the natural and fitting way to dispose of the dead. Darius thought that a sophisticated outlook should appreciate the differences between cultures. One day, to teach this lesson, he summoned some Greeks who happened to be at his court and asked what it would take for them to eat the bodies of their dead fathers. They were shocked, as Darius knew they would be, and replied that no amount of money could persuade them to do such a thing. Then Darius called in some Callatians and, while the Greeks listened, asked them what it would take for them to burn their dead fathers' bodies. The Callatians were horrified and told Darius not to speak of such things.[31]

The above example expresses the claim that the same things which are helps to one person towards the cultivation of his higher nature are hindrances to another. This further expresses the view that what is right for one may be wrong for others and vice versa. This is a scenario where we have the same action but different reactions because different people have different cultural/moral values.

Also, from claim ***e*** above, Mill's view that there are differences among human beings is also a way of justifying the diversity thesis defended by the normative cultural relativists. As a matter of fact, interestingly, the above view is a restatement of the claim **I** by the normative cultural relativists that different societies have different moral codes. This is so because there are differences among people of different cultures and societies. This could be applicable to either individual persons or individual cultures. In fact, for Wendy Donner, the societal culture thus provides the freedom and the resources for this reflection and questioning. It gives us the cultural materials needed to reach an awareness of different views of the good life, as well as the capacities required to reflect critically upon the presented options.[32]

Similarly, from claim ***f*** above, the view that the principle of tolerance ought to be encouraged due to diversity of taste by Mill is a restatement of the tolerance argument which is the claim **V** by normative cultural relativists. The claim is that it is arrogant for us to judge other cultures. We should always be tolerant of them. Just as the principle of tolerance is one of the key arguments in defense of normative cultural relativism, it is interesting to know that political liberalism also accommodates and supports tolerance as well. According to Daniel Augenstein, tolerance is one of the truest expressions of the liberal sentiment that resides in the awareness of the continuous tension between unrestrained liberty and unopposed oppression within liberalism itself.[33] The implication of the above view is that, both liberals such as Mill, and normative cultural relativists accept and preach the gospel of tolerance due to diversity of taste and values. If this is accepted, it will not be unreasonable to argue that Mill is sympathetic towards normative cultural relativism. In fact, just as Mill is interested in protecting people's rights and choices, normative cultural relativists are also interested in protecting and respecting people's rights and moral choices as stipulated by their society. Just as liberals support the protection of people with different moral codes, normative cultural relativists also support this view. As a matter of fact, the normative cultural relativists defend the view that the moral codes of different societies ought to be equally respected.

Furthermore, from ***g*** above, the view that people should be allowed to be different without depreciatory remark as if they have committed some grave moral delinquency by Mill points to the view that people's choice ought to be respected. In other words, people's moral space should be regarded. This implies that people should be permitted to make legitimate different moral choices without being accused of doing what is morally wrong. Incidentally, this is a restatement of the tolerance argument by the normative cultural relativists that was considered earlier. For example, if the people of the United States of America approve of monogamy, they should be allowed to practice such without being castigated. Also, if African people approve polygamous family life, they should be allowed to be different without derogatory remark as if they

[31] James Rachels, *The Elements of Moral Philosophy*, New York: McGraw-Hill, 2012, p.14

[32] Wendy Donner, "Autonomy, tradition, and the enforcement of morality" in ed. Ten, *Mill On Liberty: A Critical Guide*, Cambridge: Cambridge University Press,2008, p.142

[33] Daniel Augenstein, "Tolerance and Liberal Justice", in *Ratio Juris*. Vol. 23 No. 4, (2010):437–459

have committed some grave moral delinquency. This, by extension, is a way of expressing the view advanced by the normative cultural relativists that the standards of moral evaluation in each culture or society ought to be respected.

Conclusion and Further Critical Reflection

Now, in the final analysis, do the above claims identified by Mill qualify him to be a normative cultural relativist? I supply a positive response to this. The reason is simply because the normative cultural relativists can accept and need not reject the claims ***a*** to ***g*** that we derived from the chapter three of Mill's *On Liberty*. In essence, if we could infer that normative cultural relativists could be Millians, it could also be argued that Mill needs not reject normative cultural relativism as well. In fact, Wendy Donner interpreted Mill to maintain a position that Cultural practices and traditions that encourage respect for freedom and dignity, that propel human well-being and progress, and that are freely accepted by members of the society are worthy of protection.[34]

According to Mill, tolerance, cultural diversity, moral diversity, political pluralism and diversity of character are essential for any reasonable society.[35] The above view expresses the fact that Mill's views are amenable to multiculturalism. Interestingly, these same core values identified by Mill formed the essential pillars that upheld normative cultural relativism.

This article argued that even though J.S. Mill is often described as a moral universalist due to the methodology he employed in his postulation of ethical utilitarianism, one could also argue that his position maintained in the *On Liberty* qualifies him to be a normative cultural relativist. Basically, the article contended that beyond the universalist reading of J.S. Mill's works, his views are also compatible with normative cultural relativism. The article examined and interpreted the position maintained by J.S. Mill's in *On Liberty* that committed him to normative cultural relativism. It concluded that anyone that maintains such a position need not reject normative cultural relativism. Instead, Mill's own claims in On Liberty indicate that his position is consistent with core tenets of normative cultural relativism.

[34] Wendy, Donner, "Autonomy, Tradition, and the Enforcement of Morality" in *Mill On Liberty: A Critical Guide*, ed. C.L. Ten, Cambridge: Cambridge University Press, 2008, p.156

[35] J.S. Mill, *On Liberty*, p.67

References

Augenstein, Daniel, "Tolerance and Liberal Justice", in *Ratio Juris*. Vol. 23 No. 4, (2010): 437–59.

Evanoff, Richard J., "Universalist, Relativist, And Constructivist Approaches to

Intercultural Ethics" in *International Journal of Intercultural Relations*, vol. 28 (2004): 442.

Gomberg, P. (1994) "Universalism and Optimism" in *Ethics*, vol.104, No.3:536–557

Harman, G., and J. J. Thomson (1996), *Moral Relativism and Moral Objectivism*, Cambridge: Blackwell Publishers, p.5

Harman, G. (1998) "Precis of Moral Relativism and Moral Objectivism", *Philosophy and Phenomenological Research*, vol. LVIII, No.1:161

Mackie, J. L. (1977), *Ethics: Inventing Right and Wrong*, London: Penguin Book.

MacIntyre, A. (2007), *After Virtue: A Study in Moral Theory*, Notre Dame: University of Notre Dame Press.

Mill, J. S. (1859) *On Liberty*, Ontario: Batoche Book Limited.

Mill, J. S. (1906) *Utilitarianism*, Chicago: The University of Chicago Press.

Mill, J. S. (1865) *An Examination of Sir William Hamilton's Philosophy*, London: Longman.

Mill, J. S. (1991) *Collected Works of J.S. Mill*, ed. J. M. Robson, London: Routledge.

Platonov, R. S. (2019) "Moral Universality in J. S. Mill's Utilitarianism" in *Russian Journal of Philosophical Sciences*. Vol. 62, no. 11: 84–95.

Platonov, R. S. (2020) "*Moral universality in the ethics of classical utilitarianism*" in *Антиномии*, vol. 20, No.4,:45-64.

Rachels, James (2012) *The Elements of Moral Philosophy*, New York: McGraw-Hill.

Tyndal, Jason (2013) "Culture and Diversity in John Stuart Mill's Civic Nation" in *Utilitas* Vol. 25, No.1: 96.

Waldron, Jeremy (2008) "Mill and Multiculturalism", *Mill's On Liberty: A Critical Guide*, ed. C. L. Ten Cambridge: Cambridge University Press. pp. 165–84.

Wendy, Donner (2008) “Autonomy, tradition, and the enforcement of morality” in *Mill On Liberty: A Critical Guide*, ed. C.L. Ten, Cambridge: Cambridge University Press.

Wilson, Fred (2016) “John Stuart Mill”, *The Stanford Encyclopedia of Philosophy*, (Spring Edition) Edward N. Zalta.

INDIVIDUAL AUTONOMY AS NORMATIVE PROMISE AND OUTCOME OF HIGHER EDUCATION: IMPLICATIONS FROM AXEL HONNETH'S THEORY OF RECOGNITION

Asen V. Dimitrov

Institute of Philosophy and Sociology
Bulgarian Academy of Sciences, Bulgaria
email: asen.dimitrov.17@alumni.ucl.ac.uk
ORCID ID: https://orcid.org/ 0000-0003-0454-5983

Abstract:

This article suggests an explorative approach to the character and meaning of individual autonomy in higher education through a particular conception of social relations. Following Axel Honneth, a theoretical framework is outlined in which subjectivity and autonomy are results of underlying institutionalised relations of recognition in the forms of love, legal respect, and social esteem. It is argued that when higher education is seen as self-formation and as a field of relations of recognition due to its collegial, legal, and evaluative features, it is possible to conceive of its institutional environment as supportive or obstructive to recognition and, consequently, to the advancement of individual autonomy. Such a theoretical tool, based on Honneth's recognition theory, could then translate into a critical instrument able to extend the understanding of what constitutes matters of individual autonomy in and across specific institutionalised contexts, including higher education.

Key words: higher education, autonomy, recognition, students, freedom.

Introduction

The idea of multiple goals and functions of higher education has been widely established (Boyadjieva & Ilieva-Trichkova, 2016; Kerr, 1963; Zgaga, 2009). Both theory and research demonstrate that the analysis of effects or outcomes of higher education, whether private or public, extrinsic or intrinsic, requires a detailed view on the interconnectedness between individual characteristics – since education primarily concerns individuals – and the institutional and organisational features of higher education placed in a wider social setting (Marginson et al., 2023).

As "higher education is bound up in, and is a key player in, the formation of modern society" (Barnett, 1990, p. 68), its contemporary role and significance requires careful thought on the desired or desirable aspects of formation and transformation which concern individuals and societies alike. This implies an ethical dimension of higher education in addition to its commonly accepted role in social, political, economic, or purely epistemological problems. In a democratic context, for example, it is rather difficult not to conceive of higher education – education as a whole, or even democracy itself, as a matter of fact – as intrinsically linked with the notion of freedom or individual autonomy. Democratic education mandates the upbringing of informed and knowledgeable individuals able to critically assess statements and put forward arguments as well as act upon one's interests, among other goals. Yet, when it comes to practical terms, what individual autonomy stands for in institutionalised and interpersonal relations – and how these are structured and developed – is a matter of frequent dissent, hence in need of further exploration.

Some of this discord behind the notion stems from applying specific perspectives on the idea of 'self' as integral to individual autonomy, while other disagreements are based on variations in the theoretical and disciplinary treatment of the notion. For example, Ryan and Deci's (2017) primarily psychological account of self-determination admits autonomy as a possibility without insisting on the idea of moral autonomy, which would be found in Kantian approaches. The theory in focus in this article, Honneth's theory of recognition (1995) as a theory of social relations, arguably bridges different disciplinary and philosophical conceptions to come up with an explanation of how individual autonomy comes about and what conditions shape its development. His approach makes it possible to talk about autonomy without a comprehensive preset definition of the concept since individual autonomy is conceived as a result and as content of functioning relations of recognition (Ibid.). Therefore, the exploration of relations of recognition, what they mean in practical terms, and what conditions them may well be in a position to shed different light on the meaning and character of autonomy.

Institutions of society contain and reflect relations of recognition (Honneth, 1995). Looking at higher education as a social institution (Meyer, 1977; Meyer et al., 2007; North, 1991) provides for an applied social philosophical perspective on the matter of autonomy as it allows institutionalised norms or institutionalised rationality and practices in higher education to be examined with regard to individual outcomes, and autonomy in particular. Furthermore, it is individual higher education institutions (HEIs) as organisations (Clark, 1983) which realise higher education in practical terms. A two-tier account of higher education – as social institution and also in an organisational sense – further allows for a complex view on its overarching functions and roles as an idea and as a social institution (Barnett, 2022; Meyer, 1977; Meyer et al., 2007). This approach also provides solid ground for both theoretical and contextual analyses of relations of recognition as preconditions of individual autonomy, thus opening up viable perspectives for empirical work and assessment of implications.

This heuristic potential of this endeavour is contained in the juxtaposition of individual autonomy as a result of relations of recognition and the notion of individual autonomy as one of the goals of higher education. Autonomy is a constitutive value among other key values, beliefs, principles, and norms of higher education which contribute to "a *critical* concept that provides standards such that educational processes […] can be assessed as to the extent to which they fulfil the criteria implied in the idea of higher education" (Barnett, 2011, p. 3, emphasis in original). Without aiming to be extensive on the matter of 'implied criteria', higher education is considered, ideally speaking, a bastion of pursuing and upholding truth as well as a relatively safe and free space for the generation and dissemination of knowledge but also for the nurturing of the human mind and being (Barnett, 1990; 2022; Macfarlane, 2017; Marginson et al., 2023). This article adopts Marginson's notion of self-formation as an umbrella concept of these generalised missions of higher education. He addresses them as processes of self-development through immersion in knowledge to which autonomy is an integral individual trait (Marginson, 2023, p. 62). At the onset of an exploration of the conditions for self-formation in higher education, no functional features of higher education as a social institution should be discounted for their potential to support or hinder autonomy. A conscious focus on learning or teaching, as well-intentioned and thoughtful as it may be, risks partiality with regard to the development of the whole person or human mind (Macfarlane, 2017; Stojanov, 2020). A wider theoretical view could instead reflect on higher education's role in "legitimating society's cognitive structures" (Barnett, 1990, p. 8) both in terms of valid knowledge and patterns of value which get internalised by individuals in the entirety of their experience as students. This is where Honneth's theoretical work seems particularly apt.

The article relates Honneth's theoretical proposition to the institutional and organisational context of higher education. For the purpose of the critical discussion, it would be helpful if Honneth's recognition is seen as a stand-alone approach which is distinct from the works of other critical theorists, for example Judith Butler, Nancy Fraser, or even Charles Taylor, as well as one which has significantly progressed from its ontological roots if they are to be sought in 20th century critical theory (for a summary, see Wilhelm, 2021, Chapter 1). The article begins by outlining the place and role of individual autonomy as an expression of freedom, justifying the adopted assumption that individual autonomy is a normative element of higher education but also as one fundamentally belonging to the liberal democratic order. Then, Honneth's theory of recognition is briefly summarised in its part which relates to the mechanism and contents of relations of recognition and the impact of institutionalised patterns of value *on* and also *as* conditions for recognition. The theoretical projection of Honneth's recognition onto higher education would point towards a logical conclusion – if autonomy results from interpersonal relations and their institutionalised forms, then developing autonomy in higher education students, or other individuals in different institutional contexts, also has to be seen as a process dependent on the quality, or character, of these relations. Autonomy as relational does not appear as a clear-cut conclusion of social relations; it is gradually built and never complete or stable. In this sense, relations of recognition already contain some level of autonomy as they develop further. This is the normatively desired mode of being as long as social conditions support instead of hinder continuous intersubjective recognition and, consequently, autonomy. The theoretical adaption of the forms of recognition for higher education institutional and organisational rules and practices outlines such desired modes of conduct and value which, arguably, help advance individual freedom.

Individual autonomy as a normative promise of higher education

Following Honneth's theory of recognition (1995; 2007; 2014b; 2021b) but also his elaborate work on the idea of freedom (Honneth, 2014a), it might be worth starting from the rather obvious point that individual autonomy, as a norm and a value, sources its validity outside of higher education. Having in mind a liberal democratic institutional order, Honneth names "freedom, i.e. the autonomy of the individual" (Honneth, 2014a, p. 15) as probably the single most important notion which has impacted the development of modern democratic institutions. Their development has thus reflected on the legal protection of individual freedom but also on a deeper understanding of "the value of human subjects [which] lies in their capacity for self-determination" (Ibid., p. 16). In this sense, building autonomy, or conditions under which individuals could develop and exercise an enhanced capacity for individual agency and self-reflexivity, is a general political project – an endeavour which is non-specific to higher education and which, at least in theory, develops holistically within the space opened by a liberal democratic order and its institutions.

The history of higher education closely links higher education and the notion of freedom in epistemic but also in political and social terms as both integral to the processes of knowledge generation and dissemination and to the effects or outcomes of higher education which concern individuals and societies (Barnett, 2022; Marginson et al., 2023). Varying interpretations of education as personal development, gaining knowledge, and raising awareness of self and the world, where self-development has taken different forms when situated in specific historical and social contexts, date back to the writings of Aristotle through Newman, Dewey, and Jaspers, and are as relevant as ever in present-day higher education research (Arthur & Bohlin, 2005; Barnett, 1990; 2022; Macfarlane, 2017; Marginson, 2018; 2023). The specific institutional and organisational orders of democracy and modernity have led to seeing individuals as reflexive selves (Giddens, 1991) able to shape themselves and affect others and the surrounding world in a process of negotiation and conflict with established structures. The emphasis, however, rests increasingly on individual agency as a key factor of individual transformation and a key area of contribution of higher education (Marginson, 2023, p. 81).

Higher education has famously been defined as self-formation and "expansion of freedom" (Marginson, 2018, p. 3; 2023). According to this conception, self-formation is dependent on autonomy as one of its elements but it is also about generation of agency. Marginson argues that self-formation in higher education happens through conscious immersion in knowledge, as well as the ability to verbalise it, and is a mediator between self and the world (Marginson, 2023, p. 78). Self-formation could be seen as a process of development of one's capacity for reflexivity and agency as common denominators of the cognitive and behavioural aspects of the idea. The right to independently decide what development direction to take is an attribute of freedom in political, social, and epistemological terms, and so are the freedom of mind, based on positive relations-to-self, and capacity and actual ability to act upon one's life projects without undue constraints yet with respect for others. This is an idealised normative image of individual capacities which are subject to realisation in a "social-relational" context (Marginson et al., 2023, p. 4). Ideally, social conditions would support the advancement of freedom in positive and negative terms but also in its intersubjective, or social, dimension, which characterises communicative practice (Honneth, 2014a).

Freedom, however, is not a given outcome by far. In higher education and individual HEIs, different goals compete for validity and all of them are realised through institutionalised forms – norms, values, beliefs, roles, policies and procedures, as well as specific physical spaces. As desired outcomes, autonomy or even the realisation of the umbrella concept of self-formation are also influenced by existing institutional conditions. The extent to which freedom is intrinsic to a specific organisational setting in higher education or has been rationalised into its rules and practices should not be taken for granted. This paradoxical situation of having freedom as a definitive aspect of higher education without being sure freedom does characterise higher education practices allows and mandates the assessment of the institution and its organisations for their ability to deliver on their normative content.

It might help to distinguish freedom in a metaphysical sense, as an abstract freedom of mind, and freedom as individual liberty or autonomy which is shaped – sustained and curtailed – by institutions. Freedom as institutionally-embedded further possesses particular organisational characteristics, and organisations exercise what Giddens (1991) calls "regularised control of social relations" (p. 16). An important implication for institutions and their organisational structures is the ethical requirement thus imposed on processes and practices if institutions, as well as their organisations, are to support rather than hinder individuals' autonomous functioning or remain blind by choice to its social value.

If functioning relations of recognition result in individual autonomy, then, in order to assess higher education's actual performance in delivering on its own norm, it would be appropriate to critique the conditions for relations of recognition in higher education as a field of recognition (Boyadjieva & Ilieva-Trichkova, 2021). The very mechanism of relations of recognition, explained in the following section and adapted to higher education, contains and generates a notion of intersubjective autonomy which is dependent on interpersonal communication, institutionalised social relations, and patterns of value affecting and modifying relations and formal structures (Honneth, 1995; 2007; 2010; 2014b; 2021b).

The social mechanism of recognition

In a classical liberal approach, the democratic idea of the autonomous individual could be generated through the logic of rational justification of individuals' rights and responsibilities (Honneth, 2014b, p. 153; Sandel, 1982). Analogically, in higher education theory, autonomy's normativity usually comes from rationalised epistemological concerns and historical, social, and political developments which justify the endorsement of a liberal conception of the individual as an independent and potentially able knower or learner. The realisation of higher education as an extension of freedom depends on the provision of opportunities for self-realisation alongside the protection of individual rights (Macfarlane, 2017; Marginson, 2023). Honneth's theory of recognition reaches similar endpoints through different routes. It is based on an in-depth analysis of the development of western European democratic societies and is thus aligned with liberal values and principles – but also socialist and communitarian views – which are found in democratic higher education. In Honneth's work, the realisation of legitimate individual goals is clearly denoted as a simultaneous contribution to a common endeavour (Honneth, 1995), which may be seen as a devalued point in liberalism. Autonomy, as it surfaces in Honneth's diagnosis of society, is also a function and result of treating relations of recognition as constitutive of both individuals' subjectivity and the integration of societies (Honneth, 1995). Honneth's autonomy is therefore not an ideologically presupposed norm. He demonstrates that individual autonomy is a consequent outcome of specific social relations, a "certain type of individual relation-to-self that allows us to be confident of our needs and beliefs, and to value our own capacities" (Honneth, 2014b, p. 41). These specific social – interpersonal or institutionally-conditioned – relations are characterised by mutual respect and consideration of morally and also rationally justified claims for recognition of needs, rights, and aspirations or achievements which individuals, or groups, place forward against institutionalised norms when existent norms suppress legitimate individuality (Ibid.). These and other relevant aspects of Honneth's theory are explained below.

Honneth's conceptualisation of the notions of subjectivity and autonomy – respectively described as "decentred" (Honneth, 2007, p. 182) and weakened (Vatsov, 2006, p. 51) – allows for a broad, non-deterministic analytical lens on the character and contents of individual autonomy. This wide lens is, namely, relations of recognition as they arguably generate subjectivity and autonomy (Honneth, 1995). Therefore, understanding the widest possible scope of relations of recognition and their functioning or malfunctioning in specific contexts should lead to knowledge about the practical dimensions of autonomy or aspects of its failure. This knowledge could then be matched against existent images of individual or, in particular, student autonomy, with regard to students as knowers or learners as well as individual members of society.

Contrary to any critique of autonomy which sees it as resulting in a kind of social atomism, be it individualism or any other, the very existence of others is a basic condition for realising one's own autonomy (Honneth, 2014a, p. 45). Making sense of others' existence, however, refers to more than simply cognising them. In Honneth's view, which is also based on Hegel's treatment of the idea of recognition, the act of social interaction with others contains in itself not simply cognition but also recognition. Not all relations, however, are relations of recognition. Recognition is an act of "reciprocal granting of a normative status" (Honneth, 2021a, p. 57). In this sense, seeing others – either as different or identical, approvingly or not – would simply be an act of cognition, not recognition. Not all social assessment turns communicative exchanges into relations of recognition. The assessment of value in social relations would go beyond preference to an aspect of recognition if it concerns one's "normative worth" (Honneth, 2014b, p. 47) – if in some way evaluation addresses the individual as a moral being or regards the moral capacity of individuals to develop and retain minimally positive, "productive" (Honneth, 1995, p. 92) relations-to-self – self-trust, self-respect, and self-esteem. Such relations-to-self preserve individuals' capacity to interact with others and be seen as equal partners (Honneth, 1995). The understanding of social valuation which takes place within relations of recognition is thus dependent on the cognitive and behavioural features of the participants in a given communicative framework. Individuals' willingness and capacity to put forward ideas and intentions and to treat them as valuable is shaped by the reflection of traits, abilities, or life plans in the explicit or implicit assessment of others (Ibid.). The mutual evaluation and justification of social worth and its effects on interpersonal relations lie in the basis of Honneth's recognition theory and underscore the relational character of any results of relations of recognition, including autonomy.

Under relations of recognition, Honneth refers to relatively stable habitual or institutionalised patterns (Honneth, 1995; 2014b), or "bundles of behavioural norms" (Honneth, 2014a, p. 45). Sediments of recognition can be found in instances of interpersonal and institutional exchanges and are conditioned by institutionalised norms and patterns of cultural value (Honneth, 1995). Although individuals are able to use degrees of agency to overcome or align with such patterns, people tend to act somewhat unconsciously or sometimes, under unsurmountable social pressure, to follow habitual or institutionalised patterns of behaviour while internalising them with time (Honneth, 2021b, p. 143). Hence the importance of patterns, especially dominant ones. They might be plural and inclusive of individuated biographies but might also effectively act as limitations on individual life-projects. Institutionalised patterns are reflective of and materialised in the constitutive institutions of society – family, law, and work – and their corresponding forms of recognition as love, respect, and esteem (Honneth, 1995). These institutions and forms of recognition shape individual experience alongside a varying degree of individual agency. They develop and intertwine rather palpably in higher education as well.

The forms of recognition – love, respect, and esteem – are not only adaptable into higher education normative terms but are also formative for individuals, especially students. Trust and support in collegial and friendly relations which develop due to emotional proximity are reflective of the idea of nonromantic love (Honneth, 1995). Norms and rules clearly stand for law or legal respect where everyone is entitled to equal treatment and shared duties (Ibid.). Achievement and aspirations define the idea of disseminating knowledge and advancing society, which shape the legitimacy and functioning of higher education and its individual HEIs as well as their particular assessment policies and processes in this regard. The symbolic and material expressions of the forms of recognition – as internalised values but also corresponding behaviour – arguably condition the development of individual autonomy in higher education students.

Justified pluralisation of institutional policies and practices, which leave space for all students to be recognised for their legitimate needs, rights, and aspirations and thus develop their autonomy, would then depend on these broadly adapted forms of recognition in every HEI. Higher education as an institution in the sense of institutional theory overlaps semantically with the idea of institutions and their effects in the Hegelian sense, as "complexes of regulated behaviour" and "lasting embodiments of intersubjective freedom" (Honneth, 2014a, p. 53). From a normative perspective, the connection is deeper than simply semantics when higher education is normatively seen as advancing freedom, which, following Honneth, happens through functioning relations of recognition. Therefore, the shaping and application of value patterns in and through higher education and its processes is a key concern in the analysis of the relations of recognition which span beyond higher education in terms of individual and social effects. This is so because interpersonal relations are framed by institutional processes, and the institutions' norms, beliefs, values, and dominant rationalisations, or practical forms of these symbolic elements, regulate and legitimate some modes of behaviour or value patterns against others (North, 1991).

The whole possible set of processes in higher education might be responsible for effects on students in terms of regulation of behaviour and aspirations as well as internalisation of such limitations, with possible consequences on the development of agency but also, more generally, on students' relations-to-self and relations-with-others. In other words, this is a reminder that learning needs to be treated as a general psychological process of individual formation through everyday experiences, including through relations of recognition. In higher education, just like in society generally, students continue to develop self-awareness and depend as individuals on some level of positive relations-to-self, including on a positive perspective on the future. Students as individuals also "*learn* [emphasis added] to understand themselves as legal persons owed the same autonomy as all other members of society […] [and] as subjects possessing capabilities and talents that are valuable for society" (Honneth, 2003, p. 142). The intelligible patterns of value students encounter or not, whether institutionalised or habitual, appear as formative to their conception of self-value, their actual self-esteem, and the extent to which they see themselves as equal participants in social interaction.

These theoretical presuppositions, as selective or limited as they may be, raise important points concerning the conditions favouring students' autonomous choices but also nurturing the autonomy of students as individuals in society. Firstly, the very idea of justification, which easily finds normative backing in higher education, is reconfirmed as both a norm and a value. Mutuality in recognition depends on some level of shared value orientation, including when it comes to social esteem as assessment of accomplishments. Yet this shared value horizon leaves space for respectful dissent and includes a common basic understanding of the principles of justification as key to recognition (Honneth, 1995). What can be framed as "the possibility to retreat from routine practices of interaction in the lifeworld or to question them morally" (Honneth, 2014a, p. 123) could be an expression of freedom. Secondly, at any point in time individuals possess some level of knowledge and cognitive capacities which frame their epistemic and psychological capability to engage in justification. In this regard, knowledge of self, others, and the world is a valuable source of justification power, but it needs to be complemented by self-reflection and criticality as to differentiate cognitive facilities favouring social reproduction from a mode of thinking which allows agential impact on self and others. And lastly, in what is not an extensive list, these individual capabilities for justification will have to translate into behavioural terms as to gain or enhance actual agency. The success of individuals in processes of justification with consequent practical impact depends on their level of mastery of the modes of conduct which, in particular situations, would allow them to successfully justify their own moral stances and life-projects. Thus, the metaphysical notion of freedom is in fact qualified in practical terms. Freedom is firstly restricted by the dependence of individuals on institutionalised forms of behaviour and related value patterns, and secondly, by personal restraint, which characterises recognising others and engaging in justification (Honneth, 2021b, p. 140).

It is hardly difficult to imagine how individual autonomy depends on some level of participation in institutionalised forms of life, and how in higher education such participation is indeed framed by demands of mutual support, legal description of rights and responsibilities, and valuation of interests or assessment of results. Equally plausible is the image of higher education processes in which individuals do not come across one or all forms of recognition on a reciprocal basis. While functioning relations of collegiality, respect, and social esteem provide conditions for the development of autonomous individuals, their dysfunctionality hinders this process. Besides malfunctioning of relations of recognition, there is also misrecognition – instances of moral offence, disrespect, denigration, or bodily harm (Honneth, 1992; 2007), which could be systemic or symptomatic of failing conditions for recognition. The diagnosis of misrecognition is equally important to the examination of failures in recognition, be it in the narrow institutional or wider social experience of individuals. It is the act of withholding or depriving of moral or normative status – possibly seen as a reflection of social and institutional conditions rather than individual desires or choices to withhold recognition – which can be classified as misrecognition or which would characterise or underlie failing relations of recognition.

Honneth may have proposed an optimistic theory, but his analytical framework is neither being naïve of social reality nor claiming utopian purposes. Deliberation on claims for recognition would not always lead to clear, mutually agreed upon and followed solutions. Recognition cannot be full;it leaves space for ambiguity, and it is sometimes instrumentalised for adverse means and manipulation (Ikaheimo et al., 2021). The manipulative use, or abuse, of recognition could logically be expected to result in manipulated outcomes. Recognition may also be curtailed by insufficient material resources to reflect a fair assessment of needs or worth. Nonetheless, the described mechanism of recognition as generative of autonomy already reveals the imprudence of thinking about autonomy as social atomism or of encouraging autonomy as a cause – or the cause – of social alienation or disintegration. It is important to denote the difference between, on the one hand, radical individualism or forms of social pathology as failing recognition, and, on the other hand, autonomy in its definition as necessarily intersubjective and an outcome of desirable social relations of recognition. Manipulation and disregard of others would qualify as misrecognition and as damage to the very conditions for autonomy.

The idea of 'learning' how to relate and communicate with others with respect and to ourselves with respect and esteem nicely highlights a goal in the formation of individuals in higher education which complements disciplinary knowledge and professional competences. Although higher education is about advanced knowledge, it is also an institution which socialises, classifies, and legitimises (Meyer, 1977; Meyer et al., 2007). The functions of higher education, expressed through a set of processes, would provide conditions for the advancement of individual autonomy, or for relations of recognition, if the processes reflect specific "modes of conduct" (Honneth, 2015, p. 27) which characterise relations of recognition. Through 'walking the talk' and building shared normative perspectives, HEIs 'teach' their students legitimate or valued modes of conduct as they convey, or sometimes distort, higher education's institutional norms. If higher education is to advance the freedom of all participating individuals or students, in addition to furthering knowledge, it has to engage in shaping the communicative practice of recognition as behaviour, thinking of self and others, and awareness of the importance and value of recognition for the freedom of the individual.

On the practical applicability of the approach

The application of Honneth's concept of recognition to higher education, when higher education is seen as self-formation and extension of freedom, already transpires through the provided theoretical summary. Honneth's proposition, however, depicts autonomy building as a challenging continuous process. While 'full' autonomy is utopian, an intentional endeavour of fostering autonomy in individuals requires attending to a plethora of dynamic factors. Such targeted efforts concern the elimination of institutionalised misrecognition as exclusion of legitimate individual traits and unfair treatment, but also mandate the opening up of relatively safe exploratory space and expanding each one's "capacity to be open to the many sides of one's own" (Honneth, 1999, as cited in Znepolski, 2006, p. 137).

While continuing to base the argument in the field of higher education, it is possible to illustrate the applicability of Honneth's approach to seeing autonomy as institutionally embedded in principle. Building on knowledge of the practical meaning of autonomy is a continuous task under the assumption that the conditions for autonomous functioning, or conditions for relations of recognition, may evolve or change over time and place, just as claims for recognition have expanded based on the historical intensification of processes of individualisation (Honneth, 1995, p. 93). Higher education is of course only an example of a social institution in which the significance and meaning of autonomy is particularly amplified. Other institutional or organisational settings may invoke different or complementary aspects of autonomous being. Maintaining a larger point to be made, institutional features as conducive or not to recognition and autonomy ought to be seen as socially-embedded themselves, and so are the contents of autonomy as dependent on larger political, social, and cultural influences.

Autonomy, as already noted, is an element of higher education as well as a key outcome which carries meaning and sources legitimacy and content beyond the field of higher education. The autonomy of a citizen and the autonomy of a student are only different to the extent they refer to general or specific institutional orders; furthermore, autonomy in students is superseded by the idea of individual autonomy attributed to everyone in a democratic order. Education as learning and education in terms of the German *Bildung*, for example, are both processes, if at all separate, which characterise human psychology and socially constructed values before they are employed in educational terms. It is only secondary that education or learning have been structured in what is known today to be schooling institutions. This is not to disregard the importance of educational structures in any way; instead, this brings the analytical attention to the underlying psychological and social conditions for learning and self-formation in a broad context, which underlie and potentially impact the role of a particular institution in this process.

Following educational theory and the logic of Honneth's theory of social relations, education has both reproductive and transformative roles which are functions of individual agency but also of the dominant political and cultural patterns of values in a given societal place and time. To an extent, it is telling that schools are named as they are, and HEIs are indeed classified as a type of school in some countries. It is schooling that is directly underscored in their title as opposed to learning, self-development, or self-formation; schools embody certain societal practices and values which are reproduced or transformed by schooling, and this does not necessarily suggest plurality. Higher education is supposed to be normatively different from schooling, and yet, both share social context. The idea of personal development as free self-formation and the need for basic social reproduction supported by education may complement each other but could also enter into conflict and prevent legitimate social or individual transformations.

Contemporary democratic education aims to revise the relationship between reproduction and transformation, or between schooling and learning, but in no way abandons the idea of social reproduction. At the very least, educational systems are still strongly dependent on national states. This balance is revised by advancing the individual as a building block of societies. In education and especially in higher education, individual student autonomy is normatively prioritised as "a socially mediated development of mindedness which is directed toward freedom" (Stojanov, 2020, p. 1). Nurturing autonomy means preparing students to learn and think independently, but equally so it means equipping them with the knowledge, skills, and mindset to act upon their life goals alone and in collaboration with others. Such a vision of education may be promoting a normative image but in no way pushes for absolute normative alignment. Parallelly to higher education, relations of recognition develop in the spheres of the family, law, and public space or work, and are subjected to modifications based on the justification of claims for recognition. The autonomy of individuals as outcome may thus vary in substance based on the spheres of recognition prioritised by individuals in their personalised claims. For example, reasonable individual choices would sometimes be justified preferences towards the reproduction of social circumstances, oppressive or not, for example against taking perceived risks in seeking social mobility or other kinds of social transformation higher education may be in position to support. In what is not a particularly subtle example, children may decide to take care of sick parents at the expense of personal professional growth. The analysis of relations of recognition need not attribute value to the particular choice *per se* or, even less so, to the individual making it. Instead, the conditions and value patterns which underlie such choices are to be examined for offering genuine rather than subtly restricted plurality of choice and opportunity and for their potential effects on individuals and societies. While the analysis may focus on a particular form of recognition or particular social institutions, the wider social context may also have an enabling or disabling effect on individual chances, for example by lack of provision of sufficient legitimate forms of support which address mutual needs.

Honneth's recognition as a critical theory simultaneously allows for a diagnosis of the existent value frameworks at a given place and point of time and a critical assessment of the reproductive or transformative potential of social institutions. The attention to specific context, taking into consideration macro-, or institutional, meso-, as organisational, and micro-level factors at the level of the individual is crucial to understanding the conditions for recognition. In higher education, empirical analyses will need to account for larger cultural, political, or epistemological predispositions of individuals and institutions but also for specific organisational features and interpersonal dynamics. Suppression or withholding of recognition is plausible across all levels. For example, even micro-level interpersonal exchanges between peers deserve attention for their potential to damage one's relations-to-self, capacity for connecting meaningfully with others, or eventually one's ability to further participate in the processes of higher education *and* take part in them as an equal partner. Thus, potential psychological damages or benefits even of incidental communication need not be seen strictly in narrow particularistic terms, as solely one's own. Their potential for structural effects should be examined instead.

The normative content – recognising individual needs, rights, and aspirations or accomplishments – characterises the theoretical framework of recognition as open and capable of capturing dynamic interactions of individual specifics, power disbalances, and conflicting moral positions. Justifying individual claims, or "articulating needs through language" and "affective reactions" (Honneth, 2007, p. 188) has significance across social life in general and different processes in higher education, whether purely epistemic or concerning formalised, administrative processes. Nonetheless, regardless of any legitimate needs, rights, and aspirations one may hold, not all individuals are equally eloquent or gain sufficient command of cultural facilities supporting agential capacities such a relations-to-self and skills for framing personal perspectives into claims for recognition. Also, not all individuals have a good grasp on the value and role of recognition in interpersonal and institutional relations and their outcomes. The expansion of one's language competences and value horizon regarding recognition enhances the individual's ability to engage in justification and, thus, potential to reinforce one's autonomy. Therefore, in addition to expanding students' disciplinary knowledge, applied skills, and self-awareness, HEIs could improve their autonomy-building potential by targeting students' understanding of individuals' situatedness in the social world and the influence of value patterns and valuation on people's judgement and behaviour. 'Teaching' this in a classical sense may only be complemented by the inclusion of these principles and beliefs in the rules and practices which students encounter in higher education and thus tend to internalise. This non-patriarchal developmental perspective to the individual characterises aspects of the adaption of recognition to higher education's institutional context.

Of course, knowledge of 'good' practices or advances in disciplines or institutional settings are to be taken into consideration and themselves evaluated when exploring conditions for recognition and their sediments in practical life. Many existent policies and practices in higher education have already been evaluated for their potential to support students' autonomy-building. Such potential could be seen through different lenses or revitalised through the prism of recognition. For example, when conceived through the forms of recognition, the idea of 'student services' may be relieved from a sense of 'customer service' or 'serving' as failing mutuality, which is often inferred nowadays and which is sometimes used as an argument to not apply such institutional interventions. Undoubtedly, encouraging extracurricular activities or providing psychological support and sensitive career advice augment the formational impact of HEIs by addressing a larger variance of potential student needs and aspirations and complementing traditional instruction and research. If applied to student funding, the sensible consideration of needs alongside merit may provide an additional instrument to support enrolment and retention and enhance the quality of participation, hence life chances of some individuals. This should not necessarily be read as a reference to the US system although US undergraduate application and enrolment processes do provide an example of resource redistribution through fees and funding, not without undesired consequences which require redress. Thus, the design, content, and effects of processes on participating individuals can be thought through the interaction of forms of recognition in the institutional and organisational context but also in broader social context.

An example of autonomous decision-making in higher education is the right to choose majors which individuals best connect with their dynamic life-projects or the right to amend their choice as well as to participate in the design of their study programme to a reasonable yet freedom-enhancing degree. This right may be hindered by overly rigid institutional arrangements which redirect or restrict students because, at the institutional level, divergent values or ideas of valuable knowledge are prioritised against individual choice of educational and professional paths. The proposition does not imply that students should be able to momentarily alter their choices and unrestrictedly jump from module to module but argues that restrictions should be reasonable and designed with the individual student in mind. It also does not imply that such choices are completely independent of the opinion and advice of others. Extreme levels of student self-reliance may as well undermine "agentic formation" (Marginson, 2023, p.

79). Prohibiting students from following their desired paths, however, may not only limit individual potential and autonomy, but may also damage the underlying conditions for relations of recognition – one's relations-to-self, for example, or one's intrinsic motivation to connect and develop with others in particular circumstances – while studying for unwanted degrees and being prepared for employment in less desired industries, for instance.

These and other examples would benefit a more detailed elaboration within particular contexts and only serve as initial illustration of implications limited by their generality. The characterisation of autonomy as relational – dependent on relations-to-self and relations-with-others – provides an analytical lens to individuals' lifelong development as conditioned by time- and place-specific organisational, institutional, and wider social orders with their respective dominating or conflicting value models. Breaking down experiences into sets of processes in particular institutions with concrete participants narrows down the field of argument to be made. However, the influence of institutions on individuals and their own role in self-formation remains a highly nuanced problem. This complexity is fully present in higher education when it comes to defining the institutional and intersubjective conditions supporting autonomy with a plural and diverse student body in mind; diversity might have to be seen beyond traditional sociological terms in order to reflect subtler differences in needs and life projects. It also comes to show that building autonomy hardly happens by chance and that targeted efforts are required to first understand and then address legitimate needs, rights, and aspirations of individuals as a way to create and improve conditions which foster individual autonomy.

Conclusion

Relying on Marginson's definition of higher education as self-formation and Honneth's theory of recognition, this article conceptualises individual autonomy as a constituent normative promise of contemporary democratic higher education which benefits an exploration as a relational, non-relativistic yet context-dependent concept. The social mechanism of recognition, or misrecognition, traceable to institutional and organisational rules, policies, and practices, allows higher education's own norms and values to be tested for persistence in their practical application. The answer of this analytical problem would be positive as long as higher education and individual HEIs provide conditions for functioning relations of recognition. Such conditions would include the possibility, and would target the capacity of individuals, to engage in justification of claims for recognition and pursue legitimate individual life-projects without shame, fear, or undue restrictions. The empirical application of this critical framework might also illuminate the way norms and their application change or not as they are subjected to pressure from patterns of perceived social worth, direct claims for recognition, or individual preferences which warrant additional elaboration and distinction from recognition claims.

The article takes the conversation in the direction of reinforcing the normative value of individual autonomy. An additional brief argument might be worth putting forward. If the normativity of autonomy – not its exclusivity amongst the multitude of functions and purposes of higher education – in individuals' development is rejected, it is certainly unclear what other or alternative norms could replace the idea of the autonomous person, i.e. the person who is informed but also critical, reflexive, and capable of pursuing independent goals. Such a negation would be at minimum awkward if not self-defeating, especially if coming from within a field which has the pretence of surviving and retaining social value because it is built of autonomous institutional actors who provide critique to power rather than submit to it. Rather than rejecting the normativity of autonomy in higher education, a productive focus might as well be a discussion of the character and contents of autonomy as to clarify and possibly bridge conceptual (mis)understandings.

Honneth shows that the importance of autonomy, as great as it is for a democratic normative order, in practice comes as a second-order social effect – as a result of recognising individual needs, rights, and aspirations or achievements. This undoubtedly amplifies the value of relations of recognition for and in social institutions. Of course, the mechanism of recognition opens up a series of theoretical questions regarding the meaning and content of legitimate needs, rights, and aspirations, but also suggests these questions are taken up in particular social and institutional contexts. Vigilance would be required as to not turn the defence of the worth of particular individual traits into an instrument for a relativist assessment of the world.

Higher education as an institution is built on norms which do not suggest a relative reading of knowledge. Awareness of the world, others, and oneself, conditioned by rationality and mutuality, remains an underlying element of the individual capacity to engage productively in relations of recognition. It matters, however, what value is attributed to both knowledge claims and claims for recognition, as well as to their justification, as such processes characterise students' formative experience as individual members of societies. If higher education as a modern democratic institution is to deliver on its normative promise to foster agency and individual autonomy, it has to see students not just as learners but as individuals in a liberal sense, as persons of equal worth and potential although with different accomplishments. Higher education has to go beyond a narrow epistemic focus on knowledge and take account of the intersubjective character of autonomy, i.e. of accounting for individual needs, rights, and aspirations and the wider social and political context and implications of knowledge endeavours. This is indeed an extensive ground when defined as a field of ethical concern, if not responsibility, in higher education. An explicit or passive rejection of the importance of this scope is in itself an act of misrecognition when it comes to the developmental potential of individuals and communities.

Acknowledgements

This work was supported by the Bulgarian National Science Fund, National Science Program VIHREN under contract number КП-06-ДВ-2/16.12.2019, project "Dynamics of inequalities in participation in higher and adult education: A comparative social justice perspective".

References

Arthur, J. & Bohlin, K. (eds.). (2005). *Citizenship and higher education. The role of universities in communities and society*. London and New York: Routledge Falmer.

Barnett, R. (1990). *The idea of higher education.* Buckingham: SRHE and Open University Press.

Barnett, R. (2011). *Being a university*. London and New York: Routledge.

Barnett, R. (2022). *The philosophy of higher education*. *A critical introduction.* London and New York: Routledge.

Boyadjieva, P. & Ilieva-Trichkova, P. (2016). Rethinking missions and values of higher education: Insights from the capability approach and the institutional perspective. In R. Pritchard, A. Pausits & J. Williams (Eds.). *Positioning higher education institutions: From here to there* (pp. 41–62). Rotterdam: Sense Publishers.

Boyadjieva, P. & Ilieva-Trichkova, P. (2021). *Adult education as empowerment: Re-imagining lifelong learning through the capability approach, recognition theory and common goods perspective*. London: Palgrave MacMillan.

Clark, B. (1983). *The higher education system. Academic organization in cross-national perspective*. Berkeley: University of California Press.

Giddens, A. (1991). *Modernity and self-identity*. *Self and society in the late Modern Age*. Cambridge: Polity.

Honneth, A. (1992). Integrity and disrespect: Principles of a conception of morality based on the theory of recognition. *Political Theory, 20*(2), 187–201.

Honneth, A. (1995). *The struggle for recognition. The moral grammar of social conflicts*. (J. Anderson, Trans.). Cambridge: Polity.

Honneth, A. (2003). Redistribution as recognition: A response to Nancy Fraser. In N. Fraser & A. Honneth. *Redistribution or recognition? A political-philosophical exchange* (pp. 110–197). (J. Golb, J. Ingram & C. Wilke, Trans.). London: Verso.

Honneth, A. (2007). *Disrespect. The normative foundations of critical theory.* Cambridge: Polity.

Honneth, A. (2010). *The pathologies of individual freedom. Hegel's social theory*. (L. Loeb, Trans.). Princeton: Princeton University Press.

Honneth, A. (2014a). *Freedom's right. The social foundations of democratic life*. (J. Ganahl, Trans.). Cambridge: Polity.

Honneth, A. (2014b). *The I in We: Studies in the theory of recognition*. (J. Ganahl, Trans.). Cambridge: Polity.

Honneth, A. (2015). Education and the democratic public sphere: A neglected chapter of political philosophy. In O. Lysaker & J. Jakobsen (Eds.). *Recognition and freedom: Axel Honneth's political thought* (pp. 17–32). Leiden and Boston: Brill.

Honneth, A. (2021a). Intelligibility and authority in recognition: A reply. In H. Ikaheimo, K. Lepold & T. Stahl (Eds.). *Recognition and Ambivalence* (pp. 55–60). New York: Columbia University Press.

Honneth, A. (2021b). *Recognition: A chapter in the history of European ideas*. (J. Ganahl, Trans.). Cambridge: Cambridge University Press.

Ikaheimo, H., Lepold, K. & Stahl, T. (Eds.). (2021). *Recognition and ambivalence.* New York: Columbia University Press.

Kerr, C. (1963). *The uses of the university*. Boston: Harvard University Press.

Macfarlane, B. (2017). *Freedom to learn. The threat to student academic freedom and why it needs to be reclaimed.* Abingdon and New York: Routledge.

Marginson, S. (2018). *Higher education as self-formation.* London: UCL IoE Press.

Marginson, S. (2023). Higher education as student self-formation. In S. Marginson, B. Cantwell, D. Platonova & A. Smolentseva (Eds.). *Assessing the contributions of higher education: Knowledge for a disordered world* (pp. 61–87). Cheltenham: Edward Elgar Publishing.

Marginson, S., Cantwell, B., Platonova D. & Smolentseva, A. (Eds.). (2023). *Assessing the contributions of higher education: Knowledge for a disordered world*. Cheltenham: Edward Elgar Publishing.

Meyer, J. (1977). The effects of education as an institution. *American Journal of Sociology, 83*(1), 55–77.

Meyer, J., Ramirez, F., Frank, D. & Schofer, E. (2007). Higher education as an institution. In P. Gumport (Ed.). *Sociology of higher education* (pp. 187–221). Baltimore: Johns Hopkins University Press.

North, D. (1991). Institutions. *Journal of Economic Perspectives*, *5*(1), 97–112.

Ryan, R. & Deci, E. (2017). *Self-determination theory: Basic psychological needs in motivation, development, and wellness.* New York: Guilford Publications.

Sandel, M. (1982). *Liberalism and the limits of justice.* Cambridge: Cambridge University Press.

Stojanov, K. (2020). *Education, self-consciousness and social action. Bildung as a neo-Hegelian concept*. London and New York: Routledge.

Vatsov, D. (2006). *Svoboda i priznavane. Interaktivnite izvori na identichnostta.* [Freedom and recognition: The interactive sources of identity]. Sofia: New Bulgarian University.

Wilhelm, D. (2021). *Axel Honneth: Reconceiving social philosophy*. Lanham: Rowman and Littlefield.

Zgaga, P. (2009). Higher education and citizenship: 'the full range of purposes'. *European Educational Research Journal*, *8*(2), 175–188.

Znepolski, B. (2006). From individuality to subjectivity or from subjectivity to individuality. *Critique and Humanism, 22*, 131–138.

THE ROLE OF HUMAN VALUES IN THE RECEPTION OF THE ISTANBUL CONVENTION WITHIN THE EUROPEAN UNION

Hristo Valchev*, Magdalena Garvanova**

*Guangdong University of Foreign Studies
Guangzhou, China
email: hristo_valchev@yahoo.com
ORCID ID: https://orcid.org/0000-0002-1170-7621

**University of Library Studies and Information Technologies,
Sofia, Bulgaria
email: m.garvanova@unibit.bg
ORCID ID: https://orcid.org/0000-0003-2387-1442

Abstract:
In the present paper we investigate the role of human values in the reception of the Convention on Preventing and Combating Violence Against Women and Domestic Violence (better known as 'the Istanbul Convention') within the EU. By establishing a certain relationship between the most important arguments raised in favor of or against the convention and values from Schwartz's theory of human values, we formulate certain hypotheses about the influence of the importance the people from a given EU member state ascribe to certain values on the decision of that EU member state on whether the convention should be ratified. We also hypothesize that human values have played the role of mediator between the decision in question and the waves of immigrants from the Middle East and North Africa to the EU. All hypotheses are tested using the results from the European Social Survey. These results suggest that: 1) a more positive attitude towards the convention is correlated with higher importance of the value 'universalism-concern' and lower importance of the value 'tradition'; 2) the value 'universalism-concern' has played the role of mediator described above.

Key words: human values; the Istanbul convention; arguments; the European Union; immigrants.

1. Introduction

The Convention on Preventing and Combating Violence Against Women and Domestic Violence, better known as 'the Istanbul Convention', is a human rights treaty of the Council of Europe, which was opened for signature on 11 May 2011, in Istanbul, Türkiye. Its reception within the European Union (EU) has generally been positive, but it has not been universally accepted. All 27 current members of the EU have signed it, but only 21 have ratified it, and despite facing pressure from the EU[1], some of the remaining 6 countries have even formally rejected it. In 2018 the Constitutional Court of Bulgaria decided that the convention did not comply with the constitution of the country. In 2020 the convention was rejected by the parliament of Slovakia. The

[1] On 28 November 2019, the European Parliament adopted a non-legislative resolution calling on the member states that had not ratified the convention to urgently do so. On 13 February 2020, the European Parliament adopted a resolution on the priorities set by the EU for the next session of the UN Commission on the Status of Women, calling again for the accession of the non-ratifying member states and the EU to the convention.

same year the parliament of Hungary adopted a political declaration also rejecting it. Strong opposition to the convention also emerged in Czechia, Latvia and Lithuania, as a result of which the ratification in these countries was postponed. Furthermore, despite ratifying the convention in 2015, Poland has been seriously considering the option of withdrawal from the treaty.

Most of the research on the reception of the convention within the EU has been conducted from a feminist point of view. Feminist authors have interpreted the opposition to the convention as being a part of a global movement opposing human rights and gender equality and have identified different actors that have participated in this movement, as well as different strategies and frames these actors have employed. The actors in question include NGOs, churches, religious leaders, think tanks, political parties and state actors. Since all EU member states are liberal democracies, however, it can be assumed that apart from these actors, an important role in the decision of a given country on whether the convention should be ratified has also been played by the general public. It *is* recognized that the said actors have appealed to the general public, using tactics such as direct grassroots mobilization, protests and petitioning. Different arguments serving as a basis for these appeals have been identified. The effect these arguments have had on the general public, however, depends not only on their intrinsic characteristics, but also on other factors, including certain psychological facts and processes. In the present paper, we investigate the role in the decision of a given country on whether the convention should be ratified played by such psychological facts and processes.

For this purpose, we make use of the concept of human values. By establishing a certain relationship between the most important arguments raised in favor of or against the convention and values from Schwartz's theory of human values, we formulate certain hypotheses about the influence of the importance the people from a given EU member state ascribe to certain values on the decision of that EU member state on whether the convention should be ratified. Furthermore, we consider the relationship between the decision in question and the waves of immigrants from the Middle East and North Africa to the EU. This relationship is two-fold. On the one hand, one of the most influential arguments against the convention states that its ratification would lead to the negative consequence of decreasing the level of control over the inflow of immigrants. It can be assumed that the waves of immigrants make this negative consequence especially undesirable. On the other hand, research suggests that the waves of immigrants have contributed to the rise of far-right political parties (Inglehart & Norris 2017; Davis & Deloe 2017) and, in general, these political parties tend to oppose the convention. Thus, it can be assumed that the waves of immigrants have influenced the decision on whether the convention should be ratified. We hypothesize that this influence can be explained by the mediating function of human values. All hypotheses are tested using the results from the European Social Survey (ESS).

2. Human Values

The question of what human values are and how they differ from motives, needs, traits, attitudes and narrow goals is addressed in Arieli et al. (2020), Garvanova (2013, 2017a), Maio (2010), Sagiv and Roccas (2021), Verplanken and Holland (2002), etc. Human values can be understood as cognitive representations of basic motivations as goals to be pursued. As opposed to certain narrow or specific goals that people might have, human values are transsituational, i.e. apply across situations and over time. For example, if a person ascribes high importance to the value of personal security, they will seek physical, emotional and material security in every relationship and on all occasions. Furthermore, human values serve as standards for evaluating, judging and justifying one's own, as well as other people's choices. They are socially desirable; therefore, people can use them to gain cooperation in pursuing their own goals by communicating about them with others. Another significant characteristic of human values is that they vary in importance. Each person ascribes different importance to different values, thus ordering them in a hierarchy of importance.

The most widely used theory of the content and structure of human values is that of Schwartz (1992). The theory is based on the following three assumptions (see Sagiv & Schwartz 2022). First, values stem from three universal requirements for human survival: (a) biological needs of the organism, (b) interactive needs for interpersonal coordination, and (c) group needs for welfare and survival. Second, values form a coherent structure based on the degree of compatibility or conflict between the goals they express. This assumption implies that values form a circular structure in which the more compatible any two values are, the closer they are going around the circle, and the more in conflict, the more distant they are. Third, the circle of values is a continuum, in which values blend into one another rather than forming discrete entities. This implies that one can partition the value circle arbitrarily into as many or as few value categories as is convenient. The original theory partitioned the continuum into the 10 values shown in the center of Figure 1, but subsequent examination of people’s responses to value surveys led Schwartz (Schwartz 2017, Schwartz et al. 2012) to refine the theory by partitioning the value circle into 19 values for use when more fine-tuned analyses are desired. As evident from Figure 1, higher-order distinctions were also made. The distinction between personal focus and social focus values roughly corresponds to the distinction between individualism and collectivism.

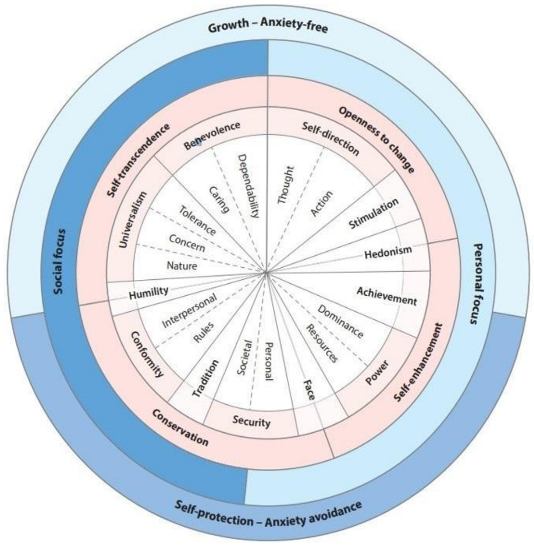

Figure 1. The structure of human value according to Schwartz’s theory (taken from Sagiv & Schwartz 2022)

This structural model was originally proposed on purely conceptual grounds, but empirical research largely supported it (see Schwartz 2015 for a summary). A total of 13 instruments have been devised in order for the importance of the values from Schwartz’s theory to be measured (see Roccas et al. 2017). Most instruments use rating as the response format, but some use ranking. Rating provides the absolute importance of each value; ranking provides its relative importance. Ranking assumes that people can distinguish the importance of every value from that of every other. Ranking’s disadvantages are that several values may actually be equally important, that one cannot apply traditional methods for assessing reliability for a ranked scale, and that people can feasibly rank only a limited number of values. Rating overcomes these disadvantages. Moreover, researchers can derive value priorities from ratings. However, rating is more subject to response biases (see below). Pairwise comparisons and best–worst scaling combine elements of both rating and ranking.

Both the absolute and the relative importance of the values from Schwartz’s theory have been shown to be relatively stable over time (see Milfont et al. 2016, Vecchione et al. 2016). Yet, they *can* change under circumstances that result from certain important events in people’s lives. Thus, migration to societies with different values may expose people to pressures to adapt their values to those of their new societies. Longitudinal studies of immigrants from Poland to the United Kingdom (Bardi et al. 2014) and from Russia to Finland (Lönnqvist et al. 2011) found mean-level value change. Changes were also found in the absolute and relative importance of values for Bulgarian students, teachers and businessmen in the context of a transition from socialism to liberal democracy (Garvanova 2005, 2013; Garvanova and Papazova 2017). Entering motherhood, women shifted their value priorities toward greater conservation versus openness to change (Lönnqvist et al. 2018). Following the global financial crisis of 2008, security, tradition, and benevolence values increased in importance, whereas hedonism, self-direction, and stimulation values decreased in importance among youth and young adults in 16 European countries (Sortheix et al. 2019). Ongoing studies are investigating possible value changes in response to the COVID-19 pandemic (Daniel et al. 2021).

The importance of the values from Schwartz’s theory has also been shown to have certain implications for attitudes, preferences and behavior, including life satisfaction (Garvanova 2017b), religiosity (Schwartz & Huismans 1995), competitiveness (Lönnqvist et al. 2013), anti-social behavior (Aquilar et al. 2018; Benish-Weisman et al. 2017; Feldman et al. 2015), pro-social behavior (Sagiv & Roccas 2021), pro-environmental behavior (Schultz & Zelezny 1998), attitudes towards immigrants (Davidov & Meuleman 2012), attitudes towards homosexuality (Donaldson et al. 2017), etc. In the area of politics, values have consistently predicted left versus right political ideology in Western societies. The values of tradition, conformity and security have been shown to predict right-wing ideology, whereas universalism has been shown to predict left-wing ideology (Piurko et al. 2011; cf. Aspelund et al. 2013; Caprara et al. 2017). Some research suggests that acts of political activism such as signing petitions or demonstrating are positively correlated with self-transcendence versus self-enhancement and with openness to change versus conservation values (Vecchione et al. 2015). Certain correlations between values and attitudes towards foreign policies have also been found (Dwyer & Coymak 2020).

3. The Relationship between Values and Practical Arguments

Research suggests that argumentation is culturally specific (Ju 2020; Ju et al. 2021). On the other hand, the specifics of a culture are directly expressed in its values. The relationship between values and practical arguments can be explained in the following way (see Valchev 2022). A practical argument is an argument to the conclusion that a certain course of action should be taken. The reasons that are put forward in support of such a conclusion can be interpreted as goals whose realization the given course of action is a means for. Thus, the structure of practical arguments can be expressed through a widely recognized argumentation scheme that can be called "Basic Instrumental Practical Reasoning" and formulated in the following way (Macagno and Walton 2018):

> Premise 1: Agent A has a goal G.
> Premise 2: Carrying out action B is a means to realize G.
> Conclusion: Therefore, A should bring about action B.

As noted above, human values can be understood as transsituational goals. Thus, 'goal G' from the Basic Instrumental Practical Reasoning argumentation scheme can be understood as referring to a certain value. It can be assumed that the higher the importance of that value for a given person or a given group, the higher the strength of the given practical argument for that person or group.

This assumption can be used as a basis for a method for analyzing qualitative data in which actors that disagree about the best course of action in a given situation raise practical arguments in support of their respective positions. Such data can be coded using the Basic Instrumental Practical Reasoning argumentation scheme. The coding process would include identifying agent A, goal G and action B. Goal G can be identified as a member of a given set of human values. Thus, each argument following the Basic Instrumental Practical Reasoning argumentation scheme would correspond to a certain value. This relationship between the arguments and the set of values can be used as the basis for formulating hypotheses about the relationship between an actor's position on which the best course of action would be and the importance the actor ascribes to certain values. More particularly, it can be hypothesized that the actor ascribes higher importance to the values that correspond to the arguments he or she has raised and lower importance to the values that correspond to the arguments other actors have raised. Such hypotheses can be tested by looking at results from empirical studies that measure the importance the actors ascribe to the values. In the present inquiry, we made use of this method.

We gathered data about the arguments raised in favor of or against the Istanbul Convention from different studies on the reception of the convention within the EU. Most work has been done by Krizsan and Roggeband, who have analyzed a large amount of data from the EU member states where opposition to the convention has been the strongest by identifying actors that have supported or opposed the convention, distinguishing between different strategies that these actors have employed and making comparisons between the different countries (see Krizsan & Roggeband 2021, 2020, 2019). Other studies that focus on the overall reception of the convention in the EU include Balogh (2020) and Leye et al. (2021). Studies that focus on the reception of the convention in individual EU member states, on the other hand, include Chiritoiu (2019), Gaweda (2017), Sutlovic (2019), Szczygielska (2019), Valchev (2022), etc.

Using the results from these studies, we identified different arguments following the Basic Instrumental Practical Reasoning argumentation scheme. For the purposes of the present inquiry, we identified only arguments that have been raised in favor of or against the convention, and not arguments that have been raised in order to support or attack certain premises of the arguments raised in favor of or against the convention. Thus, the variable 'action B' had two possible values – ratifying and not ratifying the convention. The variable 'agent A' was also interpreted as having two possible values – the proponents of the convention and the opponents of the convention. Thus, the values taken by 'agent A' and 'action B' were in a sense known. In order to identify the most common arguments raised in favor of or against the convention, we also identified the most common values taken by the variable 'goal G'.

The variable 'goal G' was interpreted as referring to one of the 19 values from Schwartz's refined theory. For convenience, we divided the process of identifying references to these values into two stages. The first stage consisted in identifying specific goals that the proponents and the opponents of the convention referred to. In most cases these goals were not stated directly, i.e. via expressions such as "we have this-and-this goal", "we want to achieve this-and-this", etc. Thus, their identification did not consist in a mere description, but in a reconstruction of the arguments under consideration. The second stage, in turn, consisted in mapping the specific goals onto the 19 values. The relationship between the specific goals and the values can be explained in two ways. On the one hand, we can say that a specific goal is an instantiation of the transsituational goal a given value consists in. On the other hand, we can also say that the realization of a specific goal would promote a given value. It must be noted that we could not expect a one-to-one correspondence between the specific goals and the values. The same specific goal might promote multiple values and the same value might be promoted by multiple specific goals.

However, since we didn't have any information about the value systems of particular proponents or opponents of the convention, the hypotheses we put forward were not about the value systems of the proponents and the opponents of the convention, but about the value systems of whole countries. We hypothesized that the values that correspond to the arguments of the proponents of the ratification would have higher importance in the EU member states that have ratified the convention and lower importance in the EU member states that have not ratified the convention, whereas the values that correspond to the arguments of the opponents of the convention would have higher importance in the EU member states that have not ratified the convention and lower importance in the EU member states that have ratified the convention. These hypotheses were tested using the results from the European Social Survey (ESS).

4. Arguments in favor of and against the Istanbul Convention

In the debates about the convention, its proponents have played a rather defensive role, mainly responding to the criticism against it. When they do raise arguments to the conclusion that the convention should be ratified (as opposed to arguments to the conclusion that a certain argument against the convention is not sound), they normally refer to the benefits of realizing the purposes of the convention. Thus, we can say that the goals that the arguments in favor of the convention refer to coincide with the purposes of the convention. It must be noted, however, that even though domestic violence is stated in the convention alongside violence against women, the convention on the whole, as well as its proponents focus mainly on violence against women. The proponents of the convention have often described in detail the severity of the problem of violence against women, citing statistics, personal experience, etc. In some instances, male partners, relatives, or acquaintances have been identified as the violators. Furthermore, for the purposes of the present inquiry, we can ignore the references to particular mechanisms such as supporting organizations and law enforcement agencies, etc. Thus, we can distinguish between three broader goals that have been referred to as reasons why the convention should be ratified – the goal of decreasing the level of violence against women, the goal of decreasing the level of domestic violence, and the goal of increasing the level of equality between men and women.

The arguments that have been raised against the convention, on the other hand, refer to both problems that are discussed in the convention and problems that are not. Balogh (2020) distinguishes between three kinds of arguments that have been raised against the convention – "feminist", "libertarian" and "conservative". The conservative arguments have played by far the most important role. Most of them have been related to the concept of gender, used in the convention. This concept is central for the convention because 'violence against women' is defined as referring to "all acts of gender-based violence that result in, or are likely to result in, physical, sexual, psychological or economic harm or suffering to women" (Article 3-a). The term "gender" is defined as referring to "the socially constructed roles, behaviours, activities and attributes that a given society considers appropriate for women and men" (Article 3-c) and is thus explicitly distinguished from the term "sex", which refers not to a social construction, but to a biological category. The opponents of the convention argue that the introduction of such a distinction between gender and sex would have two bad consequences – that of providing an opportunity for increasing the rights of same-sex couples by allowing them to get married and adopt children and that of promoting and providing a legal dimension to the idea that one can choose whether to be a man or a woman (or neither a man nor a woman).

These two bad consequences, however, are not said to be bad in themselves. The most common reason why they are said to be bad is that they are inconsistent with certain traditional and religious beliefs. Accordingly, Article 12-1 of the convention, which legislates the eradication of certain traditions, is seen as problematic. The traditional religion of all EU member states is Christianity, and Christianity in general condemns homosexuality and the possibility of choosing whether to be a man or a woman (or neither) as inconsistent with God's will. It can be assumed that all traditions that the opponents of the convention refer to stem from Christianity, but nevertheless Christianity is not always mentioned alongside these traditions. In the debates in Hungary, for instance, religion played a relatively minor role. Thus, we can distinguish between two different goals that are stated as reasons why the convention should not be ratified – preserving the national traditions and preserving the Christian faith.

Another common reason why the two consequences referred to above are said to be bad is that they are inconsistent with the existing laws of the given country. Thus, the Constitutional Court of Bulgaria decided that the concept of gender as defined by the convention did not exist in the legal system of the country, and therefore the convention was not consistent with the constitution. In the countries where same-sex marriages are not legal, the opponents of the convention also referred to the definition of marriage as a union between a man and a woman. Thus, we can say that another goal that is stated as a reason why the convention should not be ratified is that of adhering to the existing national laws. Apart from the existing national laws, the two consequences referred to above are also said to be inconsistent with 'the natural law', 'the natural order' or nature in general. The opponents of the convention argue that the convention denies 'natural differences' or 'natural human rights' and that it defines men and women 'artificially'. Even references to science have been made. Thus, we can say that another goal they state as a reason why the convention should not be ratified is that of preserving the natural order.

A further reason why the two consequences are said to be bad is that they serve the purpose of imposing on the people of the given country a certain ideology, referred to as 'gender ideology'. However, the imposition of an ideology implies an undesired cultural change, and the cultural characteristics of a society are to a large extent determined by its traditions, religion and legal system. Thus, it can be assumed that when the opponents of the convention speak of 'gender ideology', they are not referring to any goal that is different from those of preserving the national traditions, preserving the Christian faith and adhering to the existing national laws. A final common reason why the two consequences are said to be bad is that they would somehow undermine 'the family'. The term is used differently in different contexts, but it does not refer to anything different from the traditional family, the family as it is understood by the Christian faith or the family as a part of the natural order. Thus, when the opponents of the convention argue that the two consequences would undermine the family, they are not referring to any new goal either.

There are, however, two additional arguments that are commonly raised in relation to the family. The first one refers to Article 14-1 of the convention, which legislates the implementation into the educational system of topics like "non-stereotyped gender roles". Such legislation is seen as problematic, not only because it would serve the purpose of promoting the idea that one can choose whether to be a man or a woman (or neither), but also because it would deprive parents of their right to instill their values in their children. However, this right is not seen as stemming from anything different from the national traditions, the Christian faith or the natural order. Thus, this argument does not refer to any new goals either. The second argument, on the other hand, states that the convention positions men and women as two mutually opposed classes – women as victims and men as violators – and thus disturbs the peace in the family. Here we do have a new goal that is said to be a reason why the convention should not be ratified – that of preserving the peace in the family.

Finally, the opponents of the convention have also raised arguments that refer to Articles 60-2 and 66-1 of the convention respectively. Since Article 60-2 legislates granting a refugee status to people who are persecuted on grounds related to their gender, the opponents of the convention argued that it could be used as an easy way for immigrants to enter the given country. Thus, we can say that not ratifying the convention is seen as a means to realize the goal of avoiding a decrease in the level of control over the inflow of immigrants. Article 66-1, on the other hand, legislates the existence of a certain monitoring mechanism. This legislation is seen as a way of providing an opportunity for intervening in some affairs that are internal to the country. Thus, we can say that the goal the argument refers to is that of preserving national sovereignty.

5. Formulating Hypotheses and Testing Them

In the previous section we identified three specific goals that are stated as reasons why the convention should be ratified – decreasing the level of violence against women, decreasing the level of domestic violence, and increasing the level of equality between men and women. We also identified seven specific goals that are stated as reasons why the convention should not be ratified – preserving the national traditions, preserving the Christian faith, adhering to the existing national laws, preserving the natural order, preserving the peace in the family, avoiding a decrease in the level of control over the inflow of immigrants, and preserving national sovereignty. These goals are related to the 19 values from Schwartz's refined theory in the following way.

Since a decrease in the level of violence in general entails an increase in the level of security, and violence against women and domestic violence are related to the security of the individual (as opposed to the society), we can say that realizing the goal of decreasing the level of violence against women and the goal of decreasing the level of domestic violence would promote the value 'security-personal'. Furthermore, since the goal of decreasing the level of violence against women implies that women are a demographic who are somehow discriminated against and the value 'universalism-concern' includes the ideas of equality and social justice, we can say that realizing the goal of decreasing the level of violence against women would also promote the value 'universalism-concern'. The same value would also be promoted in the case that the goal of increasing the level of equality between men and women is realized.

Since the value 'tradition' covers all traditional beliefs, including those handed down by religion, we can say that it would be promoted in the case that the goal of preserving the national traditions and the goal of preserving the Christian faith are realized. As far as the existing national laws are a reflection of the national traditions and the idea of what is natural comes from the traditions and religion of the given society, the same value would also be promoted in the case that the goal of adhering to the existing national laws and the goal of preserving the natural order are realized. Furthermore, since immigrants might have different traditions and religious beliefs, they might be perceived as impeding the preservation of one's own traditions and religious beliefs. Thus, it can be assumed that the value 'tradition' would also be promoted in the case that the goal of avoiding a decrease in the level of control over the inflow of immigrants is realized. Lastly, as far as an uncontrolled inflow of immigrants and loss of national sovereignty can be seen as threats for the society as a whole, we can say that realizing the goal of avoiding a decrease in the level of control over the inflow of immigrants and the goal of preserving national sovereignty would promote the value 'security-societal'.

Thus, we can say that on the whole the arguments of the proponents of the convention correspond to the values 'universalism-concern' and 'security-personal', whereas the arguments of the opponents of the convention correspond to the values 'tradition' and 'security-societal'. However, since 'security-personal' and 'security-societal' are right next to each other on Schwartz's value continuum (together they form the value 'security' from the original theory), people generally ascribe to them similar importance. Thus, it can be assumed that their influence on the decision of a given EU member state on whether the convention should be ratified has been relatively small. Taking the importance of the other two values into consideration, we formulated the following two hypotheses:

Hypothesis 1. The importance the people from the EU member states that have ratified the Istanbul Convention ascribed to the value 'universalism-concern' at the time a decision to ratify the convention was made is higher than the importance the people from the EU member states that have not ratified the convention ascribed to the same value at the time a decision not to ratify the convention was made.

Hypothesis 2. The importance the people from the EU member states that have ratified the Istanbul Convention ascribed to the value 'tradition' at the time a decision to ratify the convention was made is lower than the importance the people from the EU member states that have not ratified the convention ascribed to the same value at the time a decision not to ratify the convention was made.

In order to test these hypotheses, we looked at the results from the different waves of ESS. If a wave of ESS was conducted in a country in the year when a decision about the ratification was made, we used the information about the values of the people from that country provided by the results from that wave. Otherwise, we checked if a wave of ESS had taken place in the country up to two years prior to the decision. If yes, then we used the results from that wave. If no, we excluded the given country from the analysis.

For the countries that have ratified the convention we assumed that a decision was made in the year of ratification. For six of these countries – Austria, Cyprus, Greece, Luxembourg, Malta and Romania – a wave of ESS was conducted neither in the year of ratification, nor within two years prior, therefore these countries were excluded from the analysis. For the countries that have not ratified the convention we checked whether the convention has been formally rejected. Three countries have formally rejected the convention. In 2018 it was rejected by the constitutional court of Bulgaria. In 2020 it was rejected by the parliament of Slovakia. That same year the parliament of Hungary adopted a political declaration also rejecting it. For these three countries we assumed that a decision was made in the year of rejection. Thus, for Bulgaria we used the results from the wave of ESS that was conducted in 2018. The results from the wave of ESS that was conducted in 2020 have not been published yet at the time of writing the present text, thus for Slovakia and Hungary we also used the results from 2018. Czechia and Latvia and Lithuania, on the other hand, have not rejected the convention formally, but around 2018, debates about the convention took place in all three of these countries, as a result of which the ratification was postponed. We interpreted this postponement as also consisting in a decision for the convention not to be ratified, and thus for these three countries we also used the results from 2018.

In order to measure the importance ascribed to the values from Schwartz's theory, ESS uses a version of the Portrait Value Questionnaire, presenting each respondent with short descriptions of people and asking them to indicate the extent to which they are like the person from the given description on a Likert scale from 1 (very much like me) to 6 (not like me at all). We can say that the absolute importance of a given value for a given country is the average Likert scale value ascribed by the respondents from that country to the descriptions that correspond to the value[2]. The relative importance of a given value for a given country, on the other hand, can be defined as the place of the value in the hierarchy of values of that country (as determined by the absolute importance of the values). Thus, for both the absolute and the relative importance of a value it is true that the lower the number expressing the importance, the higher the importance. For the purposes of the present inquiry, we took into account both the absolute and the relative importance of the values 'universalism-concern' and 'tradition'. The orders of the countries according to the absolute and relative importance ascribed to the two values are shown in Table 1 and Table 2 respectively. As evident from Table 1, the six countries that have not ratified the convention are amongst the seven countries ascribing lowest absolute importance to 'universalism-concern', as well as among the eight countries ascribing lowest relative importance to the same value. These results support Hypothesis 1. Table 2, on the other hand, shows that the four countries ascribing highest relative importance to 'tradition' have not ratified the convention, and out of the six countries ascribing highest relative importance to 'tradition', only one has ratified the convention. On the whole, however, the countries that have not ratified the convention do not ascribe higher relative importance to 'tradition' than the countries that have ratified the convention. These results support Hypothesis 2 only partially – they suggest that the hypothesis is correct only as regards the relative importance of 'tradition'.

[2] In general, the version of the Portrait Value Questionnaire used in ESS is not made for measuring the importance of the 19 values from Schwartz's refined theory. For example, it does not contain a description that corresponds to the value 'face'. Yet, it can be used to measure the importance of the values 'universalism-concern' and 'tradition'. One description corresponds to each of the two values. The two descriptions are "He thinks it is important that every person in the world should be treated equally. He believes everyone should have equal opportunities in life." and "Tradition is important to him. He tries to follow the customs handed down by his religion or his family.", respectively.

Universalism-concern (absolute)	
ES	1.62
SL	1.71
SE	1.77
IT	1.78
FR	1.85
PO	1.92
FI	1.95
HR	2.03
BE	2.05
DE	2.08
IR	2.08
NL	2.09
DK	2.11
PT	2.18
SK	2.24
LV	2.49
CZ	2.52
HU	2.57
EE	2.65
BG	2.74
LT	2.79

Universalism-concern (relative)	
FR	1
PT	1
ES	1
FI	2
SL	2
SE	2
IT	2/3
NL	2/3
BE	3
PO	3
DK	3/4
HR	4
IR	4
SK	4
DE	5
LT	5
CZ	6
HU	6
BG	7
EE	8
LV	9

Table 1. The absolute and relative importance ascribed to the value ‘universalism-concern’ by the EU member states according to the results from the European Social Survey. The countries that have not ratified the convention are highlighted. A slash (‘/’) indicates that the value had the same relative importance as another value

Tradition (absolute)	
PO	2.13
BG	2.26
IT	2.26
HR	2.36
SL	2.39
LV	2.44
SK	2.52
CZ	2.60
LT	2.60
IR	2.61
BE	2.71
ES	2.71
DK	2.76
HU	2.76
PT	2.82
NL	2.89
DE	2.91
FI	3.03
EE	3.06
SE	3.18
FR	3.35

Tradition (relative)	
BG	3
LT	3
CZ	7
LV	7
PO	7
SK	7
HR	8
IT	10
PT	10/11
DK	11
HU	11
IR	11
DE	12
NL	12
SL	12
ES	12
EE	13
BE	14
FI	15
FR	15
SE	15

Table 2. The absolute and relative importance ascribed to the value ‘tradition’ by the EU member states according to the results from the European Social Survey. The countries that have not ratified the convention are highlighted. A slash (‘/’) indicates that the value had the same relative importance as another value

It can be noted, however, that the only country amongst the seven countries ascribing lowest absolute importance to ‘universalism-concern’ that as ratified the convention – Estonia – ascribes, relatively speaking, very low absolute importance to ‘tradition’. Thus, as evident from Figure 2, if we present the absolute importance of ‘universalism-concern’ and ‘tradition’ for the different countries on a two-dimensional coordinate system, there is a particular region where all of the countries that have not ratified the convention and none of the countries that have ratified the convention fall. Yet, we can say that on the whole the results from ESS suggest that even though the importance of ‘tradition’ might be playing some role, the importance of ‘universalism-concern’ is a better predictor as regards the decision of a given country on whether the convention should be ratified. We considered the possibility of ‘universalism-concern’ also playing the role of a mediator between this decision and the waves of immigrants from the Middle East and North Africa to the EU over the period since the convention was opened for signature, including the migration crisis that started in 2015.

As we noted earlier, it can be assumed that the waves of immigrants have contributed to a more negative attitude towards the Istanbul Convention. The analysis provided above suggests that the lower the importance of ‘universalism-tradition’, the more negative the attitude towards the convention. Thus, in order for the value to have played the role of a mediator between the waves of immigrants and the decision on whether the convention should be ratified, the waves of immigrants need to have contributed to a decrease in its importance. Such a decrease can be explained through the following psychological process. It can be assumed that the waves of immigrants have been perceived as imposing a threat to economic security (as far as immigrants might take the jobs of the local people) and cultural and religious identity (as far as most immigrants from the Middle East and North Africa have a different cul-

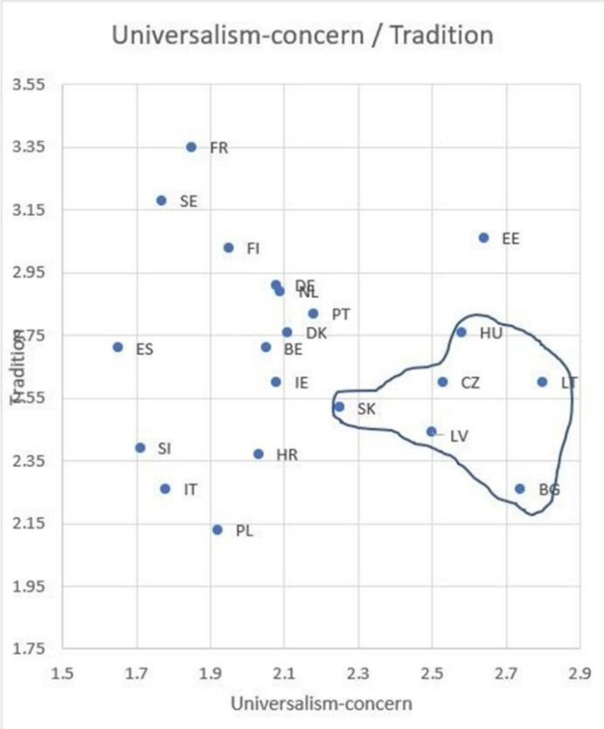

Figure 2. The absolute importance ascribed by the EU member states to the values 'universalism-concern' and 'tradition'. The countries that have not ratified the Istanbul convention are within the enclosed area

tural and religious background). This perceived threat might have made people less likely to believe that everyone should be treated equally, and this belief lies at the heart of the value 'universalism-concern'. In order to operationalize the supposed relationship between the importance of the value and the waves of immigrants, we formulated the following hypothesis:

Hypothesis 3. Since the Istanbul Convention was opened for signature in 2011, there has been a significant decrease in the importance of the value 'universalism-concern' for the EU member states; a decrease in the importance of the value is positively correlated with a more negative attitude towards immigrants.

In order to test the first part of this hypothesis, we conducted a one-factor dispersive analysis and post-hoc tests, using the results from the different waves of ESS, and in order to test the second part of the hypothesis, we conducted a correlation analysis between the Likert scale value ascribed to the description that is used in ESS as an indicator of the importance of 'universalism-concern' and the

Year	2012			2014			2016			2018			F (df), p	Post Hoc Tests, p < 0.05
	N	x	sd	N	x	sd	N	x	sd	N	X	sd		
Universalism-concern	53880	2.04	1.04	39115	2.03	1.04	43567	2.18	1.08	48452	2.20	1.09	339.32 (3,185010),0.000	2012 & 2014 vs. 2016 & 2018

Table 3. Descriptive statistics, one-factor dispersive analysis and post-hoc tests of the absolute importance ascribed to the value 'universalism-concern' according to the results from the European Social Survey. The information is presented on a yearly basis

Likert scale values for the questions from ESS that measure respondents' attitudes towards immigrants.

		Important that people are treated equally and have equal opportunities	Allow many/few immigrants of same race/ethnic group as majority	Allow many/few immigrants of different race/ethnic group from majority	Allow many/few immigrants from poorer countries outside Europe	Immigration bad or good for country's economy	Country's cultural life undermined or enriched by immigrants	Immigrants make country worse or better place to live
Important that people are treated equally and have equal opportunities	Pearson Correlation	x	,146*	,181*	,197*	-,112*	-,157*	-,140*
	Sig. (2-tailed)		,000	,000	,000	,000	,000	,000
	N	48452	45636	45580	47065	46640	46816	46595

Table 4. Correlations between the Likert scale value ascribed to the description that is used in the European Social Survey as an indicator of the importance of the value ‘universalism-concern’ and the Likert scale values for the questions from the European Social Survey that measure respondents’ attitudes towards immigrants

The results are shown in Table 3 and Table 4 respectively. As evident from Table 3, on a total level, the results clearly show a decrease in the importance of ‘universalism-concern’, two main clusters being statistically significant – 2012 and 2014 vs. 2016 and 2018. As evident from Table 4, the Pearson coefficients show a weak but statistically significant correlation ($p<0.01$) in all analyzed pairs. Considering the content of the scales, we can say that a decrease in the importance of ‘universalism-concern’ is positively correlated with a more negative attitude towards immigrants. These results support Hypothesis 3.

6. Discussion

The results of the present study suggest that human values did play a significant role in the reception of the Istanbul Convention within the EU, and more particularly, in the decision of a given EU member state on whether the convention should be ratified. The results also suggest that human values played the role of a mediator between this decision and the waves of immigrants from the Middle East and North Africa to the EU. Thus, the results are in line with research suggesting that human values do play an important role in one’s psychic life, and therefore should be a central topic of psychology. Furthermore, the results are in line with research suggesting that human values do influence the reception of certain policies. Such research provides evidence that policy makers should take human values into account in their endeavors. As far as they are supposed to serve the people, they should make sure that their policies are in accordance with the value system of the people, and as far as they are supposed to change society for the better, they should make conscious efforts to change the value system of the people so that it is in accordance with any supposedly beneficial policy.

We can say that the results of the present study also suggest that the conflict between the proponents and opponents of the convention represents a conflict between the values ‘universalism-concern’ and ‘tradition’. Both of these values have a social (as opposed to personal) focus; therefore, any conflicts between them are not conflicts between individualism and collectivism. It can be speculated that a significant number of women support the convention as a way of pursuing the individualistic value ‘power-dominance’, but do not state this explicitly in order to make their position more socially acceptable. This is an interesting proposition that can be a subject of further research. If we take the arguments of the proponents of the convention at face-value, however, we can say that the conflict between the proponents and the opponents of the convention is not a conflict between individualism and collectivism. Thus, in this case, the concepts of individualism and collectivism have little explanatory power. A better explanation is provided by a distinction between three epochs in human history – the epoch of traditional culture, modernity and postmodernity.

In the epoch of traditional culture, human behavior was governed by religious and moral beliefs, as well as by customs and traditions. We can say that high importance was ascribed to the conservation values, including ‘tradition’. This epoch was replaced by the epoch of modernity, during which human behavior was instead governed by the ideas of progress, rationality, individualism, national unity and state sovereignty. The transition from traditional culture to modernity entailed an increase in the importance of the personal focus (i.e. individualistic) values. It first began in the West, during the sixteenth and seventeenth centuries, but gradually spread across the rest of the world as well. By becoming modern, the West was able to impose its will on the rest of the world and to force it toward modernization. Modernity, however, was in turn challenged by postmodernity. Postmodernity is a very recent phenomenon, dating back only to the 1970s and 1980s. In this last epoch human behavior is governed by the ideas of emancipation, liberation from oppressive structures, justice without borders, cultural pluralism, and the celebration of difference and diversity. These ideas correspond to the universalism values, including ‘universalism-concern’. They are currently widespread mainly in the West and especially in the EU.

The EU is a postmodern political entity, which is evident from the document stating its values – the Charter of Fundamental Rights of the EU – as well as from the fact that it challenges the sovereignty of its members (Korab-Karpowicz 2019:181). Thus, membership in the EU entails a transition from modernity to postmodernity. The transition from one epoch to another, however, requires time and is often not a smooth process (Korab-Karpowicz 2019:173). Thus, membership in the EU entails conflicts between the ideas of postmodernity and the ideas of modernity. On the other hand, the EU member states located in Eastern Europe, including those that have not ratified the Istanbul Convention, started transitioning from traditional culture to modernity relatively late, which means that in these countries the ideas of traditional culture are also relatively widespread. Thus, for these countries membership in the EU also entailed conflicts between the ideas of postmodernity and the ideas of traditional culture. We can say that the conflict between the proponents and the opponents of the Istanbul Convention is one such conflict.

However, the results of the present study suggest that the value ‘universalism-concern’ has played a more prominent role in the reception of the convention than the value ‘tradition’. On the other hand, the results from ESS suggest that the importance of ‘universalism-tradition’ for the EU member states has decreased over the period since the convention was opened for signature in 2011. Thus, it can be assumed that if a given EU member state had considered the question of the ratification of the convention at a different point in time, its decision could have been different. This assumption is consistent with the fact that Poland has been seriously considering the option of withdrawal from the convention. The results from the wave of ESS that took place in 2018 suggest that the importance of ‘universalism-concern’ and ‘tradition’ for Poles at that time was quite similar to the importance the people from the countries that have not ratified the convention ascribed to the two values at the time a decision not to ratify the convention was made.[3]

It must also be noted that the present study had certain limitations. One important limitation stems from the fact that the importance of both ‘universalism-concern’ and ‘tradition’ is measured by only one question in ESS. Using Scwartz’s refined theory allowed us to conduct a more detailed analysis, but the higher level of detail came at the expense of a lower level of reliability of the way the importance of each value is measured. Accordingly, if we had used Scwartz’s original theory or any of the other higher-level distinctions between value cate-

[3] The results from the wave of ESS that took place in 2018 also suggest that the importance of ‘universalism-concern’ and ‘tradition’ for Italians at that time was quite similar, but no significant opposition to the convention emerged in Italy. Apparently, there are additional factors that play a role, but these factors lie beyond the scope of the present inquiry.

gories, each value category would have been measured by several questions, and thus the results would have been more reliable but less detailed and less interesting. Another important limitation of the study stems from the fact that we did not have any data that would allow us to compare the attitude of an individual participant towards the Istanbul Convention and the importance that participant ascribes to human values. Such data can be gathered as a part of future research on the topic.

7. Conclusion

In the present paper, we investigated the role human values play in the reception of the Istanbul Convention within the EU. First, we identified the most important arguments raised in favor of or against the convention. Then, we established a relationship between these arguments and values from Schwartz's theory of human values. We used this relationship in order to formulate two hypotheses about the influence of the importance the people from a given EU member state ascribe to certain values on the decision of that EU member state on whether the convention should be ratified. The first hypothesis stated that higher importance of the value 'universalism-concern' is correlated with a more positive attitude towards the convention, whereas the second hypothesis stated that higher importance of the value 'tradition' is correlated with a more negative attitude towards the convention. Furthermore, we hypothesized that the value 'universalism-concern' has played the role of a mediator between the waves of immigrants from the Middle East and North Africa to the EU and the decision of that EU member state on whether the convention should be ratified. All three hypotheses were tested using the results from ESS. On the whole, these results supported the hypotheses, the second hypothesis being only partially supported.

References

Aquilar, S., Bacchini, D., Affuso, G. (2018). Three-year cross-lagged relationships among adolescents' antisocial behavior, personal values, and judgment of wrongness. *Social Development*, *27*(2), 381–400, https://doi.org/10.1111/sode.12279

Arieli, S., Sagiv, L., & Roccas, S. (2020). Values at work: the impact of personal values in organisations. *Applied Psychology*, *69*(2), 230–75, https://doi.org/10.1111/apps.12181

Aspelund, A., Lindeman, M., & Verkasalo, M. (2013). Political conservatism and left–right orientation in 28 eastern and western European countries. *Political Psychology*, *34*(3), 409–17, https://doi.org/10.1111/pops.12000

Balogh, L. (2020). The ratification status of the council of Europe's Istanbul convention among EU member states. *MTA Law Working Papers*, 7.

Bardi, A., Buchanan, K. E., Goodwin, R., Slabu, L., & Robinson, M. (2014). Value stability and change during self-chosen life transitions: self-selection versus socialization effects. *Journal of Personality and Social Psychology*, *106*(1), 131–147, https://doi.org/10.1037/a0034818

Benish-Weisman, M., Daniel, E., & Knafo-Noam, A. (2017). The relations between values and aggression: a developmental perspective. In S. Roccas, L. Sagiv (Eds.), *Values and behavior: Taking a cross cultural perspective* (pp. 97–114). Springer International Publishing, https://doi.org/10.1007/978-3-319-56352-7_5

Caprara, G. V., Vecchione, M., Schwartz, S., Schoen, H., Bain, P. G., et al. (2017). Basic values, ideological selfplacement, and voting: a cross-cultural study. *Cross-Cultural Research*, *51*(4), 388–411, https://doi.org/10.1177/1069397117712194

Chiritou, A. (2019). The convoluted gender policy processes in Romania: Institutional disarray and activist resilience. In A. Krizsan, C. Roggeband (eds.), *Gendering democratic backsliding in Central and Eastern Europe: A comparative agenda* (pp. 161–199). Centre for Policy Studies, Central European University.

Davidov, E., & Meuleman, B. (2012). Explaining attitudes towards immigration policies in European countries: the role of human values. *Journal of Ethnic and Migration Studies*, *38*(5), 757–75, http://dx.doi.org/10.1080/1369183X.2012.667985

Davis, L., & Deole, S. (2017). Immigration and the rise of far-right parties in Europe. *ifo Institut - Leibniz-Institut für Wirtschaftsforschung an der Universität München, München*, *15*(4), 10–15.

Donaldson, C. D., Handren, L. M., Lac, A. (2017). Applying multilevel modeling to understand individual and crosscultural variations in attitudes toward homosexual people across 28 European countries. *Journal of Cross-Cultural Psychology*, *48*(1), 93–112, http://doi.org/10.1177/0022022116672488

Feldman, G., Chao, M. M., Farh, J. L., Bardi, A. (2015). The motivation and inhibition of breaking the rules: Personal values structures predict unethicality. *Journal of Research in Personality*, *59*, 69–80, http://dx.doi.org/10.1016/j.jrp.2015.09.003

Garvanova, M. [Гарванова, М.] (2017a). *Bulgarian at a value crossroads* [Българинът на ценностен кръстопът]. Sofia: For Letters [София: За буквите – О писменехь].

Garvanova, M. (2017b). Values and life satisfaction: Empirical evidence from Bulgaria. *4th International Multidisciplinary Scientific Conference on Social Sciences and Arts SGEM*, *17*(2), 477–484.

Garvanova, M. [Гарванова, М.] (2013). *A transformation in the values of the contemporary Bulgarian* [Трансформация в ценностите на съвременния българин]. Sofia: For Letters [София: За буквите – О писменехь].

Garvanova, M. [Гарванова, М.] (2005). A transformation of values in Bulgaria in the situation of social change [Трансформация на ценностите в България в ситуацията на социална промяна]. *Psychological Research*, *3*, 35–42.

Garvanova, M., Papazova, E. (2017). Values of Bulgarian students in emerging adulthood. *8th International Conference on Education & Educational Psychology*, doi:10.15405/epsbs.2017.10.25

Gaweda, B. (2017). *No country for losers? Gender, (in)equality, and the discursive construction of subjects and values in Polish politics* (PhD Thesis). The University of Edinburgh.

Inglehart, R., & Norris, P. (2017). Trump and the populist authoritarian parties: The silent revolution in reverse. *Perspectives on Politics*, *15*(2), 443–454, https://doi.org/10.1017/S1537592717000111

Ju, S. [鞠实儿] (2020). Generalized argumentation theory and methods [广义论证的理论与方法]. *Studies in Logic* [逻辑学研究], *13*(1), 1–27.

Ju, S., Chen, Z., & Yang, H. (2021). Political argumentation by reciting poems in the spring and autumn period of ancient China. *Argumentation*, *35*, 9–33, http://doi.org/10.1007/s10503-020-09527-x

Krizsan, A., & Roggeband, C. (2021). *Politicizing gender and democracy in the context of the Istanbul convention*. Palgrave, Macmillan.

Krizsan, A., & Roggeband, C. (2020). Democratic backsliding and the backlash against women's rights: Understanding the current challenges for feminist politics. Background paper prepared for the expert group meeting on the 25th anniversary of the Beijing Platform for Action, to inform the 64th session of the Commission on the Status of Women.

Krizsan, A., & Roggeband, C. (2019). *Gendering democratic backsliding in Central and Eastern Europe: A comparative agenda*. Centre for Policy Studies, Central European University.

Leye, E., D'Souza, H., & Meurenus, N. (2021). The added value of and resistance to the Istanbul convention: A comparative study in 27 European member states and Turkey. *Frontiers in Human Dynamics* 3, https://doi.org/10.3389/fhumd.2021.697331

Lönnqvist, J. E., Leikas, S., & Verkasalo, M. (2018). Value change in men and women entering parenthood: New mothers' value priorities shift towards conservation values. *Personality and Individual Differences*, *120*, 47–51, http://dx.doi.org/10.1016/j.paid.2017.08.019

Lönnqvist, J. E., Verkasalo, M., Wichardt, P.C., & Walkowitz, G. (2013). Personal values and prosocial behaviour in strategic interactions: distinguishing value-expressive from value-ambivalent behaviours. *European Journal of Social Psychology*, *43*(6), 554–69, https://doi.org/10.1002/ejsp.1976

Lönnqvist, J. E., Jasinskaja-Lahti, I., & Verkasalo, M. (2011). Personal values before and after migration: a longitudinal case study on value change in Ingrian–Finnish migrants. *Social Psychological and Personality Science*, *2*(6), 584–91, https://doi.org/10.1177/1948550611402362

Macagno, F., & Walton, D. (2018). Practical reasoning arguments: A modular approach. *Argumentation*, *32*, 519–547, http://doi.org/10.1007/s10503-018-9450-5

Maio, G. R. (2010). Mental representations of social values. *Advances in Experimental Social*

Psychology, *42*, 1–43, http:/doi.org/10.1016/S0065-2601(10)42001-8

Milfont, T. L., Milojev, P., & Sibley, C. G. (2016). Values stability and change in adulthood: a 3-year longitudinal study of rank-order stability and mean-level differences. *Personality and Social Psychology Bulletin*, *42*(5), 572–88, https://doi.org/10.1177/0146167216639245

O'Dwyer, E., & Çoymak, A. (2020). Basic human values and their contexts: A multilevel analysis of support for the use of armed drones in the United States, United Kingdom, and Turkey. *Political Psychology*, 41(2), 249–264, https://doi.org/10.1111/pops.12621

Piurko, Y., Schwartz, S., Davidov, E. (2011). Basic personal values and the meaning of left-right political orientations in 20 countries. *Political Psychology*, *32*(4), 537–561, https://doi.org/10.1111/j.1467-9221.2011.00828.x

Roccas, S., Sagiv, L., & Navon, M. (2017). Methodological issues in studying personal values. In S. Roccas, L. Sagiv (Eds.), *Values and behavior: Taking a cross cultural perspective* (pp. 15–50). Springer International Publishing, http://doi.org/10.1007/978-3-319-56352-7_2

Sagiv, L., & Roccas, S. (2021). How do values affect behavior? Let me count the ways. *Personality and Social Psychology Review*, *25*(4), 295–316, https://doi.org/10.1177/10888683211015975

Sagiv, L., Schwartz, S. (2022). Personal values across cultures. *Annual Review of Psychology*, *73*, 517–46, https://doi.org/10.1146/annurev-psych-020821-125100

Schultz, P. W., & Zelezny, L. C. (1998). Values and proenvironmental behavior: A five-country survey. *Journal of Cross-Cultural Psychology*, *29*(4), 540–58.

Schwartz, S. (2017). The refined theory of basic values. In S. Roccas, L. Sagiv (Eds.), *Values and behavior: Taking a cross cultural perspective* (pp. 51–72). Springer International Publishing, https://doi.org/10.1007/978-3-319-56352-7_3

Schwartz, S. (2015). Basic individual values: Sources and consequences. In D. Sander, T. Brosch (Eds.), *Handbook of value* (pp. 63–84). Oxford University Press, https://doi.org/10.1093/acprof:oso/9780198716600.003.0004

Schwartz, S. (1992). Universals in the content and structure of values: theoretical advances and empirical tests in 20 countries. *Advances in Experimental Social Psychology*, *25*, 1–65, https://doi.org/10.1016/S0065-2601(08)60281-6

Schwartz, S., Cieciuch, J., Vecchione, M., Davidov, E., Fischer, R., et al. (2012). Refining the theory of basic individual values. *Journal of Personality and Social Psychology*, *103*(4), 663–88, http://dx.doi.org/10.1037/a0029393

Schwartz, S., & Huismans, S. (1995). Value priorities and religiosity in four Western religions. *Social Psychology Quarterly*, 58(2), 88–107, http://dx.doi.org/10.2307/2787148

Sortheix, F. M., Parker, P. D., Lechner, C. M., & Schwartz, S. (2019). Changes in young Europeans' values during the global financial crisis. *Social Psychological and Personality Science*, *10*(1), 15–25, http://dx.doi.org/10.1177/1948550617732610

Sultovic, L. (2019). The impact of neoconservative activism on gender policies and the women's movement in Croatia. In A. Krizsan, C. Roggeband (Eds.), *Gendering democratic backsliding in Central and Eastern Europe: A comparative agenda* (pp. 31–77). Centre for Policy Studies, Central European University.

Szczygielska, M. (2019). "Good change" and better activism: Feminist responses to backsliding gender policies in Poland. In A. Krizsan, C. Roggeband (Eds.), *Gendering democratic backsliding in Central and Eastern Europe: A comparative agenda* (pp. 120–160). Centre for Policy Studies, Central European University.

Valchev, H. (2022). The cultural embeddedness of arguments raised as a part of the Bulgarian debate about the ratification of the Istanbul convention. *Argumentation*, *36*, 177–202, http://doi.org/10.1007/s10503-021-09560-4

Vecchione, M., Schwartz, S., Alessandri, G., Döring, A. K., Castellani, V., & Caprara, M.G. (2016). Stability and change of basic personal values in early adulthood: an 8-year longitudinal study. *Journal of Research in Personality*, *63*, 111–22, https://doi.org/10.1016/j.jrp.2016.06.002

Vecchione, M., Schwartz, S., Caprara, G. V., Schoen, H., Cieciuch, J., et al. (2015). Personal values and political activism: A cross-national study. *British Journal of Psychology*, *106*(1), 84–106, https://doi.org/10.1111/bjop.12067

Verplanken, B., & Holland, R. W. (2002). Motivated decision making: effects of activation and self-centrality of values on choices and behavior. *Journal of Personality and Social Psychology*, *82*(3), 434–47, https://doi.org/10.1037/0022-3514.82.3.434

Book review

De Brigard, Felipe. *Memory and remembering*. Cambridge University Press, 2023, 79 p.

ISBN: 9781108955447

[Reviewed by ***Kourken Michaelian,***

Centre for Philosophy of Memory,
Institut de Philosophie de Grenoble, Université Grenoble Alpes, France
and
Institut Universitaire de France, France
email: kourken.michaelian@univ-grenoble-alpes.fr
ORCID ID: https://orcid.org/ 0000-0002-5205-3046]

This brief book (part of the Cambridge Elements in Philosophy of Mind series) by a leading philosopher of memory is an important contribution to a field to which, due simply to its youth, few introductory texts have so far been devoted. Indeed, the one introductory book that philosophers of memory have so been able to give to their students is Senor's excellent *Critical Introduction to the Epistemology of Memory* (2019). De Brigard's book, since it approaches memory from the perspective of philosophy of mind rather than that of epistemology, nicely complements Senor's. Indeed, given that both books are relatively brief, they might effectively be assigned together in a single course. Naturally, this book, like Senor's, might also be assigned on its own (in an undergraduate-level course) or in combination with more advanced readings (in a graduate-level course). Having myself used it in the latter manner, I can confirm that the book does an outstanding job of introducing students who have no prior familiarity with the field to the central concepts, debates, and theories in the philosophy of memory, getting them up to speed to an extent sufficient to enable them to navigate the increasingly voluminous and technical professional literature without getting hopelessly lost.

De Brigard begins, in section 1 of the book, by providing a short introduction. In section 2, he focuses on the nature of memory understood as a capacity, its relationship to other cognitive capacities, and the relationships among different kinds of memory. The section begins with an overview of the history of philosophical attempts to distinguish between memory and other capacities in terms of the content of the former. The central problem for this approach, as De Brigard sees it, is that it either fails to include some instances of memory — the content of memory is, for example, not always about past events — or, if different kinds of memory are introduced, corresponding to different kinds of content, leads to an implausible multiplication of kinds of memory. The section then turns to recent attempts to distinguish among kinds of memory in linguistic terms, pointing out that these likewise encounter problems. The view that memory is factive, for example, has been challenged in a variety of ways, and there are reasons to suppose that linguistic considerations do not distinguish cleanly between remembering that and remembering how. Turning next to empirical distinctions, De Brigard reviews the standard taxonomy of memory, which distinguishes between declarative memory, including episodic and semantic memory, on the one hand, and a variety of forms of nondeclarative memory, on the other hand. Though it remains influential, there is accumulating empirical evidence against the standard taxonomy, including the distinction between episodic and semantic memory. After a brief review of attempts to distinguish among kinds of memory in phenomenological terms — by, for example, invoking the notion of autonoetic consciousness — this section of the book concludes with a mixture of pessimism and optimism: De Brigard is pessimistic about our ability to develop a workable taxonomy of memory but optimistic about our ability to make progress by focusing on paradigmatic instances of the kinds of memory in which we are interested even absent such a taxonomy.

Section 3 of the book sets aside the nature of memory understood as a capacity to consider the nature of remembering understood as a process. Readers familiar with the philosophy of memory will expect this section to introduce standard philosophical theories of episodic remembering, including the epistemic, causal, and simulation theories. While the section does indeed cover these theories, it begins, somewhat surprisingly, by introducing the distinction between representationalism and nonrepresentationalism — including direct realism — and goes on to treat the epistemic, causal, and simulation theories as species of representationalism. This is the one point at which the book might arguably have been better organized. Representationalism and nonrepresentationalism, on the one hand, and the epistemic, causal, and simulation theories, on the other hand, answer different questions, the former concerning the direct objects of memory or the nature of the relationship between memories and their objects, the latter concerning what it is to remember, and, while standard versions of the epistemic, causal, and simulation theories do take representationalism for granted, it is not obvious that they are bound to do so. It would thus perhaps have been more useful to clearly distinguish the debate between representationalists and nonrepresentationalists from the debate among epistemic, causal, and simulation theorists. It likewise would have been useful to clearly distinguish these debates from debates over memory traces and memory markers, which the section sometimes runs together. Nevertheless, the section provides a helpful overview of the epistemic, causal, and simulation theories and of the major problems for these, pointing out that the epistemic theory problematically treats memories as being propositional in character, that the causal theory is challenged by empirical evidence pertaining to the constructive character of remembering and to the relationship between episodic remembering and episodic future thinking, and that the simulation theory faces a variety of objections by causalists. The section concludes with a welcome discussion of recent approaches to the function of memory, a topic that has been receiving increasing attention, observing that it is not obvious that the function of memory can straightforwardly be identified with remembering.

Section 4 of the book focuses on the question of what we remember, a question that De Brigard interprets as primarily concerning memory traces. Following a review of debates over the very existence of traces and the need to invoke these in a philosophical theory of remembering, the section provides detailed discussions of more specific debates, including the debate among partisans of different ways of understanding the relationships between traces, perceptual experiences, and memories in light of the constructive character of remembering, the debate between partisans of local and distributed conceptions of traces, and the debate between partisans of explicit and implicit views of the content of traces. Finally, in a subsection that will be of particular interest to researchers, the section articulates a view of memory traces intended to be congenial to simulationism. De Brigard concludes, in section 5, with a brief summary.

Writing a volume of this sort — which must cover a great deal of ground in very little space — is no mean feat, but De Brigard has pulled it off wonderfully, selecting key issues and introducing them to students in a manner that consistently manages to remain accessible while avoiding oversimplification. He has thus contributed a key element to the developing infrastructure of a new field. I have no doubt that I — and many others — will continue to teach from the book regularly for many years to come.

References

Senor, T. D. (2019). *A critical introduction to the epistemology of memory*. Bloomsbury.