

LOGOS

The Cornell Undergraduate Journal of Philosophy

*Layers of Logical Consequence: An examination of the false dichotomy
between proof-theory and model-theory*

Grace Field, University of Toronto

Phenomenology and Conceptualism

Bernardo Portilho Andrade, Middlebury College

*A Description-Based Solution to the Non-Identity Problem in the Context
of Mitigating Historical Injustice*

Brian Wong Yue Shun, Pembroke College, University of Oxford

God's Dream: Implications of Idealistic Theism for Non-Theists

Max DuBoff, Rutgers University

In Defense of the Structural Matching Thesis

Molly Elder, Johns Hopkins University

Volume XIV - Spring 2018

Logos : λόγος

Volume XIV • Spring 2018

Copyright © 2018
Logos: The Undergraduate Journal of Philosophy
orgsync.com/72275/chapter
Cornell University Ithaca, NY

An independent student publication

Logos: The Undergraduate Journal of Philosophy, an independent student organization located at Cornell University, produced and is responsible for the content of this publication. This publication was not reviewed or approved by, nor does it necessarily express or reflect the policies or opinions of, Cornell University or its designated representatives.

Mail:
Logos: The Undergraduate Journal of Philosophy
c/o The Sage School of Philosophy
218 Goldwin Smith Hall
Cornell University
Ithaca, NY 14853

Email:
journal.logos@gmail.com

Logos : λόγος

Volume XIV • Spring 2018

Editor-In-Chief:

Haoxuan Ma

Club President:

Noam Weinreich

Vice President:

Arlinda Shehu

Joshua Nathanson

Treasurer:

Rohin Garg

Secretary:

Tianyang Xu

General Editorial Staff:

Weston Barker

Lorenzo Benitez

Rebecca Ebiana

Ashley Gasdow

David Kahn

Kwesi Kwapong

Madilynn Madigar

Scott Partington

Jae Haeng Rhee

Faculty Advisor:

Harold Hodes

Contents

Editors' Introduction	9
Layers of Logical Consequence: An examination of the false dichotomy between proof-theory and model-theory <i>Grace Field</i> <i>University of Toronto</i>	11
Phenomenology and Conceptualism <i>Bernardo Portilho Andrade</i> <i>Middlebury College</i>	25
A Description-Based Solution to the Non-Identity Problem in the Context of Mitigating Historical Injustice <i>Brian Wong Yue Shun</i> <i>Pembroke College, University of Oxford</i>	45
God's Dream: Implications of Idealistic Theism for Non-Theists <i>Max DuBoff</i> <i>Rutgers University</i>	61
In Defense of the Structural Matching Thesis <i>Molly Elder</i> <i>Johns Hopkins University</i>	75

Editors' Introduction

The staff of *Logos* is proud to present the fourteenth volume of Cornell University's undergraduate journal of philosophy. After carefully considering the submissions we received over the past year we have selected an exemplary set of five articles chosen for their creativity, cogency, and depth of philosophical inquiry.

This year's selection pool was full of quality submissions, and we received inquiries from over seventy undergraduates situated across the English-speaking world. All of the papers contained within this volume were carefully reviewed and selected because of their exceptional quality and varied subjects. The fourteenth volume of *Logos* features papers whose topics fall under the philosophy of logic, phenomenology, personal identity, philosophy of religion and epistemology. We are delighted to be able to publish such a broad set of articles while bringing the best new undergraduate work to public view.

We would like to thank and acknowledge the authors of our chosen submissions: Grace Field for her submission entitled "Layers of Logical Consequence: An examination of the false dichotomy between proof-theory and model-theory", Bernardo Portilho Andrade for his submission entitled "Phenomenology and Conceptualism", Brian Wong Yue Shun for his submission entitled "A Description-Based Solution to the Non-Identity Problem in the Context of Mitigating Historical Injustice", Max DuBoff for his submission entitled "God's Dream: Implications of Idealistic Theism for Non-Theists", and Molly Elder for her submission entitled "In Defense of the Structural Matching Thesis".

We are grateful to the Sage School of Philosophy whose funding supports *Logos* this year. We are deeply indebted to the staff of the Sage School of Philosophy, particularly our advisor Harold Hodes, as well as Pamela Hanna and Dorothy Vanderbilt, for assisting with publication and the day-to-day of running the journal; and to our undergraduate staff without whom none of this would be possible.

Haoxuan Ma
Editor-in-Chief

Layers of Logical
Consequence:
An examination of the
false dichotomy between
proof-theory and
model-theory*

Grace Field
University of Toronto

*This is a previous edition of a paper for which St. Andrews Philosophy Society's *Aporia* retains the original publishing right.

ABSTRACT

Truth, on its surface, is not obviously complex. It is used flippantly in everyday conversation, and we like to think it forms a concrete basis for rational thought. However, what is the essence of truth? This question has been hotly debated for decades by some of the brightest minds in analytic philosophy. Most fight for one of two sides – proof-theory or model-theory. Proof-theory characterizes truth in terms of logical implication. Namely, according to proof-theory, the statement “A implies B” is true if there exists a proof from A to B. In contrast, model-theory characterizes truth based on possible states of the world. By model theory, “A implies B” is true if it holds for every possible A and every possible B. The theories have always been approached as two mutually exclusive alternatives. In *Layers of Logical Consequence*, I argue that we should get rid of this dichotomous attitude. If we separate the metaphysical nature of truth from its epistemic nature, each theory falls into place as a partial description of truth. Namely, truth can be characterized as epistemically model-theoretic, and metaphysically proof-theoretic.

I. INTRODUCTION

There exists a long-standing rivalry between model-theoretic and proof-theoretic versions of logical consequence. A model-theoretic view frames logical consequence in terms of truth preservation across cases. A proof-theoretic view holds that consequence reduces to the existence of a formally valid argument between the premises and conclusion. There are clear arguments for both – and it is difficult to refute either on purely technical grounds. In light of this conflict, what is logical consequence? I believe our primary concern here is not over which view is correct, in a technical sense. Solid arguments can be made for either side of that debate, as we have seen in the extensive literature on the subject. We aim to identify the conception of logical consequence that most effectively describes its full nature. What is the nature of valid inference between a set of premises and a conclusion? Pure model-theory and the pure proof-theory have persevered in parallel for decades. Neither is entirely counter-intuitive and in fact many philosophers have suggested some combination of the two. It even seems reasonable to argue that we should be searching for a compromise: both views have benefits, and neither seems technically wrong. In this paper, I will identify the relevant compromise with a distinction between epistemic and metaphysical nature. My aim is to defend logical consequence as epistemically model-theoretic, and metaphysically proof-theoretic.

II. BACKGROUND

My view fits into a diverse landscape of existing compromise-style theory – many contemporary authors do not commit themselves to pure proof-theory or pure model-theory. Logical pluralists hold that there are several correct notions of logical consequence. Beall & Restall, Cook & Shapiro and Carnap have each defended a unique version of the view.¹ Relativists have suggested that the correct notion might instead be domain-relative – variable but determined by the field we are working in. Some versions of relativism even identify logical domain with culture, resulting in culture-dependent logic.²

Kreisel’s squeezing argument for logic makes an unorthodox compromise – and of existing views on logical consequence, his is most similar to the view I wish to defend. A squeezing argument exploits the following reasoning: a concept can be characterized by a set of necessary properties, and a different set of sufficient properties, if there exists a separate relationship to show that the necessary properties imply the sufficient properties.³ Kreisel claims to successfully run a squeezing argument on logic, where our informal concept of logical validity can be squeezed between the existence of proof, as a sufficient condition, and the absence of a counter-model, as a necessary condition.⁴ Here logic is squeezed between proof-theory and model-theory.

Hartry Field has suggested that a squeezing argument applied to logic can fully describe the everyday notion of logical consequence, by reconciling our sympathy for proof with our reliance on a case-by-case method.⁵ But, as Peter Smith notes, this statement goes too far. A squeezing theory is not enough if our goal is to describe real, intuitive logical consequence.

It would force us to squeeze intuition with proof-theory, and then squeeze proof-theory with model-theory, placing far too many formal restrictions on our intuition in the process.⁶ Instead, I propose we should admit that our intuition has two sides – one epistemic, and one metaphysical.

Epistemology and metaphysics are two central and distinct areas of inquiry, but there is no universally accepted definition for either. In general,

¹ J. C. Beall and G. Restall, “Logical Pluralism,” *Australasian Journal of Philosophy* 78, no. 4 (2000): 475-493; R. T. Cook, “Vagueness and Mathematical Precision,” *Mind* 111, no. 442 (2002): 225-247; S. Shapiro, *Vagueness in Context* (Oxford: Oxford University Press, 2006); R. Carnap, *The Logical Syntax of Language* (Paterson: Littlefield, Adams, & Co., 1959).

² G. C. Lokhorst, “The Logic of Logical Relativism,” *Logique & analyse* 41, no. 161-162-163 (1998): 57-65; C. Greiffenhagen and W. Sharrock, “Logical Relativism: Logic, Grammar, and Arithmetic in Cultural Comparison,” *Configurations* 14, no. 3 (2006): 275-301.

³ P. Smith, “Squeezing arguments,” *Analysis* 71, no. 1 (2011): 23

⁴ Smith, “Squeezing arguments,” 23-24.

⁵ H. Field, *Saving Truth from Paradox* (Oxford: Oxford University Press, 2008), 47-48; Smith, “Squeezing arguments,” 26.

⁶ Smith, “Squeezing arguments,” 26-27.

epistemology is the study of knowledge; metaphysics is the study of fundamental being. To address the epistemic and metaphysical nature of logical consequence, we must begin with a clear definition for each type of property. For my theory of logical consequence, a property is epistemic if it concerns knowledge or justified belief. Metaphysics is slightly more abstract by nature, and this is sure to be reflected in any attempt at a definition. I will stick to the following admittedly elusive, but meaningful notion: a property is metaphysical if it concerns the most abstract essence of a concept.

III. EPISTEMIC NATURE VS. METAPHYSICAL NATURE: MOTIVATION AND DISTINCTION

My view is motivated by a strong belief that the full nature of a concept cannot be reduced to its epistemic properties. Neither can we claim to describe the nature of a concept just by describing its metaphysical essence. Epistemic properties determine experiential understanding, and metaphysical properties determine abstract essence. Therefore a complete description – which is what we aim for in any discussion of a concept’s nature – must include epistemic and metaphysical sub-descriptions. These are distinct sketches, to characterize the two distinct sides to any concept. When taken together, they blend to create a full image. Later I will defend against the sceptic who argues that the metaphysical sketch is meaningless or unnecessary. For now, I assume my anti-sceptic intuition on this issue and use it to motivate my view – any description that does not include both epistemic and metaphysical qualities must be incomplete.

On the epistemic side, logical consequence seems model-theoretic. Our epistemic access to logical consequence is defined by the limits on our actual inference process. And I propose that our actual inference process follows a representationalist model-theoretic program, where models represent possible states of reality.⁷ I make this proposal in light of the intuitive dependency between experience and epistemic access to the world. Carnap notably supports such dependency, in his comprehensive work *The Logical Structure of the World*. In his words, “I can make an ‘epistemic evaluation’ of any experience I have had by stating to what extent this experience has added to my (theoretical) knowledge. This addition consists not only of the theoretical content of the experience itself but also of whatever I can infer from this content with the aid of my earlier knowledge.”⁸ I consider his statement to be intuitively and practically accurate. We come to know, or believe, through assessing the nature

⁷ G. Y. Sher, “Did Tarski Commit ‘Tarski’s Fallacy?,” *The Journal of Symbolic Logic* 61, no. 2 (1996): 658-661.

⁸ R. Carnap, *The Logical Structure of the World* (Chicago: Open Court Publishing Company, 2003), 309.

of the world based on real experience, or imaginary but believable extension of experience. And an experience is a snapshot of a possible state-of-the-world. In this sense, the epistemic side to any inference seems to be based on case-by-case analysis of historically possible and hypothetically possible states of reality.

Logical consequence might still be metaphysically proof-theoretic. In fact, we naturally see cases at the bottom of a hierarchy dominated by coherence with the abstract law, following a point made by Prawitz.⁹ In this hierarchy case-by-case relationships do not define the abstract nature of logical consequence; instead they derive from it. To reject meaningful argument-based abstract nature would be to turn the hierarchy upside down. Our intuitive understanding of consequence, therefore, hints at the proof-theoretic metaphysical essence. This hint will be essential in defending against a sceptic who questions the existence or relevance of metaphysical essence for logic.

IV. MODEL-THEORETIC EPISTEMIC NATURE & PROOF-THEORETIC METAPHYSICAL NATURE

To offer any plausibility to my position, I must be able to show it is at least possible for consequence to fit the strict technical requirements of model-theoretic dogma in its practical use, and fit the equivalent requirements of proof-theory in its fundamental nature. As Beall & Restall note, model-theory requires reliance on cases and counter-cases: “The model-centred approach to logical consequence takes the validity of an argument to be absence of counterexample.”¹⁰ In technical terms, a consequence is model-theoretic if it is effectively described by the following equivalence relation: B is a consequence of A iff for any model m , if m satisfies A then m satisfies B.¹¹ Proof-theory instead relies on the existence of argument-based reasoning. Again appealing to Beall & Restall, “On the proof-centred approach to logical consequence, the validity of an argument amounts to there being a proof of the conclusions from the premises.”¹² Technically, a consequence is proof-theoretic on the condition that an argument from A to B is valid iff there is a proof of B from A. Clearly these definitions establish a set of minimal conditions for my theory of logical consequence. At the very least, I must be able to identify the epistemic side of logical consequence with a set of possible models and a corresponding satisfaction relation. And I must be able to identify its metaphysical nature with existence, or absence, of proof.

⁹ D. Prawitz, “Logical Consequence: A Constructivist View,” in *The Oxford Handbook of Philosophy of Mathematics and Logic*, ed. S. Shapiro (Oxford: Oxford University Press, 2005), 671-695.

¹⁰ J. C. Beall and G. Restall, “Logical Consequence,” *Stanford Encyclopedia of Philosophy*, published 2013, accessed May 3, 2017, <https://plato.stanford.edu/entries/logical-consequence/>.

¹¹ A. Tarski, “On the concept of logical consequence,” in *Logic, semantics, metamathematics: papers from 1923 to 1938*, (Indianapolis: Hackett Publishing Company, 1983): 417.

¹² Beall and Restall, “Logical Consequence,” published 2013.

How does this work for the epistemic side to logical consequence? What are the models, and what is the satisfaction relation? As Varzi notes, logic aims to act as a universal requirement for theoretical reasoning. As a result, its models must be universally applicable in some sense; “logic is a uniquely ambitious theory... It aims to be the theory included in every other theory... its models want to include the models of every other.”¹³ I propose we can access this universality by appealing to Carnap’s distinction between individual and general concepts. An individual concept is specific to a certain space-time coordinate, while the corresponding general concept has many possible space-time representations.¹⁴ Carnap’s pet dog Luchs exists at a given location at a given instant. This individual concept is one among many possible representations of the corresponding concept dog.¹⁵ The sky looks blue at 1 pm on a Tuesday afternoon from my window; the sky’s blueness at that point in space and time can be seen as an individual concept representing the general concept blue. For a discrete state that does not explicitly depend on space or time, we must add to Carnap’s conception. The individual concept in such cases cannot be the discrete state itself. Instead, we may identify individual concept with our acknowledgement or experience of that discrete state. Consider the number two, for example. Two does not have an explicit space-time coordinate – but still, it intuitively represents the general concept number. Here our experiential acknowledgement of the concept two at a given place in a given instant in time can be seen as the relevant individual concept. The individual-general concept distinction is universally applicable if it includes this adjustment for discrete states without inherent space-time coordinates.

We may adopt this distinction as a backbone for the epistemic nature of logical consequence, partly in virtue of its universality. I have claimed that epistemically, logical consequence is based on satisfaction of models. I propose we should make the following identifications: the model with the above-described notion of individual concept, satisfaction with representation, and the subject of satisfaction with the above-described notion of the general concept. Then my model-theoretic thesis amounts to the claim that we observe individual concepts to come to know and/or believe consequence relationships between general concepts. In technical terms, a consequence relationship from A to B is epistemically valid on this view if and only if for all individual concepts m, if m is a representation of the general concept A, then m is a representation of the general concept B.

The minimal requirement for proof-theoretic metaphysical nature is more straightforward to satisfy. We merely need to show it is possible to think of

¹³ A. C. Varzi, “On Logical Relativity,” *Philosophical Issues* 12, Realism and Relativism (2002): 199.

¹⁴ Carnap, *The Logical Syntax of Language*, 247-248.

¹⁵ Carnap, *The Logical Syntax of Language*, 247-248.

the abstract nature of logical consequence in terms of a proof from premises to a conclusion. I believe this to be more or less self-evident. We can imagine the existence of an argument from a set of premises to the corresponding conclusion, in the same way we can imagine the existence of an abstract mathematical proof for $x + x = 2x$. The difficulty for the proof-theoretic side of my view is not whether we can possibly imagine a metaphysical relationship between proof and consequence. The difficulty will be to establish whether we can plausibly assert the existence of that relationship.

A satisfactory proof must be both rigorous and formal. Beyond that basic requirement, I do not claim to subscribe to any particular type of proof when I suggest proof-theoretic metaphysical essence. Any discussion of that kind raises myriad problems of its own; and if I were to choose one type of proof, my entire proposal for logical consequence would become opaque for anyone unsympathetic to my choice. Instead, I propose a view in which logical consequence is epistemically model-theoretic, in the sense of individual-general concept modelling outlined above, and metaphysically proof-theoretic, according to some rigorous and formal type of proof.

V. SCEPTICAL OBJECTION & ANTI-SCEPTIC RESPONSE

I admitted earlier that my view rests on an anti-sceptic intuition about the necessity and relevance of both epistemic and metaphysical nature, if our goal is to provide a full description of logical consequence. This intuition is open to challenge; even if you grant my view as it has been developed so far, you may object to this crucial assumption at its basis. Here we encounter a central challenge to my theory – a challenge tied up with the ancient debate on realism versus antirealism towards the existence of abstract universals. If we come to know and believe using a certain notion of a concept, why should we think there exists anything more abstract or fundamental to that concept's full nature? A sceptic might claim: logical consequence is metaphysically thin. I will argue that such scepticism is unwarranted. We do have reason to believe in a non-trivial metaphysical essence for logical consequence.

We will be in a good position to identify gaps in the sceptic's reasoning if we first examine the source of their inspiration. Hume's treatise on causation and induction forcefully encourages doubt about the existence of abstract metaphysical essence for the causal law. In *An Enquiry concerning Human Understanding*, he argues that we are fundamentally limited in our ability to identify causal relationships a priori, since all knowledge stems from a case-by-case examination of experience.¹⁶ This position has since been widely

¹⁶ D. Hume, *An Enquiry Concerning Human Understanding* (Oxford: Oxford University Press, 2008), 18-23.

acknowledged by scientists and philosophers alike.¹⁷ Hume then settles on a description of cause-and-effect as mere cosmic regularity, questioning the notion of cause-and-effect as a manifestation of abstract causal law.¹⁸ Causation is not equivalent to logical consequence, but I believe we can construct an analogy here. Epistemic reliance on the case-by-case method in science has led to widespread scepticism about the existence of any abstract scientific law beyond functional dependency. It seems that similar doubt could apply to the existence of metaphysical essence for logical consequence – assuming its epistemic nature is model-theoretic, and therefore reliant on a case-by-case analysis.

I could respond by restricting the breadth of logical consequence. Limits on a priori reasoning inspire Hume's scepticism about the abstract law. But these limits plausibly do not apply when working in purely abstract domains – in mathematics or classical symbolic logic, for example. I could respond to the sceptic's concerns by restricting logical consequence to these domains. However, this response would be too weak, since the notion of logical consequence extends beyond strict mathematical manipulation. I believe that even its everyday manifestation is genuine, and I aim to provide a theory to describe this full sense of the concept. As Beall and Restall write, "not all logic is a matter of form".¹⁹

Furthermore, the problem extends beyond strictly Humean considerations. Scepticism about the existence of abstract nature, beyond what we are able to directly know and believe, can be traced back at least as far as the seventeenth century. Historian of science Steven Shapin writes, "Throughout the seventeenth century there were influential voices sceptical of the legitimacy of mathematical 'idealizations' in the explication of physical nature as it actually was."²⁰ This scepticism has persevered in the philosophy of science, and in philosophy as a whole.

Any modern philosopher sympathetic to anti-realist views on the existence of universals might be inclined to adopt the sceptical position. Conceptualists, including Poincaré, Brouwer, and Weyl, see universals as mere human constructions.²¹ A formalist goes even farther – according to Hilbert, for example, a universal is simply a useful matter of notation.²² Your position on the spectrum of this debate may have technical ramifications for your use of universals in formal logic. Regardless, it is a strong indicator of your sympathy

¹⁷ Carnap, *The Logical Structure of the World*, 265

¹⁸ D. M. Armstrong, "A World of States of Affairs," *Philosophical Perspectives* 7, Language and Logic (1993): 438; D. M. Armstrong, *What is a Law of Nature?* (Cambridge: Cambridge University Press, 1983), 4; Carnap, *The Logical Structure of the World*, 264.

¹⁹ Beall and Restall, "Logical Pluralism," 480.

²⁰ S. Shapin, *The Scientific Revolution* (London: The University of Chicago Press, Ltd., 1998), 59.

²¹ W. V. Quine, "On What There Is," *The Review of Metaphysics* 2, no. 5 (1948): 33.

²² Quine, "On What There Is," 34.

towards the weight of abstract metaphysical concepts in general. If you are an anti-realist for universals, you will likely sympathize with a view on which logical consequence is metaphysically thin. Here we find a further source of scepticism. And the weak response outlined above would fail on this front, even if I were willing to constrain logical consequence to abstract domains.

A strong response must head in a different direction, from a different starting point. We must admit that, by analogy to Hume's work on causation, fully model-theoretic epistemic nature implies the existence of epistemic boundaries on a priori reasoning. In other words, a model-theoretic epistemic nature restricts our epistemic access to any proof-based abstract nature that might exist. But to submit to the sceptic would be to extend Humean and anti-realist ideals too far. To be unable to fully know about an entity is not to know that it does not exist. We can appeal to work in logic-as-modelling to begin our defence. But perhaps more importantly, we must acknowledge the distinction between epistemic and metaphysical nature. From this distinction, we are able to argue that a case-by-case epistemic method is not inconsistent with the non-trivial dominance of abstract proof for metaphysical nature. Combined with our intuition on the authority of logical consequence as a fundamental concept, I argue that such dominance is likely, even if it is not obvious in our epistemic approach to inference.

Having conceded fundamental limits on practical a priori reasoning given model-theoretic epistemic nature, we can begin to redeem my theory by examining the effects of these limits. I have emphasized that, if epistemic and metaphysical nature do both carry weight, they are not equivalent properties. One useful way to characterize this difference allows us to draw from Shapiro's work on logic-as-modelling. Namely, we can easily view the epistemic nature of logical consequence as a model. Here, I mean model in the colloquial sense of approximate representation – I am not referring to the individual-general concept introduced earlier for my treatment of model-theory. In general epistemic nature is defined by the boundaries of knowledge; therefore, it is in some sense a knowledge-based approximation of a concept that might have deeper metaphysical qualities. Key here is the approximate in approximate representation. A model is not necessarily a complete representation of the concept it stands for; "there is almost always a gap between a model and what it is a model of."²³ This point helps Shapiro to develop his logic-as-modelling view, but I use the same point for a different purpose.

If the epistemic nature of logical consequence is model-theoretic, it is particularly limited – as the sceptic notes – in its access to a priori reasoning.

²³ Shapiro, *Vagueness in Context*, 50.

Here I can use the sceptic's point for my own benefit. If the epistemic nature of logical consequence is a fundamentally limited description, it seems to be a model by definition – in the same way that “a collection of point masses is a model of a system of physical objects, and the Bohr construction is a model of an atom”.²⁴ And if it is a model, it is most likely an approximate representation. But any approximate representation must represent some non-trivial exact concept. I suggest that this exact concept is precisely the metaphysical nature of logical consequence. Epistemic nature covers everything we are able to know about logical consequence relationships. And everything we are able to know creates a model of the non-trivial abstract metaphysical reality, which is outside the realm of our direct knowledge. Therefore the sceptic's emphasis on epistemic constraint offers indirect support for my defence of non-trivial metaphysical nature.

Furthermore, and crucially, there is no fundamental inconsistency in a view that includes non-trivial essence from both sides, epistemic and metaphysical. Here we are encouraged by various influential philosophers, who subscribe to Hume's work on a scientific method yet maintain realism on the existence of the universal law. Again we work with an analogy, comparing the Humean notion of scientific method with epistemic model-theoretic essence, and the existence of fundamental causal law with the existence of non-trivial metaphysical essence.

From the early 20th-century, Carnap and Russell are among many explicit advocates for Hume's view on the epistemic limits of a priori reasoning. Still, they do not deny the existence of the fundamental law, as metaphysical nature. See the following excerpt from Carnap on the metaphysical essence of causal correlation: “Here we do not simply ask between what object the relation obtains, but what it is between the correlated objects, by virtue of which they are connected.”²⁵ And, even more explicitly, “the essence problems belong to metaphysics.”²⁶ Still he sees our epistemic experience of specific causal relationships as a manifestation of mere functional dependency.²⁷ Russell, whose position on epistemic empiricism is steadfast, still does not rule out the existence of causal law as an abstract mathematical formulation.²⁸ He writes, “there is no a priori category of causality, but merely certain observed uniformities.”²⁹ However, “In all science we have to distinguish two sorts of laws: first, those that are empirically verifiable but probably only approximate; secondly, those that

²⁴ Shapiro, *Vagueness in Context*, 49.

²⁵ Carnap, *The Logical Structure of the World*, 35.

²⁶ Carnap, *The Logical Structure of the World*, 35.

²⁷ Carnap, *The Logical Structure of the World*, 264.

²⁸ B. Russell, “On the Notion of Cause,” *Proceedings of the Aristotelian Society* 13 (1912-13): 14.

²⁹ Russell, “On the Notion of Cause,” 24.

are not verifiable, but may be exact.”³⁰ He clearly acknowledges the possibility of exact abstract law, in spite of its inherently unverifiable nature.

From the more modern tradition, Armstrong and Davidson hold a similar position. Both endorse Humean views on the method but accept the abstract law.³¹ In Armstrong’s words, “It is true that there appears to be no a priori argument that takes one from singular causation to law.”³² However, “It may be noted that the unity of the space-time world is not constituted by the mere conjunction of the state of affairs... The real unity is given by the fact that all the particulars are directly or recursively linked to each other by real, that is external, relations.”³³ Davidson emphasizes that we are often limited in our epistemic access to external causal law – but “very often, I think, our justification for accepting a singular causal statement is that we have reason to believe an appropriate causal law exists, though we do not know what it is.”³⁴

Armstrong presents an explicit argument for his subscription to the separate existence of abstract law as a metaphysical basis for correlation. We either see correlation as an instantiation of a universal and abstract entity, or instantiation of a mere regularity.³⁵ The abstract entity in the former view acts as a basis for explanation.³⁶ Mere regularity does not – any attempt at explanation would be circular, explaining correlation in terms of correlation.³⁷ Armstrong believes the former view is inherently more desirable, simply because a genuine explanation is desirable.³⁸

For logical consequence we have further reason to prefer the former view, even if we do not share Armstrong’s belief in the intrinsic appeal of explanation. Genuine explanation implies a degree of modal transparency. A genuine explanation can describe why a particular consequence relation must, or should hold. Thus by maintaining an anti-sceptic position on the existence of abstract metaphysical nature for logical consequence, we automatically ease modal problems associated with a purely model-theoretic view. These problems have been examined at length by Prawitz.³⁹

We have already identified substantial grounds to believe in non-triviality of the metaphysical essence, both in general and for logical consequence

³⁰ Russell, “On the Notion of Cause,” 16.

³¹ Armstrong, “A World of States of Affairs,” 438;

D. Davidson, “Causal Relations,” *The Journal of Philosophy* 64, no. 21 (1967): 701-702.

³² Armstrong, “A World of States of Affairs,” 438.

³³ Armstrong, “A World of States of Affairs,” 435

³⁴ Davidson, “Causal Relations,” 701.

³⁵ Armstrong, *What is a Law of Nature?*, 40-41.

³⁶ Armstrong, *What is a Law of Nature?*, 40-41.

³⁷ Armstrong, *What is a Law of Nature?*, 40-41.

³⁸ Armstrong, *What is a Law of Nature?*, 40-41.

³⁹ Prawitz, “Logical Consequence: A Constructivist View”.

in particular. There exists an even more fundamental basis for this belief, in an intuition raised by Prawitz. As he points out, the purely model-theoretic consequence would be intuitively back-to-front. On such a pure view, “we cannot really say that we infer the truth of the conclusion by the use of a valid inference. It is, rather, the other way around: we can conclude that the inference is valid after having established for all inferences of the same form that the conclusion is true in all cases where the premises are.”⁴⁰ I believe this is a crucial and telling point to acknowledge. Intuitively, we want some part of logical consequence to be a fundamental property of relationships between abstract ideas. We lose this if the validity of an inference depends purely on running through all possible interpretations or states of the world. In some sense we are demoting logical consequence, from fundamental to dependent - in particular, to dependent on possible discrete states of the world. On my view, the logical consequence is instead metaphysically dependent on rules for argument and proof, from which possible discrete states are derived. Here we see a key benefit of my view, for solving Prawitz’s intuitive problem with model-theory: namely, the ability to push logical consequence back to its intuitive status, without denying Humean limits on our epistemic access to a priori reasoning.

VI. CONCLUSIONS

It may seem as if the essence of logical consequence is model-theoretic if we subscribe to a model-theoretic program when we come to know or believe inferences. But it is possible to see logical consequence as metaphysically proof-theoretic even if we endorse a model-theoretic program for its epistemic nature. We can examine the potential for existence or dominance of fundamental a priori causal law as a basis for analogy. We have seen that on this issue, the historic rivalry between empirical vs. a priori views is not, in fact, a strict rivalry. Limits on our epistemic access to abstract nature might apply, given a model-theoretic epistemic nature. But if such limits do exist, they seem to imply an approximate epistemic representation of some non-trivial exact metaphysical entity. Non-trivial abstract nature might still exist and metaphysically dominate. And intuitively we understand logical consequence as a fundamental relation, from which cases should derive, not follow. I propose we can adopt a hybrid view, one that will satisfy each of these appeals to analogy and intuition. By allowing for a distinction between epistemic and metaphysical nature, we can characterize logical consequence as epistemically model-theoretic but metaphysically proof-theoretic.

⁴⁰ Prawitz, “Logical Consequence: A Constructivist View,” 675

BIBLIOGRAPHY

- Armstrong, D. M. "A World of States of Affairs." *Philosophical Perspectives* 7, Language and Logic (1993): 429-440.
- Armstrong, D. M. *What is a Law of Nature?* Cambridge: Cambridge University Press, 1983.
- Beall, J. C. and Restall, G. "Logical Consequence." *Stanford Encyclopaedia of Philosophy*. Published 2013, accessed May 3, 2017. <https://plato.stanford.edu/entries/logical-consequence/>.
- Beall, J. C. and Restall, G. "Logical Pluralism." *Australasian Journal of Philosophy* 78, no. 4 (2000): 475-493.
- Carnap, R. *The Logical Structure of the World*. Chicago: Open Court Publishing Company, 2003.
- Carnap, R. *The Logical Syntax of Language*. Paterson: Littlefield, Adams, & Co., 1959.
- Cook, R. T. "Vagueness and Mathematical Precision." *Mind* 111, no. 442 (2002): 225-247.
- Davidson, D. "Causal Relations." *The Journal of Philosophy* 64, no. 21 (1967): 691-703.
- Field, H. *Saving Truth from Paradox*. Oxford: Oxford University Press, 2008.
- Greiffenhagen, C. and Sharrock, W. "Logical Relativism: Logic, Grammar, and Arithmetic in Cultural Comparison." *Configurations* 14, no. 3 (2006): 275-301.
- Hume, D. *An Enquiry Concerning Human Understanding*. Oxford: Oxford University Press, 2008.
- Lokhorst, G. C. "The Logic of Logical Relativism." *Logique & analyse* 41, no. 161-162-163 (1998): 57-65.
- Prawitz, D. "Logical Consequence: A Constructivist View." In *The Oxford Handbook of Philosophy of Mathematics and Logic*, edited by S. Shapiro. Oxford: Oxford University Press, 2005.
- Quine, W. V. "On What There Is." *The Review of Metaphysics* 2, no. 5 (1948): 21-38.
- Russell, B. "On the Notion of Cause." *Proceedings of the Aristotelian Society* 13 (1912-13): 1 - 26.
- Shapin, S. *The Scientific Revolution*. London: The University of Chicago Press, Ltd., 1998.

Shapiro, S. *Vagueness in Context*. Oxford: Oxford University Press, 2006.

Sher, G. Y. "Did Tarski Commit 'Tarski's Fallacy?'" *The Journal of Symbolic Logic* 61, no. 2 (1996): 653-686.

Smith, P. "Squeezing arguments." *Analysis* 71, no. 1 (2011): 22-30.

Tarski, A. "On the concept of logical consequence." In *Logic, semantics, metamathematics: papers from 1923 to 1938*. Indianapolis: Hackett Publishing Company, 1983.

Varzi, A. C. "On Logical Relativity." *Philosophical Issues* 12, Realism and Relativism (2002): 197 – 219

Phenomenology & Conceptualism

Bernardo Portilho Andrade
Middlebury College

Remember the time when, standing in front of a Pollock drip painting, you entered a formless world devoid of recognizable objects. Recollect the struggle to find a word for the taste of a fruit, or for the smell of your childhood home. Think of how impossible it was to explain the flex of serving the tennis ball just right; or how mysterious it felt when you realized, from a swelling of your heart, that love had arrived. Examples like these might tempt one to posit a nonconceptual content of experience; such experiences are either too vague or too sacred to be leveled to the conceptual realm of everyday life and practical decisions. Several contemporary phenomenologists, whose chief concern lies in describing complex experiences like these, have rejected doctrines that demand the conceptual investment of all experience. In response to them, I contend that there is no need for this resistance towards the conceptual. John McDowell, for whom the content of experience is all conceptual, can coherently accept their phenomenological descriptions; and these descriptions, when appropriately corrected by his conceptualist claim, can enrich his position by providing vivid accounts of the experience he shows to be conceptual.

I begin this paper by presenting McDowell's argument, in *Mind and World* (1996), against non-conceptual content (Part I). I then consider and refute two phenomenological critiques of McDowell, both of which take issue with his conceptualist doctrine on grounds that it does not account for certain phenomena related to the mind's interaction with the world. First, I respond to Hubert Dreyfus's critique that conceptualism cannot do justice to instances of acting "in flow," such as absorbed coping and situated normativity (Part II). Second, I reply to Walter Hopp's argument that McDowell's conceptualism blurs the distinction between perception and thought (Part III). In response to Dreyfus and Hopp, I argue that the phenomena they describe must be called "conceptual" in McDowell's sense, but that their descriptions can contribute to McDowell's picture by offering a detailed taxonomy of the varieties of experience.

I.

McDowell aims to clarify the relation between mind and world in light of a central epistemological concern: the *justificatory relation* between perceptual experience and the beliefs we have about the world. McDowell's principal concern is to account for how experience can serve as a source of knowledge. The problem at stake traces back

to Early Modern philosophy: it can only be solved by reconsidering the fundamental relation between thought and perception. The idea of conceptual capacities operative in receptive experience is supposed to do just that.

According to McDowell, epistemology has suffered for centuries from a certain dilemma regarding how a belief can be justified (McDowell, 1996, 5). The dilemma concerns the different roles philosophers have assigned to sensory experience. On an empirical foundationalist account, sensations are conveniently two-legged: while informing our conceptual thoughts, they also offer a foothold in a non-conceptual, external reality. Granting sensations this double role would putatively allow experience to warrant beliefs. To this extent, experience would belong to what Sellars calls “the logical space of reasons,” the space of justifying and being able to justify what one says (Sellars, 2003, 131). At the same time, however, sensations would also partake in the physical-causal processes of the “space of nature,” the space of non-epistemic facts (Sellars, 2003, 131). Along with Evans (1982) and Davidson (2009, xvi), McDowell follows Sellars in deeming this double-role of sensations unacceptable. To avoid this picture, Davidson has opted for a coherentist account, according to which beliefs can be justified by other beliefs only. For McDowell, however, this alternative does not fare much better than empirical foundationalism; it suggests imageries of confinement within our thoughts, denying our rational faculties access to the empirical world (McDowell, 1996, 15). McDowell intends to break out of the perpetual oscillation between the “Myth of the Given,” which grants sensations a dubious double role, and Davidsonian coherentism, unconstrained by the external world. I will elaborate on the two horns of this dilemma in some detail.

In the empiricist picture sketched above, the justification for our beliefs would supposedly come from an encounter with a bare, extra-conceptual presence in the external world. McDowell claims that this non-conceptual Given cannot serve as a rational justification for judgment, since we cannot “understand the relations in virtue of which a judgment is warranted except as relations within the space of concepts” (McDowell, 1996, 7). By relying on a nonconceptual Given, we would “extend the scope of justificatory relations outside the conceptual sphere,” so that the reasons for our conceptual worldview would lie beyond the reach of our conceptual capacities (McDowell, 1996, 7). McDowell points out the problem with this picture: it leaves us, at best, with a brute and non-conceptualized impact from the external world as the explanation for how we are caused to believe what we believe. A

belief, however, may be *caused* by a natural event, but it is not thereby *justified* by it. Based on claims about the causal origin of our beliefs, McDowell says, we can only provide *exculpations* where we want *justifications* (McDowell, 1996, 8), i.e., we can only dispossess ourselves of epistemological responsibilities for holding the beliefs we hold – as if we acquired beliefs in the same way that we might stub our toe against a piece of furniture. To ensure that the justification for our judgments is rational – or what amounts to the same, to ensure that it *is* a justification – McDowell insists that conceptual capacities must be involved from the start, when world first meets mind.

To avoid the Given, we must therefore grant that whatever belongs to the space of nature cannot rationally justify a belief in the space of reasons. This, however, brings us to the other horn of the dilemma. For it may appear that, if we drop the Given, we must opt for a coherentist account, according to which beliefs can be justified by other beliefs only. Davidson says that “nothing can count as a reason for holding a belief except another belief” (Davidson, 2009, 141), and by that he means (as McDowell reads him) that experience cannot count as a reason for holding a belief. For Davidson, since “sensations” are not propositional attitudes, the “relation between a sensation and a belief cannot be logical” (Davidson, 2009, 143). Rather, he thinks that the relation between sensations and beliefs must be *causal*. McDowell points out the undesirable consequences of this view when he explains that the world must exert a *rational* influence on thinking if our conceptual worldview is “to pass a scrutiny of its rational credentials” (McDowell, 1996, 31). Otherwise, we have no *rational* constraint from outside the sphere of thought, and might then worry that “we have no convincing way to credit ourselves with empirical knowledge” (McDowell, 1996, 15). By denying any rational influence of the world on our beliefs, coherentism falls prey to skepticism about empirical reality; it threatens to depict our thinking as “a frictionless spinning in a void” (McDowell, 1996, 42). To escape this problem, we must posit that our thinking is subject to a *rational* (not merely causal) constraint from the outside, in which case there must be rational relations between empirical judgments and the experience that *rationally* grounds them. This entails that the space of reasons cannot be sealed within the boundary of our thoughts, across which rational relations with the external world are supposed to hold (McDowell, 1996, 52). Rather, it must extend all the way to include our experience of the world, if the world is to exert a rational constraint on our thoughts, and if we are to credit ourselves with empirical knowledge.

I have explained the dilemma to which McDowell responds:

either we commit to the naturalistic fallacy that non-rational deliverances of the senses can enter into rational justificatory relations with beliefs (empirical foundationalism), or we commit to the equally fallacious view that thoughts enter into rational relations only with other thoughts, thereby lacking any anchoring in the world (Davidsonian coherentism). To circumvent the perpetual oscillation between these two problematic positions, McDowell suggests that we regard experience as *already* invested by the relevant conceptual capacities. For him, it is only by conceiving of the contents of experience as already conceptual that we can understand how perception gives us reasons or justifications to believe that something is or is not the case (McDowell, 1996, 10). This way, we can maintain thought's bearing on reality *without* committing to the Given. McDowell finds in Kant an idea of intuition (i.e., experience) which preserves our senses' contact with the external world, while simultaneously describing the immediate contents of intuition as conceptual, thus avoiding the Given. The external world, in this case, is not one located outside the realm of the conceptual (McDowell, 1996, 54). In a way, for McDowell, reality *itself* is invested with concepts – but it is nonetheless an external reality, with which we are in touch through the Kantian faculty of sensibility. McDowell thus follows Kant's view on the cooperation between two faculties: the faculty of sensibility or receptivity (responsible for the objective purport of experience) and the faculty of spontaneity (responsible for its conceptual form and thereby for its liability to enter into rational relations). For McDowell, the conceptual capacities of spontaneity are drawn on in receptivity, and hence conceptual capacities are operative in the sensations that 'make up' our perceptual experience (McDowell, 1996, 9). There is then an inextricable cooperation between receptivity and spontaneity, so that "receptivity does not make an even notionally separable contribution to the cooperation" (McDowell, 1996, 51).

An important feature of McDowell's conceptualism is his rejection of what could be called a two-stage model of conceptual involvement, according to which perceptual experience comprises two temporally discrete stages: a first stage of sheer sensory awareness, in which a manifold of bare sensations is presented through a self-standing exercise of our faculty of sensibility, followed by a second stage in which these bare sensations are 'brought under' concepts.¹ Evans's treatment of perceptual content, McDowell may say, succumbs to a version of this model in the sense that it attempts an impossible separation between the

¹ I am greatly indebted here to James Conant's notion of a two-stage reading of Kant's *First Critique* – one of the standard Anglophone readings of Kant that he (and McDowell) vehemently reject. See Conant, 2016, 90-91.

contributions of receptivity and spontaneity (McDowell, 1996, 51). Evans claims that perceptual experience provides non-conceptual informational states independently of the operations of spontaneity, and that only *later* do concepts arrive, when we make judgments about experience (Evans, 1982, 227). This view diametrically opposes McDowell's position, in which the understanding is inextricably intertwined with all deliverances of receptivity. In contrast to Evans's division between non-conceptual perception and conceptual judgment, McDowell claims that what happens in perception is an "opportunity for judging" (McDowell, 2013a). This means that judgments of experience do not introduce a new kind of content, but simply endorse the conceptual content that is already possessed by the experience on which it is grounded (McDowell, 1996, 48-9). In fact, for McDowell, having things appear at all can only take place through the operation of conceptual capacities, for "in experience one takes in, for instance sees, *that things are thus and so*" (McDowell, 1996, 9). I do not see, say, a formless desert with intermingling shapes, only then to conceptually structure this chaos; my perception already comes in the form, say, of the paper in front of me and of the pen by my side. I believe these objects to be here with their specific colors and shapes because I see them *thus and so*. No matter how hard I find it to describe the taste of a fruit, or the smell of my childhood home, these experiences are nowhere *blind*² – words might fail me, but I *taste* a specific fruit (which I may know to be the taste of a tangerine and *not* of a pineapple), and I recollect the *smell* of my bedroom (which I can judge to belong to my childhood and *not* to my early adulthood). These experiences must come, so to speak, "saddled" with concepts to qualify as experiences at all, about which I can judge. Evans's view of experience as non-conceptual informational states prevents experience from rationally justifying judgments.³ By suggesting that perception "produces its content-bearing states independently of the operations of spontaneity" (McDowell, 1996, 51), Evans commits to a flawed two-stage model of conceptual involvement.

What I have mentioned so far captures the core of McDowell's conceptualist thesis. However, one could still question what the conceptual investment of experience amounts to exactly. With regards to this, it is worth distinguishing between kinds of conceptual content and determining what McDowell says about them. McDowell does

² My usage of the word "blind" here refers especially to A50/B74 of Kant's first *Critique* and to page 4 of McDowell's *Mind and World*, where one reads: "intuitions without concepts are blind." This means that experiences unguided by concepts – i.e., 'blind intuitions' in the Kantian jargon – do not exist.

³ As McDowell is forced to admit, Evans's position here tips the seesaw back to the Myth of the Given (McDowell, 1996, 51), since the evidence for our empirical judgments would come from non-conceptual informational states.

not say that we need to apply previously-learned lexical concepts to have an experience at all; for instance, we do not need to possess the concept of “cardinal” to perceive the cardinal that settles on our window (McDowell, 2009, 261). If McDowell were arguing that perceptual content is composed entirely of lexical concepts – so that we do not experience anything for which we do not have a word – his view would fall prey to fineness-of-grain arguments, which claim that experience transcends our conceptual powers in its detail.⁴ Instead, McDowell says that one can employ a demonstrative concept that is exactly as fine-grained as what is experienced – e.g., “I prefer *that* shade of red,” or “I am *this* tall” (McDowell, 1996, 57). It is thus a mistake to think that claims about which specific concepts play a role in experience (and which do not) constitute the primary concern of conceptualism. Rather, it is experience’s *openness to reason* – whether it be through lexical or demonstrative concepts – that throughout has taken center stage in the conceptualist doctrine. We find this expressed in several passages:

It is essential to conceptual capacities, in the demanding sense, that they *can be exploited* in active thinking... When I say the content of experience is conceptual, *that* is what I mean by “conceptual” (McDowell, 1996, 47, my italics).

An intuition’s content is all conceptual, in this sense: it is the intuition in a form in which one *could* make it, that very content, figure in discursive activity (McDowell, 2009, 265, my italics).

[This] is what it means for capacities to be conceptual in the relevant sense: they are capacities whose content is of a form that *fits it to figure* in discursive activity (McDowell, 2013b, 42, my italics)

The above three passages express a distinctive theory of experience which has been called “epistemic conceptualism” (Mazijk, 2017, 64), according to which experience has conceptual content because it belongs to the space of reasons, i.e., because it is *open to judgment*. Although there could exist many variants of epistemic conceptualism which involve an additional claim specifying the exact role of conceptual capacities in experience (say, requiring the operation of lexical concepts), I take McDowell to demand simply that the contents of all experience be open to judgment. I take *this* sense of epistemic conceptualism to be McDowell’s principal thesis.

⁴ For a great discussion of fineness-of-grain arguments, see Evans 1982, 229, Kelly 2001 and Tye 2006.

II.

McDowell has already had a notable influence on those interested in bringing phenomenology into dialogue with other areas of research, and with the analytic tradition in general. However, many phenomenologists have expressed concerns about McDowell's exclusion of non-conceptual content, often assuming (wrongly, I will claim) that his theory of conceptual intuitions rests on an over-intellectualization of human experience (Scheur, 2013, 294-299). One of these critics is Hubert Dreyfus (2013), who turns to the works of Heidegger and Merleau-Ponty for examples of pre-reflective, skillful action in order to challenge McDowell's conceptualism. For Dreyfus, the idea that the understanding is "inextricably implicated in the deliverances of sensibility" (McDowell, 1996, 46) is fundamentally at odds with the phenomenological description of what he calls skillful or absorbed coping.

Dreyfus presents two arguments that charge McDowell with over-intellectualization. The first one points to examples of absorbed or skillful coping. A particularly famous example from

Heidegger is that we do not have to think about the doorknob in order to use it to enter the room. "Once a skill is acquired," Dreyfus writes, "concepts used in learning the skill need play no further role" (Dreyfus, 2013, 18). The specific nature of these skillful copings seems to involve a kind of absorption foreign to conceptual apprehension. The doorknob, in fact, does not have to be apprehended at all. These copings are *mindless* activities; they have no intentionality, that is, no distance between a subject and an object.⁵ For Dreyfus, this means that it is inappropriate to characterize their contents (in so far as they would have any content at all) as conceptual. To acknowledge the phenomenological structure of absorbed coping is to allow experience to outstrip the operations of the understanding. Dreyfus's desire to account for such experiences leads him to conclude that McDowell's theory of conceptual intuitions rests on an overintellectualization of human experience that disregards the phenomenology of skillful coping.

Dreyfus's second argument focuses on situated normativity (Dreyfus, 2013, 23-7). Like his phenomenological account of embodied action, this argument contains several descriptively rich examples. For instance, Dreyfus claims that to know one's distance from someone else

⁵ I must acknowledge that there is something vague and metaphorical about this definition of intentionality as a *distance* between subject and object. However, this is precisely Dreyfus's terminology. See, e.g., Dreyfus's phrase: "For the player in action, the soccer field is not an 'object'... He has no *distance* from it" (Dreyfus, 2013, 17).

in an elevator, and to behave appropriately according to that knowledge, does not require any concepts

(Dreyfus, 2013, 23). In other words, by knowing how far I am supposed to stand from someone else in an elevator, I act appropriately in an immediate and mindless way: I do not stretch my arms very widely; I make some room for the other to reach the elevator buttons upon entering; I look in a certain direction to avoid staring. In these situations, I lack awareness of my actions, but once I step back to grasp it conceptually, I no longer find myself drawn with ease to the culturally appropriate act, but likely end up positioning myself inappropriately (Dreyfus, 2013, 23). As before, the premise is that conceptual involvement requires an intentional distance between subject and object. Such a distanced attitude of the mind towards the world does not obtain in situated normativity – and hence, Dreyfus thinks, conceptualism cannot account for it.

Dreyfus has thus offered two phenomenological arguments that deny conceptual content to absorbed actions. I will now analyze whether his criticism poses a genuine problem for

McDowell's conceptualism. The basic structure of his arguments appears to be as follows:⁶

(P1) Exercising conceptual capacities requires an intentional distance between subject and object.

(P2) Absorbed coping and situated normativity preclude an intentional distance between subject and object.

(C1) Absorbed coping and situated normativity are not available to conceptual capacities, i.e. are non-conceptual.

The argument against conceptualism rests in part on the first premise (P1): that conceptual capacities require an intentional distance between subject and object. This assumption, however, is denied by several philosophers, including Noë (2013), Crane (2013) and McDowell himself

(2013b). McDowell explicitly says that Dreyfus mistakenly saddles him with a picture in which “a rational agent is always at least marginally monitoring what she is doing, standing ready to intervene

⁶ I base this syllogism largely on Schear's reading of Dreyfus (Schear, 2013, 294). See another rendering of this syllogism in Mazijk (2017, 57).

with full-blown monitoring if need be” (McDowell, 2013, 45). This picture, of course, would leave no room for “total absorption” (Dreyfus, 2013, 28), and McDowell agrees that an account of agency with no room for acting in flow would be disastrous (McDowell, 2013, 45). Dreyfus’s understanding of McDowell’s conceptualism as requiring a subject-object distance would entail that, in McDowell’s account, a subject always has a detached contemplative relation to the world and a self-reflexive monitoring of her actions. However, Dreyfus cannot hold his argument against McDowell for several reasons, which I will go on to describe.

In defining *epistemic conceptualism*, I explained that McDowell only requires that the contents of experience have the appropriate structure to figure in judgments or discursive activity (McDowell, 2009, 265; McDowell, 2013b, 42). In the case of skillfully opening a door, this requirement is met by two facts. First, I can make the skillful action of using the doorknob figure in a judgment. For instance, if asked what I am doing, I immediately reply that I am using the doorknob to open the door; and if one denies that I used the doorknob to open the door, I will judge that claim to be wrong. Secondly, I can give reasons for having performed this action in hindsight. For example, I can explain that I wanted to hand in a paper to a professor in the room, and that I needed to turn the doorknob to the right instead of to the left to open the door. Other cases of “total absorption” likewise meet the requirements of McDowell’s conceptualism. A skillful squash player, e.g., may “become one with a phenomenal field” (Dreyfus, 2013, 17), but if she suddenly uses her racket to hit flies in the court instead of the ball, her opponent may ask her to justify herself and explain her actions. Playing Squash, like playing any game, involves abiding by a set of rules that justify some actions while forbidding others. Because our actions in a game can always be subject to normative evaluation, even the actions performed “in flow” must be, at least in principle, *open* to judgment and to discursive articulation. None of Dreyfus’s examples of absorbed coping or situated normativity escape McDowell’s conceptualist requirement (1) that I can make my action figure in judgment and (2) that I can give reasons for having performed it in hindsight. If McDowell’s conceptualism demands only the openness to reason of all experience, Dreyfus cannot use his examples to argue for non-conceptual content.

I have made the point that Dreyfus’s examples of “total absorption” cannot be invoked to argue for non-conceptual content, for they meet the requirements of McDowell’s epistemic conceptualism – that is, they can (1) figure in judgments and (2) be justified in hindsight. One might think, however, that Dreyfus would not find this argument

satisfactory. After all, for him, McDowell's demand that all experience be open to reasons requires that the rational agent be always monitoring her own actions from a critical distance. As a result, in Dreyfus's reading of McDowell, the rational agent would never enter a state of full absorption. Therefore, Dreyfus might say that the skillful coper is not open to reasons *during* the state of absorption. Though the actions of the skillful squash player must abide by normative evaluation, these actions are not open to judgment *while* the player is absorbed in the game – for if she stops to consider whether she has violated a rule, she loses her state of absorption by over-intellectualizing her experience.

The premise in the criticism above is that the exercise of conceptual capacities demands an active stance on the agent's part. However, McDowell nowhere defines conceptual capacities as solely active exercises. On the contrary, he says that the involvement of conceptual capacities in experience “is exactly not a matter of activity on the part of the subject” (McDowell, 2013b, 57). Dreyfus seems to miss out on the fact that, for McDowell, human beings can *passively* employ capacities that belong to the understanding, that is, without actively contemplating the relevant content involved, and hence without distorting the specific phenomenology of skillful coping and situated normativity (McDowell, 1996, 87-95). According to McDowell's definition of *Bildung*, human beings engage in a process of cultural development by which they develop a second nature, i.e., a vast collection of habits of thought that structure experiences independently of the agent's deliberation. Such a second nature allows for “having one's eyes opened to reasons” (McDowell, 1996, 84), that is, for having sensory experiences that constitute reasons for belief and that have the appropriate conceptual structure to figure in belief states. McDowell wants to convince us that experience is ‘shot through’ our conceptual capacities due to our cultural upbringing, which results in the openness of experience to reasons. He would readily accept that I do not have to actively conceptualize a rapidly moving squash ball to successfully hit it. His point is that concepts are *passively* drawn upon in skillful coping rather than being actively employed. Hitting a squash ball, like using a doorknob, is an absorbed action that cannot be performed without having first developed the relevant background knowledge by a process of *Bildung* or cultural training. When Dreyfus discusses a game of chess, for instance, he states that a chess master may be “directly drawn by the forces on the board,” without making his move for any consciously entertained reason (Dreyfus, 2013, 35). McDowell's response to this example is the same as to all others: “cultivated rationality... is operative in his [the player] being drawn to make his move by the

forces on the board” (McDowell, 2013b, 48). McDowell, then, does not reject the phenomenology of skillful coping and situated normativity. He simply claims that absorbed actions possess their passive character due to cultivated rationality, which operates independently of the agent’s deliberation.

Aside from the unwarranted assumption that conceptual capacities demand an active stance on the agent’s part, there is yet another problem with the notion that conceptual capacities are not operative *during* the state of absorption. We may wonder, following McDowell’s worries, whether this picture would not tip the seesaw back to the Myth of the Given. For if Dreyfus is right, and conceptual capacities are not operative during the state of absorption, *whence* would come the reflective judgments that players make *after* matches? McDowell himself mentions that a chess master should be able to say, for instance, “That was a good move, because it threatened my opponent’s queen” (McDowell, 2013b, 47). Dreyfus, on his part, says that the chess master would only be able to utter phrases of the sort: “I made this move because I was drawn to make it” (Dreyfus, 2013, 35). The point, however, is that both phrases contain a rational explanation for the chess master’s move – they both constitute a *giving of reasons*. “Of course,” McDowell admits, “the chess master will need to break the flow to engage in this kind of conversation... [But this] is irrelevant. He is giving expression to something he already knew if he explains his move as a response to the forces on the board” (McDowell, 2013b, 48). The problem with the idea that no concepts are operative *during* the state of absorption is that the player’s giving of reasons, *after* breaking the flow, would supposedly give expression to something “beyond the reach of [his] conceptual capacities” (McDowell, 1996, 5). When the chess master judges his move to be good, or explains it by saying that he was simply drawn by the forces on the board, he is reflecting on his absorbed experience. However, *if* his judgment about the game indeed belongs to the space of reasons (as McDowell thinks it does), then the player’s state of absorption (the content that the judgment is *about*) cannot lie outside the reach of his conceptual capacities. Otherwise, the player’s conceptual judgments would be grounded on an extraconceptual presence, which is precisely the danger that McDowell’s conceptualism is called to resolve. In a word, if the player can pass judgment on his moves or explain his tactics, his state of absorption *must* be conceptual in McDowell’s sense – that is, it must be *open to reasons*.

III.

A critic might still argue that I have only shown that Dreyfus’s

phenomenological description of absorbed action does not violate McDowell's conceptualism. I have not considered, however, whether there are any other experiential phenomena to which we simply cannot do justice once we abide by McDowell's conceptualist doctrine. In other words, though McDowell can account for skillful actions, his demand that the understanding be "inextricably implicated in the deliverances of sensibility" (McDowell, 1996, 46) might make it impossible for him to do justice to some other phenomena. This leads us to a second critique of McDowell, namely Walter Hopp's argument for non-conceptual content based on Husserl's early work.

Hopp charges McDowell with failing to explain "why perceptual experiences play such a distinctive role in the production of knowledge" (Hopp, 2011, 2). Hopp draws on Husserl's writings on the structure of fulfillment to argue that McDowell cannot account for the specific kind of epistemic input that only perception can offer. To put it simply, Husserl thinks that thoughts belong to a class of empty acts. In thinking about a red apple, for instance, I intend something "emptily," that is, without the apple being actually present to me "in the flesh," i.e., in perception (Hopp, 2011, 103). Acts of perception, on the contrary, primarily involve the kind of fulfillment that thought lacks. Thinking about an apple and subsequently seeing it fulfills the emptiness of the thought by the actual givenness of the object. As Hopp puts it, the object is no longer "merely meant or represented, but presented" (Hopp, 2011, 103). Husserl's distinction between perception and empty acts, Hopp thinks, allows us to get a grip on the nature of conceptual content and to understand why conceptualism is misguided (Hopp, 2011, 104).

According to Hopp, McDowell's conceptualism abides by two theses that, when seen in conjunction, *must* be false (Hopp, 2011, 106). The first of these, the *detachability thesis* (DT), states that "C is a conceptual content only if it is a detachable content," that is, if it is possible for C to serve as the content of a mental state in which the relevant objects that C is about are not perceptually present to the subject of that mental state (Hopp, 2011, 105). Hopp corroborates this thesis by pointing to instances where McDowell seems to endorse DT. For example, McDowell says that we can ensure that we have the concept of a certain color if the "capacity to embrace [that] color in mind can in principle persist beyond the duration of the experience itself" (McDowell, 1996, 57). Hopp concludes from this that, on McDowell's view, a conceptual capacity must be such that "it can be exercised in the perceptual absence of the object that it bears upon" (Hopp, 2011, 106).

The second thesis that Hopp attributes to McDowell, the *conceptualist principle* (CP), establishes that the justification of a belief is determined solely by (1) the belief's conceptual content and (2) the conceptual contents of those mental states to which its content is inferentially related (Hopp, 2011, 107). To understand CP, it is perhaps helpful to consider it in terms of what it rules out. CP rules out the possibility that, if a perceptual experience has *any* non-conceptual content, that content could contribute to the reason for one's beliefs. McDowell, of course, denies that perceptual experience has *any* non-conceptual content. According to him, it is a hopeless, albeit seductive, mistake to "extend the scope of justificatory relations outside the conceptual sphere" (McDowell, 1996, 7). McDowell thinks that any property an experience might have, beyond its possession of conceptual content, is irrelevant from an epistemic standpoint – and, as Hopp indicates, this means that McDowell abides by CP (Hopp, 2011, 109).

Now comes Hopp's argument: If CP is correct, then any two beliefs with the same conceptual content justify to the same degree. However, if all conceptual contents are detachable, then all the epistemic work in knowledge could potentially be carried out by empty acts. (Hopp, 2011, 109). In other words, if the sorts of conceptual contents present in perception were the same as those of possibly empty acts, then we should be able to "zap all the perceptual states, and replace them with mere beliefs with the same conceptual contents" without thereby affecting the degree to which a belief is justified (Hopp, 2011, 110). However, this is clearly absurd: empty thoughts can never play the justificatory role that perception can. Hopp illustrates his argument with an example:

Suppose Jones believes that it has recently rained on the basis of 1) his perception that Beacon Street is wet and 2) his belief that if Beacon Street is wet, then it has recently rained. Now suppose we zap his perception and replace it with a mere belief with the same content: Beacon Street is wet. In doing so, we also effectively annihilate his justification for believing that it has recently rained. But if the contents of perception were the same as those of possibly empty intentional states, then we would not have altered his justification in the slightest (Hopp, 2011, 110).

Hopp argues in this passage that perception must have an extra-conceptual content, since the conceptual contents of empty thoughts can never deliver that distinctive contribution in the process of justifying beliefs.

For this reason, Hopp thinks that McDowell is unable to address the different epistemic roles played by perception and thought respectively.

Hopp summarizes his critique of conceptualism by asking the following question: “Why is basing the belief that *p* on the perception that *p* a sound epistemic policy, while basing it on itself is not... if they have the same conceptual content?” (Hopp, 2011, 110). In response to Hopp, McDowell might emphasize that his treatment of perceptual experience depends not *only* on spontaneity, but *also* on the contributions from receptivity, which ensure that such experience is not reducible to empty thought. Of course, the conceptual content of receptive experience is different from the conceptual content of purely spontaneous activity. McDowell’s point is simply that, for perceptual experience to have justificatory import, spontaneity must be inextricably implicated in the deliverances of receptivity. This nowhere implies that one’s experience of the taste of a fruit (which puts into operation both receptivity *and* spontaneity) has the same content as the anticipation that I may have of that fruit’s taste (which involves *solely* my faculty of spontaneity). McDowell’s ‘only’ point – though it is one with several implications – is that the very difference between perceptual experience and purely spontaneous exercises must be conceptually accountable if it is to have justificatory or rational significance. If McDowell were denying the specific function of receptivity in knowledge, he would fall into the kind of frictionless coherentism that he is eager to avoid. Hopp, it seems, misunderstands McDowell’s unwillingness to discuss the contributions of spontaneity and receptivity in terms of separable contents as indicating that there is no contribution of receptivity at all in perceptual experience.

Furthermore, if we conceive of perceptual experiences as fulfilling empty thoughts (as Hopp suggests we do), then perception would have to play a confirming or negating role in relation to the conceptual content of the empty thought it fulfills (or frustrates).⁷ For instance, if I think that there is milk in the refrigerator, and then open the fridge and see that there is none, my perception of the fridge’s interior negates the conceptual content of my empty thought.⁸ The experience which fulfills (or frustrates) the conceptual content of my empty thought *cannot* involve an extra-conceptual presence, for otherwise it would not enter into rational relations with the content it is meant to fulfill. In a word, Hopp’s idea that perception delivers its distinct contribution thanks to extra-conceptual contents would lead us back to the Myth of the

⁷ My argument in this paragraph is greatly indebted to Michael D. Barber (2011, 69-70).

⁸ Hopp gives precisely this example when introducing Husserl’s notion of fulfillment (Hopp, 2011, 103).

Given. The very idea of fulfillment would not be possible if perceptual experiences were bare givens apart from conceptualization. It is *because* McDowell wants perception to fulfill empty thoughts that he rejects a non-conceptual Given that cannot fulfill anything. By demanding that receptive experiences involve the passive activation of conceptual capacities, McDowell's conceptualism accounts for how perception can deliver its distinct contribution to knowledge. Thus, Hopp's discussion of fulfillment does not contravene, but rather requires, McDowell's conceptualism.

I have argued that neither Dreyfus nor Hopp successfully criticize McDowell's conceptualist thesis. My argument does not entail, clearly, that we should reject Dreyfus's descriptions of absorbed actions or Hopp's distinction between perception and thought. After all, McDowell himself never attempts to reject Dreyfus's account of what it means to act in flow; rather, he repeatedly states that their disagreements are based on a misunderstanding (McDowell, 2013b, 54). I also do not think that McDowell would ever deny Hopp's point that perception and thought differ in phenomenologically obvious and epistemically relevant ways (Hopp, 2011, 103). My claim is that Dreyfus and Hopp can coherently accept conceptualism in the form proposed by McDowell without doing the least harm to their accounts of experience; McDowell, in turn, can profit from these phenomenological accounts, for they serve as counterexamples to any claim that conscious deliberation or distanced representation is essential to conceptual understanding, which would rule out some specific *accounts* of conceptual understanding.

To understand my claim that Dreyfus's and Hopp's accounts of experience both cohere with and contribute to McDowell's conceptualism, it is helpful to distinguish between two approaches to conceptual content. Joseph Rouse, commenting on the McDowell-Dreyfus debate (Rouse, 2013), asserts that whereas McDowell is interested in an epistemologically *normative* account of mental content, Dreyfus focuses only on *description* (Rouse, 2013, 252). Descriptive accounts take conceptual content to be something actually present or operative in specific performances by concept users. To use a concept, on this view, is to have something in mind; for instance, concept use might involve having token mental states that possess representational content (Rouse, 2013, 250).⁹ Normative approaches to conceptual content, by contrast, identify the conceptual domain with those performances that are appropriately

⁹ Jerry Fodor (1998) is exemplary of this approach. As Rouse indicates (2013, 250), Fodor begins his book on *Concepts* by saying, "The scientific goal in psychology is to understand what mental representations are... Nothing about this has changed much since Descartes" (Fodor, 1998, vii).

assessed according to rational norms. In this case, whether certain kinds of representations or structures are actually present in a particular thought or action does not matter, but only whether that thought or action is potentially responsive to rational or conceptual assessment. Rouse thinks that this distinction provides an illuminating context to understand the misunderstanding between Dreyfus and McDowell (Rouse, 2013, 251) – and, I think, between Hopp and McDowell as well.

Dreyfus's examples highlight the "mindlessness" of expert understanding: skillful players "in flow" do not have representations in mind, but instead respond directly to the solicitations of a situation on the field. According to Rouse, Dreyfus's examples are only relevant challenges to *descriptive* accounts of the conceptual domain (Rouse, 2013, 254). The point of these examples is that chess masters and skillful players need not, and even cannot, have concepts "in mind" and take up a stance of reflective detachment while in a state of flow. McDowell could and does agree with Dreyfus that conceptual capacities operate without the subject explicitly attending to a concept and its application. For what matters to McDowell's normative account of conceptual understanding is not whether concepts are actively employed during one's actual performances. The only issue is whether those performances are accountable and responsive to the relevant rational norms. Having this distinction in view allows us to see how Dreyfus's examples might contribute to the phenomenology of conceptual understanding: they would be counterexamples to any claim that the exercise of conceptual capacities requires conscious deliberation. Rather than *limiting* the scope of conceptual understanding (as Dreyfus proposed), his examples would lay out new and subtle aspects of conceptual experiences. On several occasions during their conversations, Dreyfus asked McDowell "how the pervasiveness of conceptual norms is actually experienced by perceivers" (Rouse, 2013, 268). In the picture that I have been recommending, McDowell's response should be that Dreyfus himself has already described that experience on his behalf.

Similarly, McDowell would agree with Hopp's distinction between merely thinking about something emptily and having it present in perception. Hopp's argument differs from Dreyfus's in that it is both descriptive *and* normative. To the extent that Hopp emphasizes the phenomenological difference between perception and belief, he coheres with McDowell. To the extent that Hopp defends an extra-conceptual content of perception, he opposes McDowell. For him, reason-giving relations "do not hold solely among mental states with conceptual content" (Hopp, 2011, 2), and this clearly violates the core

of conceptualism. If the reader accepts my reply to Hopp – i.e., that we cannot understand fulfillment without committing to the view that what fulfills an empty thought must be conceptual – then we can retain the distinction between perception and thought while endorsing the notion that perceptual experience is conceptual.

My argument in this paper allows the phenomena described by Dreyfus and Hopp to bring into view the diverse range of belief-justifying perceptual experiences. Though McDowell's project does *not* involve describing the varieties of conceptual experience, and though Dreyfus's project does *not* hope to show how experience can be accountable to rational norms, one may pursue a *holistic* account of perceptual experience that unifies both descriptive and normative approaches to conceptual content. This holistic account of perception would, on the one hand, explain how experience can serve as a source of knowledge; on the other hand, it would lay out in detail all the subtle and nuanced aspects of that experience. Choosing either McDowell's conceptualism or Dreyfus's and Hopp's phenomenological accounts would leave out half of this project. However, as I have argued in this paper, we may retain Dreyfus's and Hopp's wealth of description while endorsing McDowell's epistemological commitments.

BIBLIOGRAPHY

- Barber, M. D. (2011). *The Intentional Spectrum and Intersubjectivity: Phenomenology and the Pittsburgh Neo-Hegelians*. Athens, OH: Ohio University Press.
- Conant, J. (2016). Why Kant Is Not a Kantian. *Philosophical Topics*, 44 (1), 75-125.
- Crane, T. (2013). The Given. In J. K. Schear (Ed.), *Mind, Reason, and Being-in-the-World: The McDowell-Dreyfus Debate* (229-250). New York: Routledge.
- Davidson, D. (2009). *Subjective, Intersubjective, Objective*. Oxford: Oxford University Press.
- Dreyfus, H. L. (2013). The Myth of the Pervasiveness of the Mental. In J. K. Schear (Ed.), *Mind, Reason, and Being-in-the-World: The McDowell-Dreyfus Debate* (15-40). New York: Routledge.
- Evans, G. (1982). *The Varieties of Reference*. Oxford: Oxford University Press.
- Fodor, J. (1998). *Concepts: Where Cognitive Science Went Wrong*. Oxford: Oxford University Press.
- Hopp, W. (2011). *Perception and Knowledge: A Phenomenological Account*. Cambridge: Cambridge University Press.
- Kelly, S. (2001) Demonstrative Concepts and Experience. *Philosophical Review*, 110 (3), 397-420.
- Mazijk, C. van (2017). *Husserl on Concepts in Perception*. Groningen: Rijksuniversiteit Groningen.
- McDowell, J. (1996). *Mind and World*. Cambridge, MA: Harvard University Press.
- McDowell, J. (2009). *Having the World in View: Essays on Kant, Hegel, and Sellars*. Cambridge, MA: Harvard University Press.
- McDowell J. (2013a). Are the Senses Silent? Agnes Cuming Lectures, UCD School of Philosophy. Retrieved from URL: <<https://www.youtube.com/watch?v=fBQHEGg5JSo>>
- McDowell, J. (2013b). The Myth of the Mind as Detached. In J. K. Schear (Ed.), *Mind, Reason, and Being-in-the-World: The McDowell-Dreyfus Debate* (41-58). New York: Routledge.

- Noë, A. (2013). On Overintellectualizing the Intellect. In J. K. Schear (Ed.), *Mind, Reason, and Being-in-the-World: The McDowell-Dreyfus Debate* (178-193). New York: Routledge.
- Rouse, J. (2013). What is Conceptually Articulated Understanding? In J. K. Schear (Ed.), *Mind, Reason, and Being-in-the-World: The McDowell-Dreyfus Debate* (250-267). New York: Routledge.
- Schear, J. K. (2013). Are We Essentially Rational Animals? In J. K. Schear (Ed.), *Mind, Reason, and Being-in-the-World: The McDowell-Dreyfus Debate* (285-302). New York: Routledge.
- Sellars, W. (1991). *Science, Perception and Reality*. Atascadero, CA: Ridgeview Publishing Co.
- Tye, M. (2006). Non-Conceptual Content, Richness, and Fineness of Grain. In T. S. Gendler and J. Hawthorne (Eds.), *Perceptual Experience* (504-530). Oxford: Oxford University Press.

A Description-Based
Solution to the
Non-Identity Problem in
the Context of Mitigating
Historical Injustice

Brian Wong Yue Shun
Pembroke College, University of Oxford

ABSTRACT

Parfit (1984)'s Non-Identity Problem poses a central threat to the questioning of mitigating historical injustice: taken at face value, it appears to be the case that descendants of victims of historical injustice have no claims to compensation, on the basis that they would not have come into existence had the initial injustice not occurred. This article advances a *de dicto* descriptive solution to the challenge: through conceiving of the descendants' claim to compensation as grounded upon their relations to their *de dicto* counterparts in worlds where they have higher net welfare, this account circumvents the Non-Identity Objection by pinpointing the individual description as the subject of harm under historical injustice.

INTRODUCTION

The Counterfactual Account of compensation posits that descendants of victims are owed compensation as they are left *worse off* as a result of the initial acts of injustice perpetrated upon their ancestors (Nozick, 1974). There exists a wide range of objections to the Counterfactual Account – that it is epistemologically indeterminate and self-defeating (Waldron, 1992); that it does not account for why present generations or descendants of the perpetrators have the obligation to compensate victims (Butt, 2013); that it erroneously assumes that the counterfactual would necessarily have had the descendants lead better lives, and – finally – that compensation is unjustified in cases where the descendants would not have existed, had it not been for the initial act of injustice (the Non-Identity Problem) (Parfit, 1984). This article concedes all of these issues, but focuses predominantly on showing why – at the very least – the Non-Identity Problem does not function as a valid objection to the account. Compensation is still justified on grounds of *de dicto* harm committed towards a general description, which in turn wrongs the particular individual that fits that description in the actual world.

THE QUESTION OF COMPENSATION

Consider the following paradigmatic case – *Slavery*. As a historical process, the slave trade coercively shipped millions of individuals from Africa to Europe and the Americas, against the will and wishes of the slaves. A commonly raised political claim is that the descendants of the slaves who are alive today are owed compensation for the historical

injustice performed upon their ancestors. Whilst this claim appears to be intuitively plausible and valid, the underlying theoretical question remains – *why* are these descendants owed such compensation?

THE COUNTERFACTUAL ACCOUNT

The Counterfactual Account posits that descendants of victims have claims of compensation because they would have been *better off* had the initial acts of injustice perpetrated upon their ancestors *not occurred*. As per the standard view of compensation (Nozick, 1974), “something fully compensates X for Y’s action A if X is no worse off receiving it, Y having done A, than X would have been without receiving it had Y not done A”¹ To the extent that X has been comparatively harmed as a result of an injustice carried out towards X’s ancestor, it appears that X has the *prima facie* claim right on some unspecified agent to be compensated. Let it be such that an individual suffers *counterfactual-comparative harm* if they are worse off than they otherwise could have been “in the most probable outcome if the injuring act had not occurred”²

There are several plausible mechanisms through which historical injustices can result in counterfactual-comparative harm for the descendants of their victims, through impacts on i) *financial inheritance (of property and wealth)*, ii) *capability and physical capacity*; iii) *psychological unity* or iv) *wider social and political structures*. i) to iv) are all cases where it seems that the welfare of the descendants *could have been much better had the initial injustice not occurred*. The Counterfactual Account above therefore seems to support the view that they are entitled to at least some claims of mitigation.

THE NON-IDENTITY PROBLEM

A key objection to this claim is the Non-Identity Problem (Parfit, 1984); this will be the primary objection considered and addressed in this piece. Three assumptions underpin this objection, which effectively seeks to show that descendants have no claim to compensation if – in a world without the initial injustice – they would never have been born. Firstly, the *fragility* of existence: had the conditions precipitating

¹ pg. 184, Sher, 2005

² pg. 6, Butt, 2010

an individual's conception³ been even slightly different, the *very same individual would not have existed*, and would have been replaced by another individual (even if they share largely similar characteristics and attributes). Roberts (2007) claims that the same individual could be generated despite differences in their genesis conditions; however, the *Butterfly Effect* of causation suggests that even a slight and seemingly irrelevant change at some time t_0 *could have affected the exact identity* of the individual who is conceived at a time later than or equal to t_0 . Trivially, an individual conceived at a slightly earlier or later time than t_0 does not share the exact same attributes (e.g. time of conception) as the individual conceived at t_0 . Secondly, the *comparative account of harm* by some action A: an individual X is not harmed unless X has less utility in a world with A than in a world without A. As such, an individual brought into a flawed existence (e.g. a child born into slavery) – so long as they are living above the threshold of 0 utility – is not harmed by their existence, because they would not have existed in the counterfactual. Thirdly, the *assumption that existence is preferable to non-existence*: a flawed life, whilst ridden with suffering and pain, is still preferable to the absence of the life. The implicit assumption is that flawed existences still incur positive utilities, whilst the absence of existence incurs none.

Given these three assumptions, it appears that descendants of victims of injustice are not harmed by events that precede their birth. Had these events not occurred, they would not have existed. To the extent that their welfare levels are better in this world as compared to any other world in which they do not exist, it appears that they are not harmed by the initial act of injustice. To the extent that any claims to compensation require the existence of some comparative harm, it allegedly appears that no descendant of victims of injustice has a legitimate claim to compensation.

A potential response to the above problem is the shift of the justificatory basis for compensation from *harm* to *wrong* – i.e. consider the Modified Counterfactual Account: descendants of victims have claims of compensation because they would not have been *wronged* had the initial acts of injustice perpetrated on their ancestors *not occurred*. The shift from *harm* to *wrong* expands the scope of potential features of events that warrant compensation, thereby potentially (as will be discussed) avoiding the specific objection that the individual would not have existed (and hence 'worse off') had the initial act of injustice not been perpetuated. Two types of wrongness accounts will be considered:

³ Assumed to be the moment upon which the identity of a life becomes fixed and does not change

the Uncomparative Wrong Account and the De Dicto Wrong Account. The first will be rejected; the second will be shown as preferable.

THE UNCOMPARATIVE WRONG ACCOUNT

Consider first the proposition that individuals could be wronged in a non-comparative manner, through the violation of their entitlement to living in a particular way. Woodward (1986) raises the view that "... it was possible to wrong a person by violating a specific obligation owed to that person even though one's actions advantageously affect that person's other interests in such a way as to make him, on balance, better off than any other action one might have taken."⁴ Woodward justifies this claim by appealing to a series of thought experiments⁵, which effectively seek to demonstrate that an individual can be wronged by a particular action X even though X improves their aggregate welfare. When applied to the cases of historical injustice (as outlined above), these obligations may include – the obligation to not introduce an individual into an existence with severe leg ailments, or governed by oppressive societal structures. It is worth noting that Woodward is deliberately ambiguous with respect to the subject of such obligations – these obligations could either arise from specific parental and procreative obligations (hence parent-specific), or from correlated, primitive rights to not be introduced into such forms of existence (i.e. these obligations arise from irreducible claims held by the individual). Both are possible subjects that render the obligation argument viable.

Parfit (1984) raises an objection to this argument based on the absence of regret. He reasons that it is unlikely that individuals who live flawed existences would nonetheless regret having been born, for they intuitively find their existence preferable to non-existence in the first place. To the extent that there is an absence of regret, it is unclear why they are justifiably owed any compensation for their admittedly flawed, but non-regrettable existences. This objection undermines Woodward's claim, because it posits that to the extent that even though such prima facie obligations are violated, the absence of actual regret in the flawed existence suggests that no compensation is required.

There are three possible replies to Parfit that will allow Woodward's uncomparative wrong account to withstand this challenge. The first reply

⁴ pg. 812, Woodward, 1986

⁵ e.g. the racist airline and the Nazi concentration camp

is the *hypothetical denial* – it is deeply unclear if individuals indeed do not regret their flawed existences. Trivially, individuals who are born under severely oppressive social structures inherited from slavery or colonialism may emotively resent their existence, to the point of regretting it. A potential rejoinder to this reply may be to posit that Parfit's claim is not a postulation about whether individuals *actually would regret* their existences, but whether it is *reasonable for them to do so*. However, this rejoinder is susceptible to the following two objections.

Firstly, the argument from *intrinsic wrongness* – an event X necessarily entails Y; whilst X is not regrettable, the intrinsic wrongness in Y can still be regrettable and hence a legitimate ground for compensation, in spite of its being the necessary consequent of X. To visualize this, consider the *Rape*. As a result of an unconscious rape of which he is unaware and has no memory of (i.e. there is no change to his subjective conscious states), a man becomes substantially wealthier than he would have been otherwise. The man's becoming substantially wealthier (X) entails (logically) the occurrence of the rape (Y). However, the man still has a legitimate claim to compensation upon his rapist, for the very fact that the rape's violation of his autonomy and consent is intrinsically wrong. The intrinsic wrongness sources from a violation of the man's non-phenomenal attributes (i.e. attributes that exist independent of his experience). Even if he experiences only net positive utility (from becoming significantly wealthier), he is still owed compensation for the specific act of wrongness that cannot be mitigated by the benefits that necessarily follow (Woodward, 1986).

Secondly, the argument from *temporal misplacement* – Parfit at best demonstrates that it is unreasonable to regret one's existence *after* one's being brought into existence. However, it is worth noting that a defendant of compensation could circumvent this challenge by positing that an individual has claimed to compensation at some time t_1 after their conception, based on whether or not they could justifiably object to their existence at some time t^* *prior to their conception*. The logic here parallels most appeals to ideal choice theory – *had the individual been able to choose at a time prior to their conception, would they have chosen otherwise?* To the extent that individuals may object to being born as children of slaves and/or individuals with severely disfigured legs due to historical injustice, it appears that the regret-based challenge can be largely mitigated.

Whilst Parfit's immediate challenge (1984) appears to be tentatively resolved, this account as a whole remains unconvincing. Consider first

the *temporal misplacement* challenge above – whilst this challenge successfully bypasses Parfit’s objection by shifting the time at which the regret test is applied, it has a gaping metaphysical flaw: prior to the time of conception, the individual *does not exist*. As such, the question of whether the *individual* would object to their own existence *prior to the time of conception* is inherently self-defeating, because the subject (the individual) does not exist. Furthermore, consider the *intrinsic wrongness* argument above – it is deeply unclear if most descendants of victims of historical injustice are indeed violated to the same extent as the victim in the *Rape* case. Moreover, the account appeals to some arbitrary list of ‘intrinsic wrongs’ that is – at best – highly contestable and undefined. To the extent that a defender of Woodward replies with the claim that there may exist an obligation to ‘not introduce an existence who is a descendant of a victim of historical injustice’ (a crafty attempt to define *out* the problem), it is unclear whether or not such an obligation even exists. This obligation seems to suggest that there exists an obligation to prevent any victims of historical injustice from having any children – but this seems at best counterintuitive, and at worst deeply contrary to the basic understanding of procreative liberties. Woodward (1986)’s version of uncomparative wrongness therefore seems to be futile in responding to the Non-Identity Problem, let alone justifying compensation to descendants of victims of historical wrongs.

An alternative version of the Uncomparative Wrong Account is a *thresholdist* view (Steinbock, 2009) that an existence-introducing act may wrong a person by introducing them to a life that is *barely worth living*, but still beneath the *threshold* to which a reasonable person is entitled. In other words, descendants of victims may lead lives that incur net positive utilities, but are not worth living despite the net positive utilities. This seems intuitively justifiable from an appeal to sufficientarianism: the bare ‘minimum’ of any life that is *worth living* should encompass a set of sufficiently satisfied capabilities and capacities (cf. Nussbaum). A life *worth living* intuitively seems to exist above a threshold of zero net utility, at which point the individual remains indifferent to whether they are alive or non-existent. Any life deemed to be *worth living* should be pursued and preferred *actively* to death. It seems, therefore, that the actual *threshold* should be at least somewhat higher than 0 utils. If this account is true, then individuals can be wronged by being brought into existence – even if their existences incur positive utilities (that nonetheless come beneath the thresholds). If individuals are wronged, then they by the above criteria – they can be owed compensation.

There are two issues with this response. Firstly, it is deeply

unclear whether such a threshold exists in a meaningful sense. If the threshold is set at a point that is too high (Steinbock, 2009), it appears to insinuate that a large quantitative number of lives that *currently exist* are not worth living. Whilst the variabilist claim that *prior existences* matter more, and hence the threshold of harm required to remove a life is substantially higher than that required to not introduce a life, this still counterintuitively imposes a judgment that many individuals are not leading lives that are *otherwise worth living*, even if they personally disagree. The reason why this counterintuitiveness matters, is because it also practically implies that individuals can be owed compensation even if they do not *feel the need for such compensation*. To this effect, it is unclear how high the threshold can go prior to becoming absurd or excessive. If the threshold is set at an uncontroversially low level, it appears to largely resemble the current ‘0-utility’ baseline that is already used in the welfare comparisons, as well as failing to explain away the injustices perpetuated to descendants of victims

Secondly, it is unclear whose interests could be used or appealed to in determining the constituents of such a threshold. To the extent that the threshold is modeled after a hypothetical ‘ideal’ person, it is unclear why the *ideal person* should have any bearing on what compensation the *non-ideal individual* is owed in actuality; more importantly, it empirically appears that there exist many descendants of victims of historical injustice that are clearly not living beneath such thresholds in the Status Quo. Yet the intuition remains that they deserve some form of compensation – this suggests that this thresholdist conception of the Uncomparative Wrong Account is explanatorily either severely limited, or fails to correspond at all to the ground intuitions that underpin reparative justice.

THE DE DICTO WRONG ACCOUNT

An alternative account of wrongness is the Parfitian view that person P commits some wrong upon P’s failing to satisfy some principle X. Such a principle X may constitute the principle of utilitarianism – i.e. maximising the aggregate utility given a particular set of options or choices available to an individual. It may perhaps be tempting at this point to stray into a general discussion of the Non-Identity Problem and its various manifestations. However, it is imperative to refocus the discussion on the question of compensation. The Parfitian account (1984) alone seems to be largely irrelevant to the problem of compensation outlined above – perhaps those who committed a historical injustice also committed a

wrong by causing a flawed existence – but why should the *impersonal wrong* of the act translate into a justifiable claim to compensation by the individual living the flawed life?

Abandoning Parfit's impersonal claim but retaining the view that the claim to compensation is grounded – at least partially – in the actions of the original wrongdoer, the De Dicto Wrong Account offers a viable solution to the Non-Identity challenge. It is worth introducing here the following metaphysical concepts: *de dicto* identity refers to two or more objects sharing the same *description*; *de re* identity refers to two or more objects sharing the same *numerical identity*. For instance, the statement 'The 45th President of the USA could have been someone other than Donald Trump.' is possibly true on a *de dicto* reading (after all, Hillary Clinton could have won the election), but a contradiction on a *de re* reading (Donald Trump tautologically could not have been not himself, by the law of identity). Descriptions are *general* – they can map onto different individuals in different possible worlds, assuming a Lewisian framework (Lewis, 1973). The *particular* refers to a specific counterpart that maps onto a certain *general description* in a possible world (e.g. in this world, Donald Trump is the *particular* of 'The Successor of Barrack Obama', which is a *general*).

Applying the concepts outlined above, descendants in this world are *particulars* who are owed compensation because they are *wronged* by the *harm* to their *generals* (Hare, 2007). Note here that the particular individual is wronged (non-comparative), whilst their de dicto general is harmed (comparative – comparing the particulars across different worlds). This abstract claim could be grounded using the following example: *Mother's Only Child* – in World X, some injustice occurred to Mother, leaving her only child (Child X) born into a life of misery and suffering; in World Y, no injustice occurred, which enabled her child (Child Y) to be born into a life of happiness and fulfillment. Note that the Non-Identity Problem suggests that Child X has no claim to compensation even though Child Y leads a life of comparative happiness and fulfillment. On the contrary, this essay posits that this conclusion is untrue. By holding the Child X and Child Y as *de re different* (different in particular) but *de dicto identical* (identical in their general description as the 'Child of the Mother', where 'the Mother' is used indexically), this argument identifies a comparative harm incurred to the *general* – the 'Child of the Mother' is worse off in World X as compared to in world Y. As such, the historical injustice has *wronged* the actual 'Child of the Mother' in this particular possible world (Hare, 2007).

THREE OBJECTIONS TO THE DE DICTO ACCOUNT

There are three immediate objections to this account. Firstly, the *Queerness* objection: the comparative ‘harm’ to the ‘Child of the Mother’ appears to be fundamentally queer – it is unclear how a non-existent, general description could be harmed by a particular decision. Secondly, the *Irrelevance* objection: why should the comparative harm to the ‘Child of the Mother’ have any moral bearing on the particular? Thirdly, the *No Ground* objection, to the extent that the particular is affected, it is unclear why this particular form of influence constitutes a justifiable ground for compensation. Resolving these three objections could substantially strengthen the account of De Dicto wrongness, as raised by Reiman (2007) and Hare (2007).

Consider first the *Queerness* objection. This objection challenges the validity of the claim that a particular description can be harmed, in the sense that the description is non-actual and does not exist in actuality. At best, the description refers to a collection of potential persons connected by a common description – yet it is unclear, if none of these persons is harmed by the particular action, why the description is therefore comparatively harmed. This objection can be resolved through a metaphysical clarification – the view that only actual or existent objects can be harmed is often grounded in the metaphysical intuition that harm must necessarily manifest in some concrete, visualisable physical processes; however, if we suspend the assumption that harm must require some active causal, physical process and adopt a purely comparative account of harm (cf. earlier discussions), it is not implausible to suggest that a description can be ‘harmed’ in so far as there exists a discrepancy between its particular in *this possible world* and an alternative particular in *another possible world*. Therefore, there is nothing inherently queer about harming a description. More importantly, this clarification also highlights a particular strength of the De Dicto account – it enables the preservation of the intuition that there exists some form of harm to some ‘individuals’ (i.e. those who conform to the *general descriptions*) without requiring the demonstration of harm to the particular individual who exists in actuality. In other words, the harm to a description is itself a sufficient wrongness-causing feature for the particular manifestation of the description that in fact exists.

The *Irrelevance* objection echoes Parfit’s concern, that there exists an apparent explanatory gap between the view that the initial act of injustice makes things worse for the ‘Child of the Mother’, and that the injustice enactor has actively wronged her child *in this world*. Parfit

posits that there exists no “familiar moral principle”⁶ that supports this hypothetical linkage. A potential response may draw upon the idea of *opportunity cost* – consider, for instance, the *Recommendation Letter*. Bob could write a recommendation letter for Ruth, such that Ruth receives a job offer that pays her 100,000GBP per year; alternatively, Bob could write her a letter that allows her to be offered thrice the salary. Bob chooses the former for no particular reason, causing Ruth to be deprived of the *opportunity* of earning 300,000GBP. It appears that Ruth-in-this-world earns merely 100,000GBP per year, whilst Ruth-in-a-possible-world earns 300,000GBP per year. Given the disparity between the two scenarios, and given that Bob’s choosing otherwise would not have incurred a substantial cost upon him (if at all), it appears that Bob may owe Ruth 200,000GBP. Yet the *Recommendation Letter* thought experiment is apparently flawed on two levels: i) it is unclear if Bob in fact owes Ruth anything, given that any benefit accrued to her is the outcome of a voluntary, supererogatory act from him; ii) the *Recommendation Letter*, unlike the Non-Identity Problem, is not identity-changing – it is possible to identify the very same Ruth across both ‘possible worlds’, whilst it is impossible to do so for cases where the Non-Identity Problem arises. As such, a direct response to the *Irrelevance* objection seems untenable (Parfit, 1984).

It may be worth moving onto the *No Ground* objection prior to returning to the *Irrelevance* objection – as will be illustrated, the solution to the former also offers an indirect yet sufficient response to the latter. Given that it has been shown that the general ‘Child of the Mother’ is *harmed* by the initial injustice, it may be posited that there exists a *prima facie* claim to mitigation by the ‘Child of the Mother’. As noted above, however, it is *de facto* impossible to compensate a general description. To the extent that there exists a duty to compensate the ‘Child of the Mother’, and given that the ‘Child of the Mother’ does not exist as a general description, the *next best way to discharge such a duty* is to compensate the particular individual who is most proximately related to the general. In less jargonistic language, the child born in this world is the most intimately connected to the description ‘Child of the Mother’, as compared to any other person that currently exists. As such, there exists a ground for compensation (*contra* the objection); thus, the particular in this world is owed compensation even in spite of the *Irrelevance* objection above. More intuitively, this argument reflects the general intuition that in cases where direct compensation is impossible, compensating the most proximately related agent may be normatively required (Lewis, 1976).

⁶ pg. 359, Parfit, 1984; pg. 530-533, Wasserman, 2008

What accounts for the unique relation that renders the child in this world most intimately connected to the ‘Child of the Mother’? A viable explanation for the connection is that the description ‘Child of the Mother’ can be decomposed into a conjunction of all of its particulars existing across different possible worlds – i.e. ‘Child of Mother’ = {Child-in-this-world, Child-in-W1, Child-in-W2, ... etc.}. As such, the child in this world (the particular) *constitutes* a part of the ‘Child of Mother’ (the general). A duty owed towards the general is discharged through compensating a component of it. Whilst this solution apparently seems rather counterintuitive, it has two distinct analytical advantages: i) firstly, it allows for the preservation of the intuition that the compensation is targeted towards *some wrong* carried out towards *some individual (the general)*; ii) secondly, it achieves i) without the need to prove either that an individual is wronged according to some arbitrary or under-substantiated absolutist metric, or that a particular individual is *comparatively worse off* (cf. the original Non-Identity Problem).

TWO FURTHER OBJECTIONS

A further objection is that this account appears to ‘prove too much’. If the ground for compensation requires merely that the general description is *worse off* in this world, it appears to suggest that whenever the welfare of the general description is *not maximised by an action in this world* (i.e. there exists some possible world where the particular is better off than the particular in this world), the particular in this world is owed compensation. If this were the case, it would suggest that so long as there exists some possible world where some manifestation of a description is better off, the particular manifestation in this world should be compensated – e.g. if there exists a ‘Child of the Mother’ who leads an incredibly extravagant life in another possible world, then the child in this world should be compensated for the difference between his life in this world and his extravagant life in that optimal world – even if there was no act of ‘injustice’ that precipitated his existence. Call this the *Excessive* objection.

The *Excessive* objection is misguided, in that it can be resolved through some basic clarifications of the compensative question. The question of compensation seeks to identify what follows from a *particular voluntary action* – i.e. what can be *directly attributed to that particular action*. In the original *Mother’s Only Child* case, the injustice to the mother *directly influences* the welfare of the ‘Mother’s Child’ – in

that had the act of injustice not occurred, Child Y – with higher levels of welfare as compared to Child X – would have been born. Therefore, the amount for which the child ought to be compensated (on grounds of reparative justice, at least) purely corresponds to the disparity between Worlds X and Y (worlds with or without the act of injustice). On a more general note, the amount for which the descendant of the victim can claim compensation is constrained by the choice options available to the victimiser – to illustrate this, it would be unreasonable to posit that the Mother’s Child must be compensated for the initial injustice through covering the difference between the extravagant life (led in some, ‘maximal’ possible world) and the actual life he lives; instead, the comparative should simply be one where the initial injustice is not committed. Any further compensation on other grounds (e.g. grounds of luck egalitarianism) falls outside the parameters of compensation, and ought to be ignored. Furthermore, note that the question of compensation is merely but one amongst many desiderata of justice; there may exist considerations of reasonability and other side constraints that restrict the extent to which compensation can justifiably be sought and claimed from individuals in real life. As such, it is untrue that this account ‘proves too much’ with respect to how much compensation can be claimed.

The second objection is that this account is deeply inconclusive with regards to certain cases. Consider the *Conquest*. In World X, a historical conquest left an originally very affluent country heavily plundered and savaged; in order to replenish the national wealth, the Conquistador mandates that every family gives birth to at least four children. Some couple A and B hence give birth to C, D, E, and F. In World Y (where the conquest does not occur), the couple would only have one child – and name them C*. The De Dicto account seems to suggest that each of C, D, E, and F are entitled to the welfare C* would otherwise have had – but this almost appears as if C, D, E, and F are *over-compensated*, given that the counterfactual only contains one C*. Call this the *Inconclusiveness* objection.

The first rejoinder to this objection is to note that it is possible to restrict the scope of the de dicto identity by further specifying it as ‘The First Child of A and B’, such that only C would be compensated. This avoids the concern of over-compensation, but also illustrates a further point – it is always possible to modify the direct description by specifying it along principles such as identical quantity (cf. Parfit’s Same Number cases), proximate qualities, and relevant contexts. An alternative rejoinder is to bite the bullet, and accept that C, D, E, and F ought to all be compensated – note that it is merely a matter of moral luck that

D, E, and F were not born in place of C as the first child of A and B. A luck-egalitarian modification could support the view that all of them are equally entitled to the compensation that they *would have received*, had it not been for the *element of luck*. In any case, one thing remains constant across both replies: at least one of the descendants of the victims has a claim to mitigation.

CONCLUSION

In conclusion, this article addresses the Non-Identity Objection to counterfactual-based accounts of compensation for historical injustice. Accounts that seek to highlight uncomparative wrongness have been rejected, on the grounds that they either fail to justify the validity of their postulated wrongness-making features (Woodward's rights-based account) (Woodward, 1986), or that they are unable to reconcile their thresholdist metric (Steinbock, 2009) with common intuitions concerning what lives are worth living. The alternative proposal of De Dicto Wrongness (Hare, 2007) is shown to be consistent and advantageous, in that it is able to retain *somewhat* the person-affecting intuition without assuming the burden of demonstrating comparative harm to an actual person. The counterfactual account may be problematic for various reasons on a wider scale – but at the very least, the Non-Identity Problem has been shown to be much less of a valid objection than it may initially seem.

BIBLIOGRAPHY

- Butt, Daniel. "Inheriting rights to reparation: compensatory justice and the passage of time", *Ethical Perspectives* 20, 2 (2013), 245-269
- Lewis, David, 1973. *Counterfactuals*, Oxford: Blackwell Publishers
- Hare, Caspar, "Voices from Another World: Must We Respect the Interests of People Who Do Not, and Will Never, Exist?", *Ethics*, 117 (3) (2007): 498-523
- Nozick, Robert, 1974. *Anarchy, State, and Utopia*. Basic Books.
- Parfit, Derek, 1984. *Reasons and Persons*. Oxford University Press.
- Reiman, Jeffrey, "Being Fair to Future People: The Non-Identity Problem in the Original Position" (2007)
- Roberts, Melinda, "The Nonidentity Fallacy: Harm, Probability and Another Look at Parfit's Depletion Example," *Utilitas*, 9 (2007), 267-311
- Waldron, Jeremy. "Superseding Historic Injustice." *Ethics* 103, no. 1 (1992): 4-28.
- Woodward, James. "The Non-Identity Problem," *Ethics*, 96 (1986): 804-31.
- Steinbock, Bonnie. "Wrongful Life and Procreative Decisions" (2009)

God's Dream:
Implications of Idealistic
Theism for Non-Theists

Max DuBoff
Rutgers University

ABSTRACT:

Idealism has remained an attractive theory throughout the history of philosophy, particularly by those seeking to justify theism. The nineteenth-century Hasidic master Rabbi Mordechai Yosef Leiner, the Ishbitza Rebbe, envisions the world as comparable to a dream in the mind of God. Although multiple contemporary thinkers have examined the merits of his theology, I examine it from a non-theistic perspective, arguing that the Ishbitza Rebbe's idealism, a form of *idealistic theism*, should be more plausible to non-theists than traditional theistic accounts of the world are. I address various common objections to theism and explain how idealistic theism helps solve them, and I explain why common arguments against idealism do not hinder idealistic theism. I argue that idealistic theism has far-reaching implications on how we think about metaphysics, philosophy of religion, and more.

I: INTRODUCTION

We humans care a rather great deal about our reality. We yearn to be real, to matter in the broadest sense possible, to find meaning, and to be loved; theism's enduring popularity should thus come as no surprise. Theism can take many different forms, however, including some that seem to contradict theists' and non-theists' deepest assumptions about the world. One such account, offered by the Ishbitza Rebbe—the nineteenth-century Hasidic master Rabbi Mordechai Yosef Leiner (hereafter “the Ishbitza”)—implies that humans are like a dream or story in the mind of God.¹ The Ishbitza has prompted considerable discussion recently: Rabbi Herzl Hefter seeks to reclaim the ideas of the Ishbitza for mainstream Orthodox Judaism by showing that, despite ostensible conflicts, “[Rabbi] Mordechai Yosef is entirely loyal to traditional Jewish religious sensibilities”;² while Samuel Lebens, responding to Hefter, identifies multiple challenges to the Ishbitza's views and argues that “the Ishbitza is able to wield his one metaphysical assumption to overcome these problems.”³ Neither Hefter nor Lebens, however, addresses the impact of the Ishbitza's views on non-theists, instead focusing on the profound effects of the views on arguments for theism itself. I seek to show that the non-theist, given assumptions she likely already holds,

¹ Samuel Lebens, “God and His Imaginary Friends: A Hassidic Metaphysics,” *Religious Studies* 51, no. 2 (2015): FirstView Article, 1, http://journals.cambridge.org/abstract_S0034412514000298.

² Herzl Hefter, “In God's Hands: The Religious Phenomenology of R. Mordechai Yosef of Izibica,” *Tradition* 46, no. 1. (2013): 45, http://traditionarchive.org/news/_pdfs/0043-0066.pdf.

³ Lebens, “Imaginary Friends,” 4.

should find the Ishbitza's brand of idealistic theism more likely than standard theistic accounts. Although I realize that she will still consider both accounts wildly implausible, this exploration sheds light on why people do or do not subscribe to theism as well as on how they respond to assertions that contradict their entire worldview.

To measure my success, I will employ the concept of an average non-theist (hereafter NT), who, like most non-theists and theists alike, acknowledges that she cannot be truly certain about the existence of god(s) or lack thereof; she thus must assign credence to all sorts of possibilities, including theistic ones. I will consider myself successful if, based on my arguments, NT is likely to place more credence in the existence of the sort of theological entity⁴ the Ishbitza describes, with the entailed metaphysical assumptions described below, than in the sort of deity described by mainstream Judaism, Christianity, and Islam, which creates or actualizes material objects and is fully omnipotent, among other maximal properties. In Section II I will define idealism and other important terms; in Section III I will lay out the Ishbitza's basic views; in Section IV I will address four of the most common ideological barriers to theism and explain how idealistic theism makes progress toward overcoming each; and in Section V I will defend idealism against potential defeaters that could weaken the conclusions of Section IV.

II: IDEALISM

The term "idealism" is rather slippery, partly due to its informal usage in conjunction with its technical one. Moreover, different sorts of idealism have emerged within philosophy, distorting the technical term. In all cases, however, idealism conceptually or physically prioritizes the mental over the material. For the purposes of this paper, it is only necessary to establish that the Ishbitza's views fit the term. I will thus employ without argument the definition given by Paul Guyer and Rolf-Peter Horstmann of what they call "ontological idealism": the assertion that "something mental (the mind, spirit, reason, will) is the ultimate foundation of all reality, or even exhaustive of reality."⁵ This definition does not require that only mental objects exist, but it does require that what we experience as real somehow relies on the mental. As the major

⁴ For reasons that will become clear later, I wish to leave open the possibility of a godlike entity other than the traditional God of the major monotheistic faiths. Although I will subsequently refer simply to "God" for the sake of convenience, it may be assumed that I intend a relatively broad sense of the term.

⁵ Paul Guyer and Rolf-Peter Horstmann, "Idealism," *The Stanford Encyclopedia of Philosophy* (Fall 2015 Edition), ed. Edward N. Zalta, <https://plato.stanford.edu/archives/fall2015/entries/idealism/>.

monotheistic faiths have done for centuries, I will assume that God is incorporeal and primarily mental.⁶ Thus, if it is shown that God Godself is the bedrock of reality, then idealism and theism will both be correct. I will call this concept *idealistic theism*, and I will call a theistic account that posits something non-mental as the ultimate foundation of reality *standard theism*. Despite the name, standard theism thus describes common accounts, such as that God creates a material world which we inhabit and which is real for us, as well as many views traditionally considered heretical.

III: THE ISHBITZA

Like many other figures of the early Hasidic movement, the Ishbitza set forth esoteric concepts largely rejected in his own generation. Contrary to the overwhelming Jewish tendency to believe in free will, which allows humans to choose to fear and follow God, the Ishbitza asserts that “apparent freedom is an illusion.”⁷ Further, unlike some monotheists who deny free will but value heaven or even union with God, the Ishbitza establishes that “[t]he only reality which is of any consequence is the will of God.”⁸ From the centrality of the will of God and from various metaphors in the *Mei ba-Shiloa* (hereafter *MH*), the Ishbitza’s magnum opus, Hefter derives a “paradoxical two-tiered [sic] reality”⁹ whereby there is “ontological reality to the ‘illusion’ of free will.”¹⁰ To conceptualize that tiered reality, he expands the claim about free will to argue that the whole “God-created illusion possesses ontological reality,”¹¹ effectively refuting the claim that our lack of objective free will makes us meaningless. From a metaphor in the *MH* where the Ishbitza calls the world a dream, Hefter assumes that the dreamer must be God Godself. In explaining the tiered reality, Hefter shows that, according to the Ishbitza, “[a] thought in the mind of God *is* reality.”¹²

Lebens premises his defense of the Ishbitza on Hefter’s key distinction about levels of reality, and he assumes, expanding on the

⁶ See, for instance, the third of Maimonides’ Thirteen Principles of Faith: “Maimonides: The 13 Principles and the Resurrection of the Dead,” Fordham University, accessed December 4, 2017, <https://sourcebooks.fordham.edu/source/rambam13.asp>.

⁷ Morris M. Faienstein, *All Is in the Hands of Heaven: The Teachings of Rabbi Mordecai Joseph Leiner* (Piscataway: Gorgias, 2005), 28.

⁸ Faienstein, *Hands of Heaven*, 29.

⁹ Hefter, “In God’s Hands,” 62.

¹⁰ Hefter, “In God’s Hands,” 61.

¹¹ *Ibid.*

¹² *Ibid.*, 61-62 (emphasis in original).

MH, that “the world is God’s lucid dream.”¹³ Elsewhere, he likens God to an author and an artisan.¹⁴ He thus allows God control and agency over the world but does not necessarily commit his idealistic theist to traditional views about God’s power and workings.¹⁵ Lebens importantly builds on Hefter’s two-tiered reality by distinguishing between two sorts of truth: *truth_E*, i.e. “truth-relative-to-God’s-dream-or-story”;¹⁶ and *truth-simpliciter*, i.e. objective truth at the most fundamental level of reality. It is true_E, therefore, that we have free will, and false-simpliciter. Although this distinction arises in Lebens’ treatment of freedom, which is not necessarily a barrier or concern for NT, the distinction underlies many other arguments that Lebens employs. Between Hefter’s and Lebens’ interpretations, therefore, the Ishbitza can be understood as expressing a coherent idealistic theistic ideology.

IV: IDEALISTIC THEISM

Having established a firm background for idealistic theism, I will now address four common concerns that NT might have about standard theistic accounts and show that they are less problematic for idealistic theism. These concerns are: 1) the conceptual hurdle of a greatest possible or transcendent being; 2) the ostensible complexity of a theistic account of the world; 3) the existence of seemingly large degrees and amounts of evil; and 4) conflict between our moral intuitions and theistic grounding of morality. I do not claim that these four concerns constitute an exhaustive list of reasons for non-theism, particularly since such reasons vary greatly. I do claim, however, that these four are widely held and legitimate reasons for non-theism, such that I will achieve my goal of showing idealistic theism more broadly plausible than standard theism if I do so for these four concerns.

The first concern posits that a greatest possible or other sort of transcendent being, common in standard theism, simply is not very plausible. We have discovered myriad natural laws that govern our world, and none of them seem to allow for the sort of God espoused by the major monotheisms. Although God could exist anyway, either based on laws we have not discovered or in conjunction with all the laws we have, there is no reason to think God does (aside from tradition, of course). As

¹³ Lebens, “Imaginary Friends,” 2.

¹⁴ *Ibid.*, 10.

¹⁵ I do not agree with Lebens’ assertion (p. 2) that he is simply following Hefter’s interpretation of the Ishbitza, but exact fidelity to the Ishbitza’s worldview is not important for the purposes of this paper.

¹⁶ Lebens, “Imaginary Friends,” 5.

physicist Sean Carroll, for example, has declared, “The universe runs by purely naturalistic principles.”¹⁷ While plausible responses to this concern certainly exist, they often fail to convince NT because we have little or no epistemic access to God. Idealistic theism, however, can confront this concern by providing analogical examples not open to standard theism. First, Lebens cashes out his metaphor of God as an author by noting that authors often claim that their characters or stories take on lives of their own, reducing the authors’ control. Lebens thus imagines that perhaps “the *perfect* author must be sensitive to the ideas and images that passively ‘arise.’”¹⁸ Referencing the dream metaphor, Lebens also notes that “even in the most lucid of dreams, the dreamer is not fully in control.”¹⁹ God in this account appears omnipotent insofar as is possible but not to the extent of overriding natural laws, assuaging some of NT’s concerns about plausibility.

Moving away from the Ishbitza, another argument with a surprising conclusion, Nick Bostrom’s Simulation Hypothesis, can establish plausibility for God in an idealistic framework. Bostrom convincingly argues that there is some chance, albeit small, that we are living in a computerized “ancestor simulation” created by a technologically advanced posthuman society.²⁰ Given the popularity of this argument, NT likely already believes it or at least understands the concept. If we are living in a simulation, our relationship to those simulating us exactly mirrors our relationship to God in idealistic theism: it is true_s (i.e. relative to the simulation) that reality exists as we think, but false-simpliciter. Further, although Bostrom only discusses intentional ancestor simulations, it seems plausible that a posthuman could accidentally press a button that creates a simulation and thus us. If NT accepts these situations as possibilities, it should not be difficult for her to imagine God in the same role.²¹ A standard theistic account cannot employ this helpful analogy, however, since such an account does not involve a tiered reality.

Another example similar to idealistic theism occurs in Isaac Asimov’s short story “The Last Question,” wherein Asimov imagines a

¹⁷ Sean Carroll and William Lane Craig, “God and Cosmology,” filmed 2014 at Greer-Heard Point Counterpoint Forum, Video, 2:26:35, https://www.youtube.com/watch?v=07QUPuZg05I&feature=youtu.be&list=FLRhV1rWIp_m_U19bBm_2RXw.

¹⁸ Lebens, “Imaginary Friends,” 10 (emphasis in original). Although Hefter implies that the Ishbitza thinks God’s delusion must be intentional (e.g. Hefter, “In God’s Hands,” 46: “He [God] drew a veil over our eyes...in order to facilitate Divine worship”), Lebens does not mention such a distinction and implies that the delusion might be intentional or unintentional; idealistic theism as I wish to present it follows Lebens’ approach on this point, regardless of the Ishbitza’s original opinions.

¹⁹ Lebens, “Imaginary Friends,” 10.

²⁰ Nick Bostrom, “Are You Living in a Computer Simulation?” *Philosophical Quarterly* 53, no. 211 (2003): 243-255, <https://www.simulation-argument.com/simulation.html>.

²¹ And one running a simulation might be indistinguishable from, and thus functionally equivalent to, God.

future series of humans increasingly mental rather than physical along with computers that become ever more powerful.²² The imagined later stages of Man (likely similar to what Bostrom conceives of as posthuman) and computer are almost entirely mental; minds roam freely, separate from their bodies, and the AC (Automatic Computer) cannot “release his consciousness”²³ (implying that it is conscious) until it answers the eponymous last question. These mental beings resemble us as we are in an idealistic theistic account, and the conscious computer resembles (or, according to Asimov, might actually be) God. Based on examples of imaginary computers and the like, NT should thus allow for the existence of the sort of God of idealistic theism. These examples do not explain the existence of the relatively standard God that the Ishbitza wants to posit, but they do illustrate the mechanisms of idealistic theism and render idealistic theism easier to conceive of and accordingly more plausible than standard theism.

The second concern for theism relates to simplicity. Many argue that theses which require us to posit fewer entities are *ceteris paribus* more deserving of belief than those which require us to posit more, appealing to Occam’s Razor. Since God is a rather significant entity and scientific theories are, according to many, enough to explain the phenomena we observe and predict, accounts that do not posit God are seemingly simpler than and more preferable to accounts with God. Thus Carroll, explaining his preference for non-theism, argues: “If I have two possible methods that fit what we observe about the universe, and one of them has less stuff, less ideas, is more self-contained, is more rigid and well-defined than the other one, I’m going to prefer that one.”²⁴

Regardless of whether simplicity is a worthwhile criterion for evaluating accounts, however, idealistic theism is simpler than standard theism. Alan Baker formulates a weak version of Occam’s Razor as follows:

(OR₁) Other things being equal, if T_1 is more ontologically parsimonious than T_2 then it is rational to prefer T_1 to T_2 .²⁵

Although ontological parsimony is a rather complex concept, “a sufficient condition for T_1 being more parsimonious than T_2 is for the

²² Isaac Asimov, “The Last Question,” in *Nine Tomorrows* (Bantam, 1960), <https://www.physics.princeton.edu/ph115/LQ.pdf>.

²³ *ibid.*

²⁴ Sean Carroll, “God and Cosmology,” 2:26:40.

²⁵ Alan Baker, “Simplicity,” *The Stanford Encyclopedia of Philosophy* (Winter 2016 Edition), ed. Edward N. Zalta, <https://plato.stanford.edu/archives/win2016/entries/simplicity/>.

ontological commitments of T_1 to be a proper subset of those of T_2 ,²⁶ where “a theory, T , is ontologically committed to F s if and only if T entails that F s exist.”²⁷ Idealistic theism is ontologically committed, on the most fundamental (i.e. truth-simpliciter) level, only to God; standard theism is also committed to people and some kind of non-mental world, in addition to God. The ontological commitments of idealistic theism are thus a proper subset of the ontological commitments of standard theism, and idealistic theism is accordingly more ontologically parsimonious than standard theism. Based on (OR₁), therefore, it is rational to prefer idealistic theism to standard theism.

This argument, which attempts to classify physical objects as ontological commitments and mental objects as part of their originating minds, fails if objects that are only real_E count as ontological commitments, but because they are not real in some sense (real-simpliciter) even if they are real in another, they appear to be less of an ontological commitment than physical objects that are real in every sense. And although this argument can seemingly be applied to any non-idealistic hypothesis, most of us consider idealism rather unlikely and would not agree that other things are equal (as stipulated by (OR₁)), except in cases where the only variable is idealism, as when comparing idealistic theism and standard theism. Taking a different tack, the standard theist could argue that God’s creations are not separate enough from God to count as ideological commitments, but then the standard theist seemingly commits himself to pantheism, a rather unpopular theory, which posits that all things are part of God. Idealistic theism seems no simpler than pantheism, but based on simplicity NT should prefer idealistic theism to other standard theistic accounts.

The third concern, the problem of evil, involves the most serious challenge to any theism with a purportedly good God. Evil hardly needs an introduction; wherever we look, we can tell that lots of bad things happen to people of varying moral quality at every instant. Perhaps many instances of evil can be explained away by universalistic concerns, NT might concede, but is every instance of evil really necessary, and if not why does God allow such evil? On top of normal evil, philosophical discourse often focuses on grave evils, e.g. slow and cruel dismemberment by a terrorist, which are particularly egregious and do not have an obvious explanation. Theists often try to explain evil with an appeal to free will, but the idealistic theist maintains that it is false-simpliciter that humans

²⁶ *ibid.*

²⁷ *ibid.*

have free will, and the Ishbitza seemed to believe strongly that “man ultimately has no independence.”²⁸

Lebens proposes some possible theodicies for an idealistic theist which were mentioned above in response to the first concern: God might be a perfect creator, for whom it is actually best to let some ideas arise in God’s mind without interference; and God, despite being like a lucid dreamer, still cannot control every aspect of the metaphorical dream.²⁹ By assigning God less control and qualifying omnipotence and goodness, the idealistic theist allows God to permit at least some evil without reproach.³⁰ Lebens also proposes another theodicy via the metaphor of God as a storyteller, whereby the storyteller has no moral responsibility to her characters.³¹ It is false-simpliciter, in this theodicy, that God is morally responsible for the evil we and others have experienced. Lebens notes, for example, that we should not blame Anthony Burgess for the horrible evil that occurs in his novel *A Clockwork Orange*, and that perhaps we should praise him for writing such a moving work.³² Nevertheless, God’s moral responsibility or lack thereof is relatively little consolation when we chafe at the actual evil we and our fellow people experience. As long as we are conscious—and the Ishbitza and his interpreters seem never to question that assumption—we are radically different from the characters of *A Clockwork Orange* and of our dreams as well.³³ Lebens addresses this objection by positioning us as unimportant and arguing that “inside Burgess’s stories, it is true to say that his characters are sentient persons.”³⁴ Consciousness, however, seems objective in a way that free will is not; if I am conscious, that fact is true_E, but it is also true-simpliciter that there is a mental object originating from God with all my properties that possesses consciousness. Lebens’ second theodicy thus fails and God is morally responsible for the evil of our world, except insofar as God cannot prevent it, although it seems unlikely that God is incapable of preventing any of the great evil we experience, assuming (as the Ishbitza and his interpreters do) that God has any power at all.

The idealistic theist still has one trick remaining to him, however; perhaps the illusion of the tiered reality extends to evil in addition to

²⁸ Fajerstein, “Hands of Heaven,” 28.

²⁹ Lebens, “Imaginary Friends,” 10.

³⁰ Although this tactic precludes commitment to a strong notion of omnipotence, NT clearly will not mind the change, and the core of idealistic theism is unchanged by this tactic.

³¹ Lebens, “Imaginary Friends,” 10.

³² *ibid.*, 10–11.

³³ One could, of course, extend consciousness to characters in novels and our dreams, but there is no reason to think that is the case, and forcefully Lebens rejects this notion (p. 12). A major God-making feature for the idealistic theist seems to be the ability to create consciousness for mental objects.

³⁴ Lebens, “Imaginary Friends,” 11.

free will. Evil possesses ontological reality in our level of reality, but it is of a lesser sort than is unequivocally real evil. By seeking the will of God, one might even stop seeing evil. Although God might allow evil, however, perhaps even large quantities of it, the centrality of God's will shifts focus away from moral responsibility. The radical claim that "God's infinite will is the only true will"³⁵ is thus an ideological barrier for NT to idealistic theism but also a chance to worry less about to what extent we should blame God for evil. In this version of idealistic theism, evil just *is*, but it is overshadowed by God's will, a greater force in the grand scheme of the world. Unlike Lebens, who claims that the Ishbitza can use all traditional theodicies based on what is true_E, I have shut off some theodicies to the idealistic theist, those based on free will and those that deny us the moral standing of conscious beings. All other theodicies remain open, though they likely mean little to NT. I have presented NT with the opportunity to recognize that God might not want to or be able to prevent some evil as well as the opportunity to shift her focus away from God's potential moral responsibility. Both of these opportunities, which are only open for idealistic theism, make idealistic theism more likely to NT than standard theism.

The fourth concern involves the worry that, if there is a God and ethics is based on what God desires or considers correct, common-sense ethical notions that we hold dear might turn out to be false. Plato most succinctly sets forth this problem in the *Euthyphro* when Socrates puzzles thus: "is the holy loved by the gods because it is holy? Or is it holy because it is loved by the gods?"³⁶ Most theistic accounts appear committed to a sort of divine command theory of ethics, and idealistic theism, with its focus on the will of God, is no exception. Nevertheless, idealistic theism does not require that God's will be the primary ethical force; the Ishbitza predicts in the *MH* that eventually "Israel [the Jewish people] will divine God's will only from the depths of their hearts."³⁷ Ethics comes from us, though in some sense it thus comes from God, since "[a]ll human actions are...an intrinsic part of God's plan for creation."³⁸ Nevertheless, the relevant moral intuitions and reasoning come from us, alleviating NT's worry, and only come from God in a secondary sense. NT can even take heart in the realization that, by the lights of idealistic theism, we are more likely to reach ethical truth (both truth_E and truth-simpliciter) than if non-theism or standard theism were

³⁵ Hefter, "In God's Hands," 47.

³⁶ Plato, "Euthyphro," in *The Dialogues of Plato: Volume I*, trans. R. E. Allen, 41-58 (New Haven: Yale University Press, 1984), 50.

³⁷ Hefter, "In God's Hands," 48.

³⁸ *ibid.*

correct. Although Hefter attributes to the Ishbitza “a potentially dangerous doctrine of radical Divine immanence which at times justifies antinomian behavior,”³⁹ that behavior stems from our own efforts to seek God and God’s will, not from impersonal laws God has given.⁴⁰ In any standard theistic account that involves substantial libertarian free will (i.e. most of them), God must leave at least some decisions indeterminate; in many other accounts, ethics is secondary to one’s relationship with God. The idealistic theist, on the other hand, can genuinely reach ethical principles from his own ideas but *through* God’s will.

V: DEFENDING IDEALISM

I have thus shown that NT should consider idealistic theism more likely than standard theism, largely based on assumptions that NT already holds or is likely to hold. Nevertheless, that conclusion only applies if NT can seriously give credence to idealism as a concept. I will accordingly discuss three concerns that NT might have about idealism and show that they are negligible: 1) the strangeness of idealism; 2) the existence of consciousness in an idealistic account; and 3) the possibility of an infinite regress in idealistic theism.

The strangeness of idealism is probably the biggest concern for NT; idealism runs contrary to everything we are taught and experience. If our reality is composed solely of mental objects, why does everything seem so *real*? The prominent ontological idealist George Berkeley argues against this intuitive objection via the character Philonous in his *Three Dialogues Between Hylas and Philonous*. Philonous states that, regarding sensible qualities, “there can be no *substratum* of those qualities but spirit,” allowing them to exist “as a thing perceived in that which perceives it.”⁴¹ Berkeley avers that mental objects and qualities are quite real and are certainly perceived, an idea the Ishbitza also wishes to maintain. Further, Berkeley maintains that a quality, and indeed what we would consider an object, can constantly exist because it “is truly known and comprehended by...the infinite mind of God.”⁴² NT need not accept Berkeley’s strong conclusion about God’s limitless perception to agree with his preliminary conclusions about what makes something real, and

³⁹ *ibid.*, 50.

⁴⁰ Indeed, Hefter quotes the Ishbitza’s assertion that “this approach may necessitate taking action contrary to the Halakhah [Jewish law]” (49).

⁴¹ George Berkeley, *Principles of Human Knowledge and Three Dialogues* (Oxford: Oxford University Press, 2009), 180 (emphasis in original).

⁴² Berkeley, *Three Dialogues*, 177.

everything we perceive is quite real by Berkeley's definition regardless of any material existence it has.

Radical skepticism decreases the strangeness of idealism in general. René Descartes, for instance, convincingly argues that "there are no certain indications by which we may clearly distinguish wakefulness from sleep."⁴³ If that is the case, we could be radically deceived about the nature of our existence, and further, we might not be able to tell the difference between God's creations and God's dreams, reinforcing the possibility of idealistic theism. Descartes also imagines "some evil genius not less powerful than deceitful, has employed his whole energies in deceiving me,"⁴⁴ and this skeptical hypothesis is even more difficult to disprove than the previous one. We might also be deceived about our reality by existing in the sort of ancestor simulation that Bostrom describes. As noted previously, NT likely finds Bostrom's thesis plausible, and idealism, which posits that we are mental objects, is a small jump from Bostrom's notion that we are computer-simulated objects. Strangeness is thus only a *prima facie* objection to idealism, and it does not stand up when confronted with other plausible theories that deny that the reality we experience is primarily material.

A second concern for idealism involves consciousness, viz. how we can possess it if we are mental objects. Consciousness does not rely on the physical or mental, however, and can arise the same way it would if we were material: via God, natural processes, etc. Bostrom assumes the concept of *substrate-independence*: "[p]rovided a system implements the right sort of computational structures and processes, it can be associated with conscious experiences."⁴⁵ Although he notes that this theory is controversial, he also acknowledges that he only needs as a premise that whatever gives rise to consciousness is suitably replicated in a simulation; idealism only requires the same, that an appropriate cause of consciousness apply in an idealistic world, and both physical and non-physical potential causes of consciousness still do apply.

A third concern for idealism involves the possibility of an infinite regress of mental realities. If I am a mental being originating from one God, that God could be a mental being originating from another God, etc. An infinite regress does not change what is true_E, although it might require the addition of more levels to the tiered reality. Further, this

⁴³ René Descartes, *Discourse on Method and Meditations on First Philosophy*, ed. David Weissman (New Haven: Yale University Press, 1996), 60.

⁴⁴ Descartes, *Meditations*, 62.

⁴⁵ Bostrom, "Computer Simulation," II.

concern is not unique to idealism; Bostrom acknowledges that, if there are ancestor simulations, “there may be room for a large number of levels of reality, and the number could be increasing over time.”⁴⁶ For the idealistic theist, it would be disappointing in some sense if there are more than two levels of reality because God would play less of a fundamental role than anticipated; practically, however, NT should not care about how many levels of reality there are because the level she inhabits and the level that immediately underlies it will still fulfill the roles of truth_E and truth-simpliciter, respectively. The concern of a regress is thus irrelevant.

VI: CONCLUSION

The Ishbitza proposes a rather provocative way to think about our ontological status, as a dream in the mind of God. Building on the work of Rabbi Herzl Hefter and Samuel Lebens, which examined consequences of the Ishbitza’s thought for theists, I have sought to demonstrate that an average non-theist should assign more credence to an idealistic theistic account of the world than to a standard theistic account. I examined various reasons why a non-theist might hold that view and argued for each that idealistic theism comes closer to overcoming the reason than standard theism does. Although NT is extraordinarily unlikely to become a theist based on idealistic theism, we gain insight from this exploration into the justification for claims skeptical of what appears to be a material reality. Custom unsurprisingly plays a large role in our selection of accounts of the world around us, as it does in our religious beliefs. Fascinatingly, theism can seem so alien to the non-theist that an account that posits multiple seemingly implausible concepts is more likely than one that posits a single implausible concept. Despite considerable work on the topic, there is much more room to plumb the depths of the philosophy behind religious motivation, particularly in today’s rapidly changing religious landscape, and exploration of how religious beliefs (or lack thereof) fit in with other beliefs is a necessary first step.

⁴⁶ Bostrom, “Computer Simulation,” II.

BIBLIOGRAPHY

- Asimov, Isaac. "The Last Question." In *Nine Tomorrows*. Bantam, 1960. <https://www.physics.princeton.edu/ph115/LQ.pdf>.
- Baker, Alan. "Simplicity." *The Stanford Encyclopedia of Philosophy* (Winter 2016 Edition), edited by Edward N. Zalta. <https://plato.stanford.edu/archives/win2016/entries/simplicity/>.
- Berkeley, George. *Principles of Human Knowledge and Three Dialogues*. Oxford: Oxford University Press, 2009.
- Bostrom, Nick. "Are You Living in a Computer Simulation?" *Philosophical Quarterly* 53, no. 211 (2003): 243-255. <https://www.simulation-argument.com/simulation.html>.
- Carroll, Sean, and Craig, William Lane. "God and Cosmology." Filmed 2014 at Greer-Heard Point Counterpoint Forum. Video, 2:26:35. https://www.youtube.com/watch?v=07QUPuZg05I&feature=youtu.be&list=FLRhV1rWIpm_pU19bBm_2RXw.
- Descartes, René. *Discourse on Method and Meditations on First Philosophy*. Edited by David Weissman. New Haven: Yale University Press, 1996.
- Faienstein, Morris M. *All Is in the Hands of Heaven: The Teachings of Rabbi Mordecai Joseph Leiner*. Piscataway: Gorgias, 2005.
- Fordham University. "Maimonides: The 13 Principles and the Resurrection of the Dead." Accessed December 4, 2017. <https://sourcebooks.fordham.edu/source/rambam13.asp>.
- Guyer, Paul and Horstmann, Rolf-Peter. "Idealism." *The Stanford Encyclopedia of Philosophy* (Fall 2015 Edition), edited by Edward N. Zalta. <https://plato.stanford.edu/archives/fall2015/entries/idealism/>.
- Hefter, Herzl. "'In God's Hands': The Religious Phenomenology of R. Mordechai Yosef of Izbica." *Tradition* 46, no. 1. (2013): 43-65. http://traditionarchive.org/news/_pdfs/0043-0066.pdf.
- Lebens, Samuel. "God and His Imaginary Friends: A Hassidic Metaphysics." *Religious Studies* 51, no. 2 (2015): 183-204. FirstView article, 1-22. http://journals.cambridge.org/abstract_S0034412514000298.
- Plato. "Euthyphro." In *The Dialogues of Plato: Volume I*. Translated by R. E. Allen. 41-58. New Haven: Yale University Press, 1984.

In Defense of the
Structural Matching
Thesis

Molly Elder
Johns Hopkins University

ABSTRACT

What is the relationship between the structure of temporal experience and the structure of temporal events? Some philosophers, represented here by Ian Phillips, accept the assertion of the structural matching thesis that the two mimic each other. Other philosophers reject the structural matching thesis and hold that the structure of temporal experience is not necessarily closely connected to the structure of the content of experience, a position which will be represented here by Sebastian Watzl.

I begin with an overview of the structural matching thesis and a discussion of a paper by L. A. Paul, who uses the results of an experiment about perception to successfully argue for a claim concerning temporal experience. Next, I summarize Watzl's argument against the structural matching thesis, which stems from a set of two perception experiments performed by Suchow and Alvarez. Finally, I present two theories which explain how the results of Suchow and Alvarez's experiments can be explained without rejecting the structural matching thesis. The first theory belongs to Phillips, and attributes the experiments' results to crowding out. The second theory is inspired by Michael Flaherty, and attributes the results to the density of experience. In addition, I present suggestions for further experiments which could contribute to the debate over the structural matching thesis.

THE STRUCTURAL MATCHING THESIS

The structural matching thesis proposes that the structure of temporal experience closely matches the structure of the temporal content of experience, or what is in fact happening (Watzl 1011). To illustrate this concept, consider the process of putting a hibiscus teabag in hot water and watching the color (hot pink) leach out of the leaves into the water. Suppose it is your experience that this leaching process occurs quickly, and that the water quickly changes from clear to hot pink. The structural matching thesis postulates that this is equivalent to saying that the property of being hot pink is introduced into your experience of the teabag-mug-water system over a small interval of time. To put it more generally, you experience something as changing quickly only in the case that your experience of the thing is changing quickly (Watzl 1011).

There are two important variations of the structural matching

thesis. The first is layout-to-content, which contends that experience has some temporal structure which informs the content of experience. In the hibiscus tea example, the fact that the water transitions quickly from clear to pink informs the way your experience of the tea gains pinkness over a short span of time. The second variation is content-to-layout, which argues that the content of experience informs the temporal structure. In the tea example, your experience of the tea changes quickly, and this informs your understanding that tea is brewing quickly.

In other words, the structural matching thesis states that the nature of temporal experience is essentially the same as the “temporal process being experienced” (Hoerl 1). Or, again, “experiencing a process requires as much time as the process is experienced to take up,” that is, our experience of a process should not extend significantly past the time when the process ends, nor should it end before the process itself ends (Hoerl 5). It is the “concurrence between the process of experiencing and the experienced process” that the structural matching thesis proposes (Hoerl 5). I emphasize this because it is a slippery concept, and it is easy to understand the structural matching thesis to be claiming more than it actually does. For example, at least in the formulations presented by Hoerl and Watzl, structural matching does not seem to assume the existence of a special now, or of passage to supplement experiential inputs, or that time maintains a constant rate either experientially or in reality. This last non-assumption will become important later in the paper.

As a warm-up before discussion of Watzl’s paper, I will briefly address a paper by L. A. Paul in which she, like Watzl, presents a pair of illusions in order to illuminate the perception of time. Paul is motivated by a desire to undermine the idea that a privileged “now” and objective passage of time are the “only reasonable explanation[s]” for human experience of nowness and passage (Paul 338). Neither nowness nor passage is the subject of this paper, so I will not take the space to explain them. However, understanding in very general terms what Paul’s paper is about is important if one wants to understand why she is concerned with the experiments she references. Her project attempts to show that events do not need to have temporal properties like ‘occurring in the present’ for the experience of events to have those properties; that is, nowness and passage are “distinctly mental” processes, and just because they are experienced does not mean they exist outside of the human consciousness (Paul 340). As part of this project, she discusses an experiment where participants are shown

a screen with a dot alternating between flashing on the right edge of the screen and the left edge. Watching the flashing dot creates the illusion that the dot is actually moving back and forth across the screen, even though objectively it is not. The interpretation of this experience that Paul endorses is that the brain takes these static inputs and presents them to the consciousness as movements (Paul 349). Moreover, when the dot changes color as it switches from the right side of the screen to the left, the participants experience it as not only moving across the screen, but also changing color as it does so (Paul 350). It is particularly relevant to her work that the brain can create the illusion of an animated experience from a sequence of static experiences, in both position and color change. In other words, human visual systems are “capable of producing dynamic visual content” from the most “partial and fragmentary” evidence, like a movie (Dainton 364). Paul claims that these two experiments, where the experience of motion is created from merely presenting a series of static images in sufficiently quick succession, represent the nature of experience writ small. Experience of movement is just “illusory,” a “result of the brain’s need to accommodate the contrasts” between successive states of the world (Paul 352). Experience of flow or passage is an interpretive flourish added as the brain processes inputs, and Paul’s point about the status of passage is made.

It is worth noting that Watzl does not include either of the experiments discussed by Paul in his paper. I believe this is because he did not think they would support his argument against the structural matching thesis. As noted above, structural matching is not explicitly wedded to any specific ideas of nowness or passage. While the experiments Paul uses in her paper support taking a skeptical view of any absolute (as opposed to only mental) conception of nowness/passage, the results do not seem to support a rejection of the idea that the structure of experience mimics the structure of what is being experienced. Structural matching interprets the fact that A is experienced to be changing quickly as equivalent to the experience of A to be changing quickly. There is nothing in this theory that bars the changing experience of A from being a comparison of A’s color/position/size/etc. over successive static “frames.” Proving that a quick progression of similar images can be interpreted by the brain as movements does not represent a successful attack on structural matching. In the case of movement, one would have to show inconsistency between the real speed of A (which could be explained using time-indexed static states) and the speed at which one experiences A moving. This is what Watzl attempts to do, and what most of this paper will be concerned with.

WATZL'S ARGUMENT AGAINST STRUCTURAL MATCHING

While Paul attacks the idea that passage and nowness exist outside of experience, I do not interpret her as attacking the structural matching thesis more generally. In contrast, Watzl criticizes the structural matching thesis explicitly. This section will provide an overview of the experiments he cites, why he finds the results problematic for structural matching, and how he responds to several anticipated counterarguments.

Watzl's paper analyzes a series of experiments performed by Suchow and Alvarez. In the experiment that Watzl particularly emphasizes, participants were asked to focus on the center of a circle of colored dots. Each dot was changing color at a uniform rate, and the ring of dots was either kept stationary or rotated. Under rotation, participants reported that they experienced the dots changing color at a slower rate than under the stationary condition (Watzl 1015). Faster rotation caused the color change to appear slower, a phenomenon which Suchow and Alvarez labeled "silencing." Watzl argues that this result creates a problem for the structural matching hypothesis because the participant's experience of color has the same temporal layout as reality (the participant's experience of a dot's color at time t accurately reflects the color of the dot at time t) but different temporal content (the participant experiences the rate of color change differently in the two states, but objectively the rate of color change is the same). This is presented as evidence against the structural matching thesis because the content of experience seems to present a different picture of reality than the one presented by the layout of experience, and it seems that at most only one of these pictures can be accurate. While the participant's experience of color seems to approximately "keep up" with temporal reality, she experiences an objectively identical property (rate of color change) differently under the two conditions. Since the underlying rate of change remains the same, the participant's experience must not match the temporal content of reality under at least one of the two conditions.

The central claim made by Watzl using Suchow and Alvarez's data is that, when the wheel is rotating, you experience temporal content (your experience of the rate of color change) differently than when the wheel is stationary, but you experience the temporal layout (your experience of the color of each dot at time t) approximately in real time (Watzl 1017). The difference in temporal content is demonstrated in the original experiment, and the accuracy of temporal layout experience has been shown by subsequent experiments. Watzl proceeds to argue that this conundrum represents an irreparable breakdown of the temporal matching thesis.

The first counterargument which Watzl presents attempts to defend the structural matching thesis by arguing that “color information is updated less frequently” under the rotation condition than under the stationary condition (Watzl 1020). Watzl elucidates this by suggesting that under the stationary condition, color experience updates, say, ten times more often than it updates under the rotating condition, or 100 times per second compared to 10. When color experience updates less frequently, the rate of change seems to be slower. Watzl argues convincingly that this does not save the structural matching thesis, because rather than explaining why the speed of color change appears to decrease under the rotating condition, it suggests that the experience of color change is less “smooth” under rotation (Watzl 1022). He suggests that the theory confuses the rate at which experience changes from one color to the next (i.e. 10 degrees on the color wheel per second versus 100 degrees) which is what the experiment was concerned with, and the rate at which experience updates, which is what the argument addressed.

One possible response to Watzl’s critique of this argument would be to point to Paul’s work. The proponent of structural matching could argue that as long as the color information is updated frequently enough for the brain to translate what is being perceived as smooth motion, the jerkiness Watzl points to would not arise. Watzl already implicitly admits that the brain is probably smoothing a set of discrete states into an experience of smooth motion, since he raises no objection to the original (finite) number of updates. However, this response does not address how less frequent updating would lead to the illusion of slower motion. On that point, which Watzl would contend is more important, his critique is valid and quite damaging.

The second attempt to defend the structural matching thesis against the silencing experiment argues that experience of color change needs to be understood as experience of color change of specific objects, and under rotation participants are unable to track individual dots. Suchow and Alvarez constructed an experiment where the dots remained stationary but the background rotated, and found that participants still experienced the rate of color change as slower than in the condition where there was no additional rotation input, just like they experienced when the dots were rotating (Watzl 1024). This result suggests that the silencing effect cannot be attributed to a failure to track the individual dots, and that this defense of the structural matching thesis is ineffective.

The heart of this argument is that, under rotation, participants cannot fully track all the dots individually, and the circle of dots becomes

the more important perceptual entity. This suggests that the explanation for the silencing phenomenon is related to the level of stimulus participants experienced, rather than any specific stimulus, and Watzl believes that it represents an effective rebuttal of the second attempt at defense. However, this result can be used in support of the structural matching thesis at least as fruitfully as Watzl uses it against the thesis, if not more so. A more extended discussion of this can be found in the next section.

Finally, Watzl discusses the proposition that under rotation, the brain's processing capabilities are busy with the rotation input, and the color of the dots is processed as either completely indeterminate, or as a simplified color category like red rather than a particular shade of red (Watzl 1025). Complete color indeterminacy seems at odds with the evidence that participants notice if a dot's color is flipped to its complement, from orange to blue, for instance. Additionally, participants in the experiment do report that they experience color change under rotation, and it does not seem possible for color to change unless it is at least somewhat determined. Finally, Watzl argues that it is difficult to even understand what is meant by complete color indeterminacy. It seems difficult to imagine how one can experience an object without it having a determined color. With these counterarguments, the more radical form of indeterminacy is put aside, and consideration turns to the mild indeterminacy proposal. Claims that the color experience of a particular dot is constant are foiled by the same flipping argument that shuts down complete indeterminacy, and by pointing out again that participants do still experience color change, just at a slower rate (Watzl 1025). However, claims that the color changes under rotation but is not fully determinate are no more successful. Participants experience change as smooth, while color indeterminacy would suggest that the dot spends a period of time being indeterminately blue, then indeterminately green, but not that it passes through all the highly determined slices of a detailed color wheel. This would create an effect not of smoothness, but of occasional large changes in hue. According to Watzl, this is just another version of the original infrequent updating defense, and it fails for the same reasons (Watzl 1026).

Watzl's arguments are convincing here when one assumes that the participant maintains as much awareness of the indeterminate dots as she does the ones which remain determinate. Watzl uses the flipping experiment to justify this assumption, but does not consider that flipping the color of a dot could be a drastic enough change to bring the participant's attention back to that dot. Phillips' argument, introduced in

the next paragraph, goes into detail about how some dots could have indeterminate color under rotation. However, before leaving Watzl it is worth noting that he does not seem to really want to present the indeterminacy argument fairly. It is obvious to him, and to anyone else who spends a little time considering it, that there is no way to be fully aware of an object but experience its color as indeterminate. As such, it seems likely that the indeterminacy argument is more nuanced than Watzl's description of it. For this more nuanced argument, we now turn to Phillips.

PHILLIPS' DEFENSE OF THE STRUCTURAL MATCHING THESIS

Writing in response to Watzl, Phillips argues that Watzl, as well as Suchow and Alvarez, misinterprets the way the silencing effect operates to make color change appear slower. Watzl's interpretation is that the rotation condition leads us to experience the rate of color change of each dot as progressing slower than under the stationary condition (Phillips 701). Phillips suggests that a more accurate interpretation would be that as the circle of dots rotates faster, more and more of the individual dots are "crowded out" (Phillips 701). Crowding is a phenomenon in which it is difficult to recognize an object when it is surrounded by similar objects, especially when all the objects are in the periphery (as is the case in Suchow and Alvarez's experiment, since the participant is instructed to focus on the center of the circle). Phillips proposes that, under the stationary condition, the color change experience of each dot is enough to distinguish it and prevent it from being crowded out. However, under rotation, the experience of universal rotation of the system "subsumes" the individual changes in color, and some percentage of the dots are crowded out and thereby silenced (Phillips 702). Participants then experience some dots changing color at the original rate, and some not changing at all, but when asked by Suchow and Alvarez to describe the *single* rate at which all the dots are changing color, they unconsciously average the rates they are actually experiencing to provide the uniform rate that was requested.

The reasoning here is that if there is too much information, a participant's brain might not be able to fully process all the information it receives, and some of the inputs might be left out of the participant's conscious experience. This is supposed to be a way of accounting for the apparent discrepancy between the rate at which a participant's experience of color is changing (at the real rate) and the rate at which

the participant experiences the colors changing (a slower rate). It seems plausible that a scientist could design an experiment to test this assertion. For example, let us assume that the apparent deceleration in rate of color change is the result of an unconscious average of dots changing color at the original speed and dots which, due to crowding, are not perceived to change color at all. It should be possible to increase the rate of color change under rotation so that after participants average in the crowded out dots, the uniform rate of color change over all the dots is experienced as being the same as the rate experienced under the stationary condition. In contrast, if Watzl's interpretation is correct, one would expect participants to notice that not all the dots changed color at the same rate under the rotation condition. Until such an experiment has been designed and performed, it will be difficult to conclusively compare Watzl and Phillips' understandings of the silencing phenomenon.

It seems likely that Watzl would attempt to level the same criticisms against this theory that he used against the indeterminacy defense. Recall that the arguments against complete indeterminacy were that the evidence from the color flipping experiment showed that participants were tracking the colors enough to notice when they changed dramatically, that reports of color change require a color to be at minimum determined enough to undergo change, and that it is difficult to conceive of what a dot of indeterminate color means. Phillips addresses the first objection by saying that under his theory, a color flip would register with the participant because the dramatic change would cause the dot to reenter the participant's conscious experience (Phillips 700). However, Phillips considers this concern as separate from his project, and does not go into detail. This seems like a mistake. If adding the experience of rotation subsumes some of the experience of color change, as the changing colors of some dots go unnoticed, one would think that the experience of color flipping would not affect the experience of rotation, since rotation is a condition of the whole field, not of each separate dot. However, if a period of color flipping affected the experience of rotation, this result would argue against Phillips' proposal. This will be explored in more detail later in this paper.

Phillips would be able to respond to the second and third objections together. His crowding proposal would mean that some dots simply would not enter into the participant's experience under rotation, and would just melt into the entity 'the circle of dots.' It is not that they are experienced as having an indeterminate color, but that they are not really experienced at all. With this in mind, it seems that, as long as Phillips' proposal is backed up by evidence from further experiments,

it will succeed against Watzl in the long term. In the short term, Phillips successfully avoids the pitfalls which Watzl pointed out in other defenses of structural matching, and presents a sufficiently plausible account that the burden of proof shifts back to Watzl.

ALTERNATE DEFENSE OF THE STRUCTURAL MATCHING THESIS

Both Watzl and Phillips accept that participants' experience of the rate of color change is different when the circle of dots or the background is rotating compared to when there is no rotation. Both make the implicit assumption that the participants' experience of the passage of time is uniform across states. I would like to propose that doing away with this assumption can provide another defense of the structural matching thesis.

It is a well-documented fact of the human experience that the same amount of time may feel longer or shorter under different circumstances. Anyone who has been a high school student can attest that boredom can make one hour feel like three, and the opposite idea is so widely accepted that it has its own cliché, "time flies when you're having fun." Psychological experiments have shown that higher information processing loads can change "passage of time judgments," or judgments about how long a span of time feels (Wearden 2016, 133). Specifically, in two studies led by Wearden, increasing the amount of information participants were asked to process and report on was correlated with participants reporting a faster passage of time judgment (Wearden 2016, 133). It is worth noting that at least in the second study, increased information load was not so great that silencing became an issue, since it was reported that increasing the variation in context did not have a significant effect on the reported passage of time judgement, and no loss of accuracy was noted (Wearden et. al 2014, 296). A study led by Flaherty found an "S-Shaped relation between the amount of conscious information processing and time experience" (Wearden 2016, 135). The study reported that subjects experienced "protracted duration," or a stretch of time feeling longer subjectively than it was objectively, both when stimulus levels were very low, and when there was "intensified stimulus complexity" producing "greater density of experience per standard temporal unit" than under moderate levels of stimulus (Flaherty 76; Wearden 2016, 135).

Recall the experiment with which we started this paper – color-changing dots arranged either in a stationary circle or rotating around

a center point. The authors of the original paper, Suchow and Alvarez, as well as Watzl and Phillips, interpreted the results as silencing, and in their own way each attributed that to there being too much information for conscious experience to take note of it all. I would like to propose that the slower rate of color change reported by participants should be interpreted by considering the possibility that experience of time duration might change under rotation, rather than experience of color change. Although Phillips seems somewhat skeptical, he does not make a significant effort to argue against Suchow and Alvarez's interpretation of the color flip experiment, namely that participants accurately tracked the color of the dots through time (Phillips 700). As such, we will accept this conclusion as accurate. What remains to be accounted for is the rate of change discrepancy. Speed, of course, can be expressed as distance divided by time, with distance defined in this case as the number of degrees that the color of a given dot at time t is displaced from its original position on the color wheel. Suchow and Alvarez, Watzl, and Phillips all focused on the numerator, distance. I would like to focus on the denominator, duration.

Flaherty found that increasing the “density of experience,” what I shall refer to as the inputs, past a certain level made time appear to go slower. As long as the reader finds this result plausible, it is an easy step to apply it to a situation like the rotating condition where data suggests conscious experience is so overwhelmed by the perceptual inputs that it cannot process all of them. To be concrete, based on this we can expect that three seconds of rotation to feel longer to participants than three seconds during which the circle of dots does not rotate, since rotation represents a whole new set of inputs which have to be integrated into the participant's experience. Additionally, we know that participants do accurately experience the displacement around the color wheel of each dot. Taken together, this implies that participants experience the same amount of displacement over a longer perceived duration, which translates to a slower rate of color change under rotation. This is exactly what the experiment found to be the case.

Now we must consider whether this alternate explanation supports the structural matching thesis. Recall that the thesis's stance can be summarized by the statement “You experience something changing quickly only if it is the case that your experience of the thing is changing quickly.” Or, formulated more concretely, “You experience dot number 12 changing color at a rate of 70 degrees per second around a typical color wheel only if it is the case that your experience of the color of dot number 12 changes at a rate of 70 degrees per second around a

typical color wheel.” Watzl would claim that your experience of the dot’s color changes at a rate of 70 degrees per second, but your experience of the rate at which the dot changes color is 35 degrees per second, and that this contradicts the structural matching thesis. Supporters of the structural matching thesis would look to Phillips’ theory that, as long as dot number 12 escapes the effects of crowding out, your experience of the color of the dot changes at a rate of 70 degrees per second, and your experience of the rate at which the dot changes color is 70 degrees per second as well. If instead your experience of the dot’s color changes at 70 degrees per second, and you find that one objective second feels like it lasts two seconds, then your subjective experience will be of the dot changing color at 35 degrees per second. If your subjective experience were to be standardized and mapped onto a uniform timeline where each second was experienced as identical in duration to every other second, then the rate of change of your experience of the dot’s color and your experience of the dot’s rate of color change would match.

It is important to note here that I am not claiming that this interpretation is correct, only that it is at least as plausible as the interpretation taken by Watzl and the other authors cited above. Further experiments would be necessary to say more definitively which interpretation is closer to the truth. For example, an experiment could compare the original rotation condition to a rotation condition where some or all of the dots were flipping their colors periodically. If rotation appeared to continue at the same rate but the rate of color change was slower still with the added inputs, the result would favor Phillips’ explanation. Presumably the increased stimulation would lead more dots to be crowded out, but the system as a whole would remain intact, leaving rotation untouched. If rotation appeared to slow, this would support my interpretation. Increasing the density of experience should cause subjective time to move slower still. It is also plausible that there is a saturation point after which increasing the level of stimulus has no effect, so no experiential change could also potentially be explained under my proposal. Other, unexpected, results could support Watzl until proponents of the structural matching thesis provide coherent explanations.

CONCLUSION

Watzl attempts to cast doubt on the structural matching thesis using Suchow and Alvarez’s silencing experiment. It is Watzl’s contention

that the illusion from the experiment causes the layout of a participant's experience to diverge from the content of her experience, and that this is conclusive evidence against structural matching. The purpose of this paper was to show that Watzl is not successful in his attempt. There are at least two plausible explanations for the experiment's results which do not conflict with the structural matching thesis. Phillips' theory followed Watzl's strategy of analyzing how participants experienced the dots, but adopted a different interpretation. Phillips claimed that under the rotation condition, some of the dots are simply crowded out of the conscious experience, and the apparent decrease in the rate of color change is the result of averaging some dots experienced to be changing color at the original rate and the rest of the dots not experienced to be changing at all. The second theory focused on the perceived duration under the rotation and stationary conditions. It claimed that under the rotation condition, the additional input load resulted in an experience of extended duration, which led the rate of color change to appear slower.

This discussion would benefit from further experimentation. As it stands, this paper presents three potentially valid theories, and no definitively correct theories. As such, all that can be said is that Watzl's attack on the structural matching thesis has been beaten back for the moment. Watzl concludes his paper with an exhortation to continue to address time consciousness with an "interdisciplinary spirit" (Watzl 1031). I would also like to emphasize this. Suchow and Alvarez did work in the field of biology, and Flaherty was published in a journal of psychology. It is likely that other fields also have valuable contributions to make, and integrating them into this ultimately philosophical discussion can only make the picture richer and more complete.

BIBLIOGRAPHY

- Dainton, Barry. "Sensing Change." *Philosophical Issues*, vol. 18, no. 1, 2008, 362-384.
- Flaherty, Michael. "The Perception of Change and Situated Engrossment." *Social Psychology Quarterly*, vol. 54, no. 1, 1991, 76-85.
- Hoerl, Christoph. "Time and Tense in Perceptual Experience." *Philosophers' Imprint*, vol. 9, no. 12, 2009, 1-18.
- Paul, L. A. "Temporal Experience." *Journal of Philosophy*, vol. 107, no. 7, 2010, 333-359.
- Phillips, Ian. "Breaking the Silence: Motion Silencing and Experience of Change." *Philosophical Studies*, vol. 168, no. 3, 2014, 693-707.
- Watzl, Sebastian. "Silencing the Experience of Change." *Philosophical Studies*, vol. 165, no. 3, 2013, 1009-1032.
- Wearden, John. *The Psychology of Time Perception*. London: Palgrave Macmillan. 2016.
- Wearden, John, et al. "'Subjective Duration in the Laboratory and the World Outside.'" *Subjective Time: The Philosophy, Psychology, and Neuroscience of Temporality*. Ed. Arstila and Lloyd (pp. 287-306). Cambridge, MA: MIT Press. 2014.

