

Philosophy in Practice

VOLUME 8 – SPRING 2014



CALIFORNIA STATE UNIVERSITY, LOS ANGELES
DEPARTMENT OF PHILOSOPHY

Philosophy in Practice

VOLUME 8 – SPRING 2014

Cover Art: by Letizia Ragusa, “Untitled”

© Copyright 2014 by CSULA Philosophy Department. All rights reserved. Except for brief quotations in a review as permitted under the United States Copyright Act of 1976, no part of this publication may be reproduced or distributed in any form or by any means, or stored in a data base or retrieval system, without the prior written permission of the publisher. Individual copyright reverts to the authors upon further publication of their articles.

CONTENTS

Acknowledgments	v
Faculty	vi
Professor Spotlight: Sheila Price	viii
Articles:	
Troubles for a New Dispositional Account of Belief <i>Nathaniel Greeley</i>	1
Preserving Realism in the Philosophy of Science <i>Alysha Kassam</i>	21
Deflating Deflationary Meta-Ontology <i>Nicole Lavoie</i>	36
The Normative Critique of Law's Claim to Authority: A Bad Concept of Law <i>Pedro Viramontes, Jr.</i>	48
Reasons and Normativity <i>Mark Gaynor</i>	64
Phenomenology and Time: Husserl, Derrida, Zahavi <i>Jared Gee</i>	77
Why There Still May Be Such a Thing as Time-Travel <i>Samuel Batzdorff</i>	91
Reconsidering Two Notions of Analyticity <i>Neil Sanchez</i>	104
The Incompatibility of Compositional Nihilism and Direct Reference Theory <i>Lena Becerra</i>	119
A Defense of Bare Particulars <i>Ric Saenz</i>	128

Nihilism with a Berkeleyan Twist <i>Lorena Guandique</i>	147
The World Is Not Enough: How Hylemorphic Dualism Overcomes Problems with Cartesian Substance Dualism <i>Mark Morrow</i>	156
Contributors	174
Philosophy Program Information	176

ACKNOWLEDGMENTS

Philosophy in Practice is produced by students at California State University, Los Angeles. The editorial staff wishes to thank all the authors for submitting and refereeing articles, the faculty of the philosophy department for agreeing to supervise individual contributions, and the University for its generous financial support.

EDITORS:

Jared Gee
Nathaniel Greely
Neil Sanchez

FACULTY ADVISOR:

Dr. Michael K. Shim

CALIFORNIA STATE UNIVERSITY, LOS ANGELES

PHILOSOPHY FACULTY

David Pitt (2003–), Acting Chair, Ph.D. City University of New York. Philosophy of Mind, Philosophy of Language, Metaphysics

Mohammed Abed (2008–), Ph.D. University of Wisconsin, Madison. Ethics, Social and Political Philosophy, Philosophies of Violence, Genocide and Terrorism

Mark Balaguer (1992–), Ph.D. City University of New York. Philosophy of Mathematics, Metaphysics, Meta-ethics, Philosophy of Language, Logic

Talia Bettcher (2000–), Ph.D. University of California, Los Angeles. History of Modern Philosophy, Philosophy of Self, Gender and Sexuality

Anna Carastathis (2009–), Ph.D. McGill University. Social and Political Philosophy, Feminist Philosophy, Philosophy of Race, Critical Race Theory, Postcolonial Theory, Phenomenology

Richard Dean (2009–), Ph.D. University of North Carolina, Chapel Hill. Ethics, Kant's Moral Philosophy, Applied Ethics

Foad Dizadji-Bahmani (2013–), Ph.D. London School of Economics, United Kingdom. Philosophy of Science, Philosophy of Physics, Philosophy of Probability

Ronald Houts (1983–), Ph.D. University of California, Los Angeles. Metaphysics, Epistemology, Logic

Henry R. Mendell (1983–), Ph.D. Stanford University. Ancient Philosophy, History of Ancient Mathematics and Science, Philosophy of Science, Metaphysics

Joseph Prabhu (1978–), Ph.D. Boston University. Philosophy of Religion, 19th and 20th Century German Philosophy, Moral and Social Philosophy, Indian and Comparative Philosophy

Sheila Price (1964–). M.A. University of California, Los Angeles. Recent Philosophy, Comparative Religions, Medical Ethics, Environmental Ethics.

Michael K. Shim (2007–), Ph.D. State University of New York, Stony Brook. 20th Century Continental Philosophy, Phenomenology, Husserl, Modern Philosophy, Philosophy of Mind, Philosophy of Language

Kayley Vernallis (1993–), Ph.D. University of California, Berkeley. Moral Psychology, 19th and 20th Century Continental Philosophy, Feminist Philosophy, Ethics, Aesthetics, Gender and Sexuality

EMERITUS PROFESSORS

Thomas Annese (1961–1992). Epistemology, Modern Philosophy.

Sharon Bishop (1967–2004). Ethics, Political Philosophy, Philosophical Psychology, Feminist Ethics.

Donald Burrill (1962–1992). Ethics, Philosophy of Law, American Philosophy.

Ann Garry (1969–2011). Feminist Philosophy, Philosophical Methodology, Epistemology, Applied Ethics, Wittgenstein, Philosophy of Law

Ricardo J. Gómez (1983–2011). Philosophy of Science and Technology, Philosophy of Mathematics, Kant, Latin American Philosophy

George Vick (1967–1997). Metaphysics, Phenomenology, Existentialism, Philosophy of Religion, Medieval Philosophy.

PROFESSOR SPOTLIGHT: SHEILA PRICE



Small in stature but big on charm and intellect, Sheila Price has been a fixture at CSULA for 50 years now, making her the longest running member of the Philosophy Department. Given her diverse talents, the fact that we've been able to hold on to her this long must mean that we're doing something right. She was recruited right out of UCLA in 1964, a time when the climate for women in philosophy was even more challenging than it is today.

"When I was a grad student, the UCLA philosophy department finally hired one woman. I

remember a professor, commenting on the possibility of hiring a second, saying that 'one woman was enough'."

She was so captivated with philosophy, however, that she says this difficult climate never fazed her. After getting her master's degree in philosophy at UCLA and a second master's in folklore and mythology for good measure, then head of the department at CSULA, Ed Sales, phoned her up and offered her a job. She took it, and the rest is history. She found CSULA to be quite friendly to women.

"Women are certainly encouraged here and they're hired here. I think there were times when there were more women than men in the department. I've certainly been appreciated here."

Price has witnessed a number of social upheavals in her time here. She taught during the Vietnam era, and CSULA was not immune to the general turmoil that prevailed.

"There was tremendous excitement on the campuses during the Vietnam War. We had all kinds of protests. CSULA wasn't any

different than any of the other campuses in that respect. I had a lot of students who had served in Vietnam. I remember they would talk to me because they found it very hard to focus. People were less aware of PTSD in those days.”

Price even found herself the subject of government scrutiny during this time, though perhaps accidentally.

“We had a radical professor, I forget his name. He was away for one semester and I had to cover his intro to philosophy class. One student always came to class in a suit, and was always taking notes. One day he came to office hours, this was shortly after J. Edgar Hoover had died, and informed me that he had enjoyed the class but he had to go back to Washington. He was FBI. They had sent him to monitor the other professor, and when he told his bosses in Washington that the professor had changed, they told him to stick around anyway. They were hoping to find that I was deeply subversive. At the time, however, you could teach intro to philosophy any way you liked and I taught it as ancient Greek philosophy. So the FBI didn’t find anything subversive about Socrates. Tell that to Thrasymachus and the Sophists.”

Through the many changes, big and small, that have occurred in the intervening decades, Price always recounts them on a personal level, as filtered through her interactions with students. She recalls her individual conversations with Vietnam vets or with female Iranian students during the Iranian revolution.

“I’m not the kind of teacher who went out to coffee with her students. But I really care about my students and I worry about them. I’m particularly concerned in this economy that we have. They all want a better life, and I can’t guarantee that having an education will do that for them, financially at least. It used to be the case that you could.”

When asked if CSULA has a special role in helping students find a better life, she replied, “Yes, we’re accessible. A lot of people are relatively late bloomers. They may not have the grades to get into UCLA right away. They don’t like high school and they don’t get the grades, but once they start at the university they find what they’re good at. So I think CSULA plays a special role.”

Price is a popular lecturer, particularly beloved by her students. Because of that, she's had the opportunity to teach a diverse range of courses.

"I taught in a lot of different departments over the years. I taught in Asian Studies, American Studies, History, and Philosophy. You name it I taught it. Back then if you had an idea for a course, you could teach it. Things have tightened up since then. I was billed as part time, but I did all the things full time people do. I guess I just didn't get as much money."

Price says that the one thread that connects her diverse interests has always been belief. "I'm interested in the nature of belief, why people believe, what they believe, and the emotions that accompany belief." To this end she employs texts as diverse as the works of Plato, William James' *Varieties of Religious Experience*, and Philip K. Dick's *Do Androids Dream of Electric Sheep?* (the book on which the film *Blade Runner* is based).

When asked if the students have changed over the years, she replied, "Well, as you get older, students look younger. But students now are very worried about money and getting out on time. I can't blame them, but sometimes they resent the fact that this isn't a trade school. They'll say, 'I'm an accounting major, why do I need to learn philosophy?' I have to explain that at a university you're becoming a sophisticated person. You're learning more than what's required for a job. You'll never be smarter than when you're sitting in these classes. I meet professional people and they consistently tell me, 'I took a philosophy course as an undergrad. I wish I could take it again. I miss the stimulation. Now all I know is ophthalmology.'"

When asked if she would still recommend philosophy as a major, given students' economic concerns, she replied:

"Yes, I think it's a great major. It provides a deeper understanding. Whatever philosophy is, it's not superficial. The aim of philosophy is to make you smarter, and it's done by increasing your vocabulary. Having a better vocabulary makes you perceive more."

As for her plans for the future, Price says she intends to stay

right here and teach for as long as she can. Since she's officially retired, she doesn't teach as many courses as she used to. In her spare time, she reads avidly.

"I have a rule that I only read nonfiction before 5pm. After 5, I'm allowed to read fiction."

She's also studying Spanish, though she already speaks German and French.

"Time being what it is, I don't know if I'll have the time to learn it."

With boundless energy and intelligence like hers, however, surely it won't be a problem.

— N.G. et al

TROUBLES FOR A NEW DISPOSITIONAL ACCOUNT OF BELIEF

Nathaniel Greely

INTRODUCTION

I believe that George Washington was the first President of the United States. But what is a belief, and what is it for me to have one? One popular view holds that beliefs are stored mental representations. On this view, somewhere in the vast data bank of my mind there sits a belief about old George, acquired in elementary school and ready to be recalled when necessary. On this account, beliefs are accessed directly, and I either have them or I don't.

Beliefs don't always seem so clear-cut, however. It's possible that I might be mistaken or deceived about what I believe, particularly if facing the truth could prove unpleasant. It's also possible that whether I have a particular belief or not depends on the context. Dispositionalism is an alternative theory of belief that is supposed to make better sense of these more complicated situations. Dispositions are tendencies to exhibit certain behaviors, or to have certain thoughts or experiences. Dispositionalism about belief describes people's beliefs in terms of their dispositions. I will focus on a contemporary dispositionalist account offered by Eric Schwitzgebel. Schwitzgebel claims that to have a belief is to fit a "dispositional stereotype" associated with that belief. If a person's behavior, both physical and mental, lines up with what an average person would expect of someone who believes, say, that George Washington was the first President, then that person has the belief that George Washington was the first President. On this view, then, to have a belief is to fit the appropriate dispositional profile.

This means that on a dispositional account of belief, we don't have direct access to our own beliefs and can often be wrong

about what we believe. If I claim to believe that all people are equally deserving of respect while consistently treating people of lower socio-economic status with disdain, then according to Schwitzgebel I am mistaken about what I believe. When what we think we believe differs from what our dispositions tell us, Schwitzgebel claims that the dispositions give us the real story. I will offer counterexamples that suggest that Schwitzgebel is wrong on this point. There are a number of situations in which we would intuitively attribute a belief to a person even if she does not fit the dispositional stereotype. This is a problem for Schwitzgebel because he explicitly relies on folk intuitions as the proper arbiter of belief ascription. If my examples are convincing, then folk intuitions about what we believe diverge from dispositionalism in a wide range of cases. This means that, given Schwitzgebel's own criteria for assessing theories of belief, dispositionalism fails in a wide range of cases.

In Section 1, I will spell out Schwitzgebel's account of belief in detail, contrasting it with representationalism and pointing out the aspects of his account that make it vulnerable to counterexamples. In Section 2, I will offer two such counterexamples. Once these are laid out, we shall see that there exists a whole class of beliefs that prove problematic for dispositionalism. In Section 3, I will discuss some possible objections to my arguments and respond to those objections. In section 4, I will briefly return to an interesting result of one of my counterexamples. It seems that on a dispositionalist account of belief, it's doubtful that any human could be ascribed a belief in dispositionalism. Although this last point is not central to my thesis, and it's not even clear whether it constitutes an objection to dispositionalism, it does create a puzzle that may be worth further consideration.

SECTION 1 – SCHWITZGEBEL'S ACCOUNT OF BELIEF

Before I detail Schwitzgebel's dispositional account of belief, it's worthwhile to briefly explain the primary view to which he

is opposed. The representational account of belief is arguably the most intuitive account, and for this reason has become the dominant theory of belief. To explain it, it will be helpful to use a slightly worn metaphor. Many find it natural to think of a mind as a sort of computer, containing a large data bank filled with information. Unlike a silicon-based computer, the storage medium of our minds is neural, but these neural states encode mental representations. Many of these representations are beliefs that represent the world as being a certain way. In my mind, there is a belief that represents San Francisco as being north of Los Angeles. If someone asks me which way it is to San Francisco, I simply search the databank of my mind, find the right belief, and respond, “North”. Not all beliefs are easily accessible. The name of my third grade teacher might be deeply buried or lost altogether, but for the most part, if I want to know what I believe, I can find out easily and directly by accessing the appropriate representation. According to representationalism, then, having a belief is like having a file on your hard drive – it’s either there or it isn’t, and it’s usually not hard to find out if it is. Schwitzgebel disputes the representational account of belief because he holds that beliefs aren’t always so clear-cut. Instead, he proposes an account that is more flexible.

In order to understand Schwitzgebel’s account of belief, it is important to understand what he means by dispositions and dispositional stereotypes.

Dispositions can be characterized by means of conditional statements of the form: If condition C holds, then object O will (or is likely to) enter (or remain in) state S. O’s entering S we may call the *manifestation* of the disposition, C we may call *condition of manifestation* of the disposition, and the event of C’s obtaining we may call the *trigger*. (Schwitzgebel 2002, p. 250)

For example, if the condition of my dropping the hard-bound, complete works of Aristotle on my bare toes obtains, then the object known as me is likely to enter into a state of shouting obscenities and hopping around. In this example state S is an

objective, observable behavior, but on Schwitzgebel's account it needn't be. State S can also be a phenomenal state, such as feeling pain or seeing stars, or a cognitive state, such as thinking what a clumsy oaf I am. This is what distinguishes his dispositionalism from earlier, behaviorist versions and, for this reason, he calls his account a "phenomenal, dispositional account of belief" (Schwitzgebel 2002, p. 249). I have the disposition described by any such conditional, then, just in case that conditional statement is true.

Note that these conditionals are defeasible. They only hold *ceteris paribus*, or all else being equal. The conditional would be false if I were doped up on morphine. According to Schwitzgebel, this fact is no more problematic for his theory than it is for the natural sciences. Scientific generalizations only hold *ceteris paribus* as well. A dropped coin will fall to the ground unless it is in a strong magnetic field, or caught in an updraft, or is caught by a greedy bystander, and so on indefinitely. Such generalizations can be immensely useful, even if all of the counterfactual defeaters cannot be formulated explicitly.

The next important concept in Schwitzgebel's account is that of a dispositional stereotype or profile. "A dispositional stereotype is a stereotype whose elements are dispositional properties" (Schwitzgebel 2002, p. 251). A good way to think about dispositional stereotypes is to consider personality traits. A person can be described as hot tempered if she has the disposition to feel anger quickly and with little provocation, to raise her voice, and so on. Similarly, we might think of the dispositional stereotype associated with a belief as a bundle of dispositions. A standard example in the literature is the belief that there is beer in the fridge. Someone who believes that there is beer in the fridge is likely to have a number of dispositions. For example, she will be disposed to go get a beer if she wants one. She will also be disposed to offer a beer to a guest who is thirsty. When asked if she has any beer, she is likely to answer in the affirmative. She will also be disposed to feel surprise if she opens the fridge and finds that the beer is gone. The list could go on indefinitely, and we may not all agree on which dispositions are appropriate. No single disposition is either

necessary or sufficient to definitively attribute a belief to some person. Rather, there will be a cluster of dispositions that typical speakers of the language, i.e. “the folk”, are apt to associate with a given belief, and this constitutes the dispositional stereotype for that belief. If a person matches that stereotype to a sufficient degree, she can be described as having that belief.

To believe that P, on the view I am proposing, is nothing more than to match to an appropriate degree and in appropriate respects the dispositional stereotype for believing that P. What respects and degrees of match are to count as “appropriate” will vary contextually and so must be left to the ascriber’s judgment. (Schwitzgebel 2002, p. 253)

Recall that any given disposition within the stereotype holds only *ceteris paribus*. If I am particularly greedy, I might deny that there is beer in my fridge when asked. In such cases, the person in question can be “excused” from exhibiting that particular disposition, but will still be said to have the belief if enough of her other dispositions match the stereotype. Schwitzgebel’s account claims to be superior to earlier, behavioristic dispositional accounts of belief because it includes dispositions to enter various mental states. Behaviorists avoid any talk of mental states in favor of observable behavior. Behaviorism is generally thought to fail, however, when we imagine situations in which every observable behavior associated with a particular belief is constrained, but it still seems intuitively correct to say that the person has the belief. It is much more difficult to imagine a similar level of constraint over one’s phenomenal and cognitive states. I might have some reason to want to hide my surprise about the empty fridge from my guests, but I will nonetheless *feel* surprised, and so Schwitzgebel’s account succeeds where behaviorist dispositionalism fails. It provides the intuitively correct answer that I do believe there is beer in the fridge even if my objective behavior doesn’t always reflect it.

If this account of belief sounds a bit complex and fuzzy compared to representationalism, this is the point. Schwitzgeb-

el's account is supposed to capture the complexity and fuzziness inherent in many cases of belief attribution. In particular, Schwitzgebel claims that there are situations in which it is not appropriate to describe a person as either having a belief or not, which would be impossible according to representationalism. Schwitzgebel holds that sometimes people should be described as having a belief in certain contexts but not having it in others, depending upon their dispositional profile. He offers the example of a mother who suspects that her son smokes marijuana. When she defends her son against her husband's accusations, she truly believes that he is innocent. However, on days when she speaks to her therapist in confidence, she might truly believe that her son does smoke pot. A dispositional account of belief will give us this result because her dispositional profile hovers somewhere in the middle, changing from situation to situation, whereas a representational account requires that deep down, she "really" believes one or the other. Schwitzgebel cites this flexibility as one of the virtues of his theory (Schwitzgebel 2002).

It is important to note that Schwitzgebel's dispositionalism is not a metaphysical claim about what a belief is. Rather, it is a claim about when we can accurately describe someone as having a belief. Compare this to the representational account, which does make a metaphysical claim about what beliefs are. Beliefs are mental representations. The representational method of belief attribution flows from this metaphysical claim. To have a belief is to have the appropriate mental representation. Lacking such a metaphysical grounding, Schwitzgebel stakes the truth of his account entirely on the assumption that folk attributions of belief are by and large correct.

To believe that P is to act and react, and be otherwise disposed, in ways that ordinary people would regard as characteristic of believing that P. This might sound circular, but it's not. It is to ground metaphysics in folk psychology: There are some patterns in the world that ordinary belief ascribers have glommed on to, patterns that drive their apti-

tude to regard certain dispositional tendencies as characteristic of believers. (Schwitzgebel 2011, p. 8)

Schwitzgebel is not denying that there might be some metaphysical underpinning to beliefs, but he does not commit to what that might be. His heavy reliance on folk intuitions as somehow tracking the metaphysical reality, however, makes his account vulnerable to counterexamples in which our folk intuitions diverge from dispositionalism, for his claim is that dispositionalism is the right way to characterize these intuitions.

In the next section I will offer such counterexamples. They will follow a pattern set by Schwitzgebel in a 2011 paper. In this paper, Schwitzgebel points out that on a dispositional account of belief, we don't have privileged access to our own beliefs. Of course, a disposition to attribute a given belief to oneself will be part of the dispositional stereotype associated with that belief, but it will be only one disposition among many. Schwitzgebel offers an example to make the point. He posits a hypothetical character, Professor Piotr, who "will sincerely, unhesitatingly, and unqualifiedly assert, with a feeling of inward assent, that all people deserve equal respect, including those below him in status" (Schwitzgebel 2011, p. 20). In other words, Piotr has a disposition to attribute to himself a belief that all people deserve equal respect, and he's disposed to say as much. The basic claim of dispositionalism, however, is that this is not enough for us to be sure what Piotr believes. The concern is not that Piotr may be lying when he claims to believe in equality. Rather, Piotr just can't know what he believes on the basis of self-attribution alone. For us, or Piotr himself, to know what he believes, we need to look at his dispositions and compare them to the dispositional stereotype. Schwitzgebel goes on to describe myriad ways in which Piotr violates the dispositional stereotype associated with the belief that all people deserve equal respect. Piotr is short with the housekeeper, feels indignant when the University President has to fly coach, and so on. On the whole, his behavioral and mental dispositions weigh against his having the belief he claims, because to have the belief

just is to fit the stereotype.

To the extent that believing a proposition is not only a matter of being disposed to simply assert but also a matter of how one steers one's way through the world—a matter of how one acts and reacts in a wide variety of situations, what one implicitly assumes, what is exhibited in one's reasoning and emotional reactions and spontaneous behavior—to that extent, any privilege that attaches to one's knowledge of what one would say or explicitly judge extends only contingently to the matter of what one believes in too. (Schwitzgebel 2011, p. 21)

The Piotr example is, of course, designed to nudge our intuitions toward the conclusion that Piotr doesn't really believe that all people deserve equal treatment, just as a dispositionalist would conclude after adding up Piotr's various dispositions. In a 2012 paper Schwitzgebel describes similar examples and then offers something of a challenge. In contrasting “deep” accounts of belief like representationalism with the “superficial” dispositional account that he proposes, Schwitzgebel claims that when the deep account of belief diverges from the dispositional account, our folk intuitions about whether we should attribute the belief to that person will match the verdict of dispositionalism.

Faced with such an account, here's what I would do: Attempt to discover the maximum possible divergence the account allows between the deep structure and the superficial dispositional profile, and then consider whether, in such cases, dispositional profile or deep structure would produce better classificational practice given our interests. Existing deep accounts, to the extent they commit clearly enough to permit such comparisons, will, I wager, tend to lose the contest. (Schwitzgebel 2012, p. 16)

When the belief revealed upon direct introspection contradicts one's dispositional profile, Schwitzgebel bets that we will favor the verdict of dispositionalism. In fact, he must claim this.

As I have pointed out, lacking any explicit metaphysical basis, Schwitzgebel's account of belief is only accountable to the verdict of folk intuition, which he claims will accurately track whatever metaphysical truth there is regarding belief possession. In what follows, I will take Schwitzgebel up on this wager. I will offer examples in which the deep account diverges from the shallow account, but in which nonetheless our intuitions fall on the side of the deep account. Relying as he does on folk intuitions, unless Schwitzgebel can deny or explain away the apparent divergence, these examples will show that dispositionalism is not an appropriate description of the folk method of belief attribution. Since, according to Schwitzgebel, the folk method gets it right, this means that dispositionalism is wrong in these cases. I think we will find that such cases are widespread.

SECTION 2 – COUNTEREXAMPLES

Let's start with an example of extreme divergence between the "deep", representational account and the "shallow", dispositional account of belief. Imagine a professor not unlike Schwitzgebel's Piotr; call her Pietra. Pietra will sincerely, unhesitatingly, and unqualifiedly assert, with a feeling of inward assent, that philosophical determinism is true and compatibilism is false. In other words, Pietra attributes to herself the belief that every event is pre-determined by the initial state of the universe and the laws of physics, and that this is incompatible with the proposition that humans have free will. The way I have worded it is somewhat technical, but the belief is not particularly obscure or rare. Versions of determinism are at least as old as Sophocles and are common today not only among philosophers, but also scientists, depressed teens, and many Protestant denominations. In fact, it's arguably entailed by a belief in the omniscient God of classical theism, in which case there are billions of determinists.

Let's look at Pietra's behavior, however. She is indeed disposed to attribute a belief in determinism to herself, both inwardly and in conversation. She is also disposed to write papers

that profess this belief. These are all dispositions we would associate with someone who believes in determinism. But this is where the overlap stops. Pietra is frequently (nay, incessantly) disposed to deliberate before undertaking actions of any significance. She is disposed to feel and act as if those actions were free choices and that she could have done otherwise. She is also disposed to feel guilt when she has done something she considers to be wrong, and feel resentment when others do wrong. She is also disposed to admonish wrongdoers and in general to treat others as if they had the power to make free choices.

Does this behavior match the dispositional stereotype of someone who believes in determinism? Is Pietra's self-ascription of belief accurate, or is she as self-deceived as Piotr? The two cases seem perfectly isometric. In both cases the subjects match the dispositional stereotype of their self-ascribed belief only in their dispositions to self-ascription. In nearly every other aspect of their behavior they do not fit the stereotype. This seems to be exactly the sort of case Schwitzgebel is referring to when he says:

If someone is behaviorally, phenomenally, and cognitively disposed perfectly belief-that-P-ishly across the board, lacking only the tendency to self-ascribe that belief, we might well enough in most circumstances go ahead and say that she believes that P. (Schwitzgebel 2011, p. 9)

Piotr is behaviorally, phenomenally, and cognitively disposed perfectly belief-that-people-shouldn't-be-treated-equally-ishly, except in his tendency toward self-ascribing that belief, and Schwitzgebel goes ahead and says that he believes that people shouldn't be treated equally. Pietra is behaviorally, phenomenally, and cognitively disposed perfectly belief-that-humans-have-libertarian-free-will-ishly, except in her tendency toward self-ascription, and so Schwitzgebel should declare that she believes in free will.

So it seems we have the dispositionalist's verdict on Pietra, at least if Schwitzgebel is consistent. She does not believe in determinism. But where do our folk intuitions lie? I think it's fair to

say that most people would take her self-attribution at face value. If Pietra has thought it through and convinced herself that determinism is true, then she has that belief. There's no need for her behavior to change in extremely inconvenient, or perhaps impossible, ways in order to have a belief.

This example is only an extreme version of many such possible counterexamples. This is because Schwitzgebel's account leaves no room for the familiar phenomena, grounded in folk psychology, of hypocrisy and weak will. I believe smoking will kill me, I believe I want to live, and yet I smoke. I believe that buying a certain product is unethical, and I believe I should act ethically, yet I often buy this product because it's cheap and convenient. Isn't it possible that I truly hold these beliefs but I just don't have what it takes to act in accordance with them? Folk psychology would say yes. According to Schwitzgebel, however, Piotr and the rest of us are not weak, lazy, or hypocritical. Instead, we are self-deceived.

Schwitzgebel's account of belief comes with a host of *ceteris paribus* clauses that might be used against these counterexamples. To be fair, I have not even spelled out the stereotype that we would associate with someone who believes in determinism, I've just appealed to a gut reaction that Pietra doesn't fit it. But then neither does Schwitzgebel spell out the stereotype associated with a belief in equality in the Piotr case. This is his point—that unaided folk intuitions determine the stereotype, determine whether one fits that stereotype, and by the terms of his wager, determine whether dispositionalism gives us the right answer. I think in the Pietra case, the smoker case, the hypocritical shopper, and many other conceivable situations, our folk intuitions show that dispositionalism does not give us the right answer. Schwitzgebel allows that in certain situations the appropriate group of "folk" that determine the stereotype might be specialized. He also claims that in some cases we rightly weigh dispositions to self-attribution more heavily than other dispositions. I will address these and other possible objections in due time, but first I'd like to offer another example.

Take Professor S. S professes a belief in a dispositional account of belief. Just like Piotr and Pietra, he is inclined to attribute this belief to himself, both inwardly and in conversation, and he is disposed to write papers in favor of dispositionalism. But, on the whole, does he act as if dispositionalism were true? Or does he, on the whole, treat his own beliefs and others' as if they are directly accessible upon introspection? I think that the latter is the case, not only for Professor S but for any human. If this is the case, then S does not fit the dispositional profile of a dispositionalist and therefore doesn't believe in dispositionalism, according to dispositionalism.

The relevant behavior to look out for will be Professor S's ascriptions of belief. In order to fit the stereotype of a dispositionalist, he should ascribe beliefs to himself and others based upon whether their dispositions match the stereotype. When ascribing beliefs to other people, he will have to do this largely on the basis of a survey of their observable behavior. In his own case, he can compare his cognitive and phenomenal dispositions to the stereotype as well. In both cases, he would consider self-ascriptions of belief as only one among many relevant factors. Now, it's not that professor S never does this. We do sometimes survey other people's behavior in order to assess their professions of belief, but usually only when we think people are lying to us, or to themselves. Professor S may even do this a bit more than most people. On the whole, however, I submit that if Professor S is a human, he will tend to take people at their word and trust that self-ascriptions of belief are accurate. But as a dispositionalist, he shouldn't.

Here it is important to note that Schwitzgebel privileges other people's assessments of our beliefs over our own. He believes that folk intuitions about belief do accurately track the metaphysical truth, but he also claims, "our attitudes are rarely quite what we hope or think they are" (Schwitzgebel 2012, p. 21). By attitudes he means propositional attitudes like belief. So according to Schwitzgebel, insofar as our intuitive attributions of belief are about others and are based largely on behavior, we can rely on them, but when we introspect upon our own beliefs, we are

usually wrong. This means that others will usually be wrong about their own beliefs as well. Since on the dispositionalist account, there's no reason why we should accept self-attributions of belief at face value, and since most other people are even less likely than S to be careful about this, then most of the people S encounters on a daily basis are likely to be wrong about their beliefs a good deal of the time. S asks a woman where the bus stop is. She will futilely rack her brain and come up with a likely mistaken self-ascription of a belief about the bus system, and report this to S. And he will rashly take her word for it. He does this sort of thing all day long. He even accepts his own self-ascriptions about these everyday sorts of beliefs without further investigation the vast majority of the time. He might be more inclined to assess his own beliefs like a dispositionalist when in a session with his therapist, making surprising discoveries about his beliefs through painstaking analysis of his own dispositions. But can we imagine him living his life like that? One of Schwitzgebel's own examples illustrates the absurdity of this notion.

I ask Laura if she believes that there is anything wrong with gay men having consensual sexual intercourse, and she answers that she sees nothing wrong with it. What processes might drive Laura's self-ascription? Why not: lots? She might answer the question in part by thinking about whether there really is anything wrong with gay men having consensual sexual intercourse... she might be partly calling up the moral facts (or putative moral facts) from memory... Laura's self-ascription might also be partly driven by her general conception of herself as a liberal... she might visually imagine gay sex and have an emotional or aesthetic reaction... (Schwitzgebel 2011, pp. 12-13)

The passage goes on at length, and I've only selected some parts to provide the flavor of Schwitzgebel's view on the phenomenology of the self-ascription of belief. At first glance his description may seem reasonable, but he is in fact conflating several different phenomena. Some of his descriptions of what could be

going on in Laura's head more accurately describe someone who is attempting to formulate a belief that they do not decisively have yet. Others sound like someone who is calculating how to present herself to an audience. Those of us who actually have beliefs on the subject of homosexuality and are prepared to answer the question truthfully needn't go through any such acrobatics. We simply answer the question because we think that we know what we believe directly. I submit that our Professor S, like the rest of humanity, would act as though his own beliefs were directly available through introspection the vast majority of the time. It's hard to imagine anyone effectively getting through life in any other manner.

There are two conclusions we can draw from this example. First, like the Pietra case, I think that intuitively most of us would say that, whether or not S acts like a dispositionalist most of the time, S can still believe in dispositionalism. That's where our folk intuitions lie. Like Pietra, if Professor S has taken the time to consider the matter carefully and convince himself of the truth of dispositionalism, then he has the belief. Whether he acts on it is another matter entirely. But there is a further point that makes this example more interesting. If we assume that dispositionalism is true, and then concur that S, along with every other human, does not fit the stereotype of someone who believes in dispositionalism, we end up with the conclusion that, on a dispositionalist account of belief, dispositionalism cannot, as a matter of psychological or at least practical necessity, be believed. This point is not essential to my argument. In fact, it's not clear what implications it has for dispositionalism and so I will wait until Section 4 to address it. Instead, I will move on and address some possible objections to my arguments in this section.

SECTION 3 – OBJECTIONS

Remember that in the Pietra example I followed Schwitzgebel's lead in his Piotr example by not explicitly drawing up the dispositional profile for a belief in determinism. Perhaps Schwitzgebel

can plausibly draw up a dispositional profile that would give us what I claim is the intuitively correct answer—that Pietra does believe in determinism. If so, then dispositionalism and folk intuitions would agree and all would be well for Schwitzgebel. One strategy might be to claim that belief in determinism is a specialized sort of belief and that the dispositional stereotype should reflect what people who are well versed in the subject should expect. This is one explicit caveat that Schwitzgebel makes about stereotypes. Although the vast majority of dispositional stereotypes should reflect folk notions, he points out that certain “novel or half-novel or sub-culturally local property types” might require a more specialized group to draw up the stereotype (Schwitzgebel 2012, p. 9). Maybe a group of professional philosophers, familiar with the behavior of professed determinists, would be more lenient in their dispositional profile and wouldn’t expect the belief to be reflected in Pietra’s behavior.

I don’t think determinism should qualify for this treatment. Recall that when I introduced the example, I pointed out that there are many sorts of determinists, few of which were professional philosophers. Any given college undergraduate is likely to have a rudimentary grasp of the concept. Any such person is likely to draw up a stereotype that Pietra does not fit, for they are likely to have some notion of the concept of fate and what concrete implications this should have for our lives. If there were no such implications, the subject wouldn’t elicit the interest it does and it would not be taught so widely.

It might seem contradictory for me to first claim that the folk would draw up a dispositional stereotype for a determinist that Pietra doesn’t fit and then to turn around and say that nonetheless the folk would attribute to her a belief in determinism. My point is that this is the whole problem with dispositionalism. We intuitively separate beliefs and behavior. To marry the two violates folk intuition, so Schwitzgebel can’t appeal to folk intuition to justify the marriage.

The determinism case is only one among many examples we could use, most of which require no specialized knowledge

whatsoever. I only used it because the divergence between the deep and shallow accounts is so extreme in this case. Schwitzgebel's wager concerned cases of extreme divergence, and the Pietra case was designed to win that bet. Once the point is made, however, we can see that problems crop up for dispositionalism in any number of cases, such as the weak-willed smoker and the hypocritical shopper. Such cases are not as clear-cut as the determinism example. The smoker may have a number of dispositions that point to divergent beliefs. They might approach a 50/50 split in some cases. These will have to be weighed against each other by the dispositionalist before reaching a verdict. But all I need to make my point is a divergence between the final verdicts of the representational and dispositionalist accounts. If the dispositional calculus gives one answer and the representational account gives another, yet folk intuitions fall on the side of representationalism, then dispositionalism fails in those cases as well. Even the Piotr example itself, if we revisit it having considered my points, appears in a different light. Many of us profess a belief in equality, yet none of us are completely innocent of the sorts of behaviors described in the Piotr case. Maybe, on the whole, many people's behavior even tips toward the side of hypocrisy. It's not intuitively obvious that in these cases we should deny people their beliefs.

Another possible objection might be that, in the Pietra case, it is somehow impossible to act as though determinism were true, and that therefore this falls under one of Schwitzgebel's *ceteris paribus* clauses. He notes that there are excusable deviations from the stereotype, such as when a person is physically restrained and thus unable to act in accordance with the stereotype. "Certain types of conditions are regularly regarded as excusers in this sense, such as physical incapacity" (Schwitzgebel 2002, p. 254). But not all behaviors that would match the stereotype of a determinist are out of our control, if determinism is false. If determinism is true, then of course all of our behavior is out of our control and the dispositionalist account would be trivially true given the *ceteris paribus* clause, for all of our behavior would be excused from matching any dispositional stereotype. If we stipulate, however,

that in the Pietra case determinism is false, then we should expect Pietra to be able to act in accordance with the belief. Now, as a determinist, she might think that she can't alter her behavior even though she can, since we are assuming determinism is false. This might dissuade her from trying to act like a determinist, but it should also dissuade her from trying to act in general, and this is something she cannot do. She will continue to deliberate, make choices, exert herself in hopes of achieving contingent objectives, and so on. This is not consistent with the stereotype of a determinist. Also, the points I made in connection with the first objection apply here too. Since the determinism case is only a model of extreme divergence between the deep and shallow accounts, we can easily formulate less extreme examples of beliefs whose associated behavioral dispositions are more clearly within our control. The Professor S example allows for more control, and the hypocritical shopper even more. Certainly there are other possible examples.

Schwitzgebel also allows that, in assessing some beliefs, self-attributions should be weighed more heavily than other dispositions. He doesn't allow for many of these, as it "encourages noxiously comfortable self-portraits" (Schwitzgebel 2012, p. 20). Nonetheless, we can imagine situations where the only dispositions available to us to evaluate are dispositions to self-attribution. Take, for example, a belief that time is circular, or that there is a black hole at the center of some particular distant galaxy. In this case, the dispositional profile would only differ from that of a person who held the opposite beliefs (that time is linear, or that there is no such black hole) in regards to dispositions to self-attribution (i.e. writing papers, giving lectures, etc...). There just aren't many consequences for such beliefs in everyday life. We can see, then, how a belief in determinism differs immensely. Compared to these examples, the everyday consequences of determinism regarding ethics and our concept of self are great and pervasive. This is why the topic is interesting. Similarly, we have seen that a belief in dispositionalism requires an enormous expenditure of time and energy evaluating our own and others' beliefs

in terms of behavior, since the dispositionalist can't take them, or herself, at their word. Therefore it seems that such an objection is applicable to neither example.

SECTION 4 – AN UNBELIEVABLE ACCOUNT OF BELIEF

Having made my argument and responded to objections, I would like to return briefly to an interesting point that came up in the course of my argument. The Professor S example seems to show that dispositionalism, if true, can't be believed. There are three worries for dispositionalism that I would like to point out in connection with this. My argument does not depend on it, but they seem too interesting to pass over completely, though I don't have time to explore them in depth.

First, the incredibility of dispositionalism commits its advocates to a Moorean Paradox. G. E. Moore pointed out that it seems paradoxical to simultaneously assert, in the first person, "P is true and I don't believe it". Similarly, it seems that if I assert that dispositionalism is true, I must also concede that I don't believe it. The real puzzle surrounding Moore's paradox, however, is why it seems to be a paradox in the first place. After all, it's certainly possible to believe something that is false, or to not believe something that is true. It's only that we instinctively rebel against stating the two simultaneously. At any rate, Moorean paradoxes make us uncomfortable, so dispositionalism about belief should also make us uncomfortable, even if we can't quite explain why.

Second, if dispositionalism is incredible, does this mean that its negation is indubitable? The proposition that I exist seems to be indubitable, and for that reason Descartes took it to be true. The proposition that I don't exist seems incredible and therefore false. However, it's probably not the case that our psychological limitations always reflect the metaphysical facts. It seems that the incredibility of a philosophical position should factor against it in some way, but the argument for this is not straightforward. Again, I only have time to note the problem here.

Finally we might ask whether there is something special about an account of belief in particular that we want it to be believable. I'll leave these questions for us to ponder.

CONCLUSION

I have offered a number of counterexamples to Schwitzgebel's phenomenal, dispositional account of belief. These counterexamples follow a pattern that Schwitzgebel himself establishes in his example of Piotr. Schwitzgebel's point in the example is to show that when a person's self-ascriptions of belief diverge from the dispositional stereotype associated with that belief, the dispositional account denies possession of that belief in those cases. Furthermore, Schwitzgebel stakes the accuracy of his account on the claim that our unaided folk intuitions will ascribe the same beliefs as his dispositional account. I have attempted to show through my counterexamples that there are whole classes of belief for which this is not the case. It is in fact common for a person's behavior to diverge considerably from the dispositional stereotype associated with a belief, apart from that person's dispositions to self-ascription, while our folk intuitions nonetheless attribute a belief consistent with those self-ascriptions. For what it's worth, I have also pointed out that a dispositional account of belief seems to entail that no one actually believes in dispositionalism.

No doubt dispositionalism appears useful in cases where representationalism has difficulty, such as when it is hard to say whether or not someone "really" believes. It can also be a useful way to characterize subconscious beliefs. Unfortunately, where dispositionalism does not seem to be appropriate is right in the meaty middle. By and large, we seem to know what we believe directly, and we don't have to survey our behavior in order to do so.

Bibliography

Schwitzgebel, Eric. "A Dispositional Approach to Attitudes: Thinking Outside of the Belief Box." In *New Essays on Belief: Constitution, Content and Structure*, edited by Nikolaj Nottelmann (New York: Palgrave, 2013).

Schwitzgebel, Eric. "A Phenomenal, Dispositional Account of Belief." *Nous* 36 (2002): 249-275.

Schwitzgebel, Eric. "Knowing Your Own Beliefs." *Canadian Journal of Philosophy* 35 (2011): 41-62.

PRESERVING REALISM IN THE PHILOSOPHY OF SCIENCE

Alysha Kassam

INTRODUCTION

The realism debate in the philosophy of science is concerned with whether the entities and relations between objects presented by scientific theories are actual constituents of reality. Wholesale realism is the position that all or most entities posited by scientific theories exist while wholesale anti-realism is the position that all or most entities posited by scientific theories do not exist. P.D. Magnus and Craig Callender (2004) argue that the wholesale realism debate is not substantive since both sides commit a statistical fallacy called the base rate fallacy. Although Magnus and Callender specifically focus on the realist “No Miracles Argument” (NMA) and the anti-realist “Pessimistic Meta-Induction” (PMI) they conclude that wholesale realism and anti-realism are untenable positions in general. In this paper I will argue that Magnus and Callender’s deflationary position on the wholesale realism debate is not justified because the NMA and the PMI are not empirical arguments and therefore cannot be refuted by statistical fallacies. Moreover, when properly construed as non-empirical arguments, it is evident that the wholesale scientific realist and anti-realist disagree on the epistemic criteria for empirical success. Whereas the realist sets a high standard for a theory to be deemed empirically successful, the anti-realist sets a lower standard for empirical success. After illustrating how the debate is substantive in this way, I will argue why one should be a wholesale realist.

Magnus and Callender summarize the realism debate by means of two opposing inductive arguments—the realist no miracles argument and the anti-realist pessimistic meta-inductive

argument. The epistemic thesis of scientific realism is that scientific theories represent knowledge of the world. The no miracles argument supports the epistemic thesis of scientific realism as it is the claim that the empirical success of science gives us reason to believe that scientific theories accurately describe natural phenomena. According to the no miracles argument, it would be a ‘miracle’, or extremely improbable, that given the empirical success of science, scientific theories are not actual representations of reality. The anti-realist, on the other hand, argues that the history of science consists of many theories that exhibited empirical success and yet are now considered false. This pessimistic meta-inductive argument is meant to show that there is no way of confirming that current scientific theories are distinct from their predecessors. Moreover, given the large number of past scientific theories that were empirically successful yet false, we can inductively claim that our current successful scientific theories will similarly turn out to be false.

PMI AND NMA: BAD INDUCTIVE ARGUMENTS

In order to understand how the PMI and the NMA are misguided inductive arguments I will first clarify what makes a good inductive argument. An inductive argument is one in which universal statements are inferred from particular instances of observation. In order to make a good inductive argument that all Fs are Gs, it is important to consider the number of observations made of a F being a G as well as the features of the population of Fs one is observing. Consider the inductive claim, “All teenagers smoke weed.” In order to support this claim, it is important to have observed a large number of teenagers smoking weed. Moreover, it is important that the teenagers you have observed smoking weed are representative of the entire population of teenagers in general. So, for instance, you may claim, “All teenagers smoke weed,” having been in Amsterdam; yet this would not be representative of the entire teenage population because in places like Utah very few people smoke weed in general. Given these examples, the

justification of an inductive argument relies on its sample size (the number of individual observed instances of Fs being Gs) as well as the use of random sampling (every member of the population you are making inferences about should have an equal opportunity of being used in the sample).

Therefore, in order for the NMA and the PMI to be well-formed inductive arguments, they must use large random samples. Yet as Magnus and Callender point out, it is not the case that the PMI or the NMA use random sampling. The PMI is the inductive claim that, since most past successful scientific theories are false, most current successful scientific theories will end up being false as well. To support their claim, anti-realists point to past successful but false theories in science. These past scientific theories they reference are not randomly selected but, rather, are selected for the purpose of supporting their claim. The realist similarly makes their inductive argument based on theories that are intentionally selected to illustrate the empirical success of science. As a result, Magnus and Callender conclude that both the realist and the anti-realist do not use random sampling when selecting theories for their inductive arguments.

Another argument against the PMI and the NMA as inductive arguments is the claim that they commit the base rate fallacy. The base rate fallacy is the tendency to ignore the general, unconditioned probabilities, or “base rates,” when making inferences about the probability that an event will occur. For example, assume a diabetes diagnosis can be given by means of one simple blood test. Now suppose that the probability that x tests positive (Tx) given that they have diabetes (Dx) is such that $P(Tx|Dx) = 1$; or, it is certain that if x has diabetes then x will test positive. Nonetheless, it might be the case that someone who does not have diabetes will still test positive (a Type I error). Suppose the probability of a Type I error is five percent, so that $P(Tx|\neg Dx) = .05$. To commit the base rate fallacy would be to assume that if a patient tests positive for diabetes, they have a 95% chance of having the disease given that $P(Tx|Dx) - P(Tx|\neg Dx) = .95$. This is a fallacy because one must also consider the sample from which the people

tested are being drawn. One cannot assume that the $P(Ta) = 1$. For example, further suppose that in the region in which they are testing for diabetes, only 1 in 1000 people are in fact diabetic (diabetes is extremely rare in that region). Given the Type I error, 50/1000 people will test positive but not have the disease, while only 1/1000 will test positive and have the disease, so 51/1000 is the total number of individuals who will test positive. Out of these 51 people, only 1 will have the disease, so the probability that a patient who tests positive will have the disease is only 1 out of 51 or $P(Da|Ta) = .02$. This probability is significantly lower than .95. This example illustrates the false positive error, an error that occurs when neglecting the base rate.¹

I will now illustrate how the PMI and the NMA commit the base rate fallacy. As previously mentioned, the PMI gets its strength from drawing upon past theories that illustrated predictive success but are now considered false. The anti-realist claims that the conditional probability of a theory being false given that it is a candidate theory is 1. Although the extensive list of successful yet false theories seems compelling, the anti-realist using the PMI fails to realize that the base rate is critical for comparison. For example, suppose that the anti-realists gives us a list of 1,000 theories that were successful yet false. Although this large number of successful yet false theories seems to conclusively support the anti-realist claim, further suppose we find the base rate of false theories is 1/100. This means for every 100,000 theories only 1,000 of these theories support the PMI. This example is meant to show that until we have the base rate, we cannot determine the strength of the PMI argument.

From this example one can infer how the realist makes the same fallacy given that she similarly does not consider the base rate when making her argument. The realist wants to claim that empirical success is a likely indicator of the truth of a theory. To support this claim, the realist must provide a base rate that illustrates that the likelihood of a candidate theory being empirically successful and true is significantly greater than its being empirically successful and false. Yet in order to have a way of deter-

mining the approximate base rate of truth among current scientific theories, there must be an independent criterion for establishing the truth of a theory.² If an independent criterion for establishing the truth of a theory were known, then the NMA and the PMI would no longer be needed. Until the realist and anti-realist establish a base rate, both arguments rely on the assumption that either a significant proportion of our current theories are successful and true or a significant proportion of our current theories are successful and false. As a result, until a base rate is established, it is evident that both sides are relying on their intuitions concerning the ratio of true to false theories to support their arguments.

DEFLATING WHOLESALE REALISM

Since the PMI and the NMA fall victim to the base rate fallacy, Magnus and Callender conclude that they are not substantive wholesale arguments. They further argue that wholesale realism and anti-realism are implausible in general, as non-inductive wholesale realist and anti-realist arguments are also flawed. They consider the anti-realist “underdetermination of theory” argument. This is the claim that every theory has many empirically equivalent rivals and that there is no objective way of determining which of these empirically equivalent theories is the correct one (Magnus and Callender 2003, p. 27). Magnus and Callender argue that the underdetermination of theory argument should not be interpreted too strongly, because underdetermination can then be applied to all inductive inferences in general, which leads to complete skepticism. Yet the underdetermination of theory must be strong enough so that the competing theories do not refer to the same particulars. Therefore, the wholesale underdetermination of theory argument faces the problem of finding underdetermination strong enough so that different theories do not refer to the same entities and weak enough so that the argument does not lead to complete skepticism. Magnus and Callender also argue against the non-inductive wholesale realist “conjunction argument” which states that since scientific theories from different domains detect the same unob-

servable entities by different empirical means, it must be the case that these entities truly exist. Magnus and Callender point out that conjunction does not always apply to our most empirically successful scientific theories. To justify this objection, they point out that the space-time metric is dynamical in general relativity and not dynamical in quantum theory.³

Magnus and Callender argue that since the PMI and the NMA are fallacious, and other current wholesale realist and anti-realist arguments are implausible, one can safely claim that the wholesale realism debate in general is insubstantial. Yet Magnus and Callender argue that one can still be a realist or anti-realist about specific scientific theories by means of making retail arguments. Retail arguments are ones that establish the reality of particular entities posited by specific theories. So, for example, one can be a realist about electrons and an anti-realist about quarks. This type of piecemeal realism, or realism about particular entities, does not permit one to make generalizations about science as a whole but allows only for realist commitments on a theory-by-theory basis. Magnus and Callender argue that retail arguments do not commit the common statistical fallacies wholesale arguments do, as random sampling and base rates can be used for retail arguments.

RECOVERING THE WHOLESALE REALISM DEBATE

Let us consider the form of Magnus and Callender's argument. They first claim that since both the NMA and the PMI commit the base rate fallacy, both sides are talking past one another. They then infer that, given that the PMI and NMA are the main realist/anti-realist arguments, and that other wholesale realist arguments are not compelling, one can conclude that wholesale realism/antirealism is not a substantive debate and one should rather be a piecemeal realist. I will now discuss whether their premises support their conclusion. I will first argue that the NMA and the PMI are not empirical arguments, and therefore it is irrelevant whether they fall victim to the base rate fallacy. I will then illustrate that

when reformulated as non-empirical arguments, it is evident that both sides disagree about something substantive—the epistemological criteria for empirical success. Finally, I will conclude that there is good reason for being a wholesale realist.

Magnus and Callender claim that the NMA and PMI are bad inductive arguments because they cannot be empirically supported by probabilities corresponding to how many true versus false scientific theories there are. It is the case that when both arguments are construed as strictly empirical arguments there is not a substantive basis for their claims, as both positions cannot present to us why it is the case that a theory is likely to be true or false given a verifiable probability. Yet the PMI and NMA can be formulated as non-empirical arguments, and when they are formulated in this way it is evident that they disagree about something substantive. In order to see this point, let us examine the PMI as a *reductio ad absurdum*:

1. Success is a reliable test for truth (assumption)
2. Most current scientific theories are successful (premise)
3. So most current theories are true (MP 1,2)
4. If most current theories are true then most past theories are false (premise)
5. Past theories are false (MP 3,4)
6. Past theories were successful (premise)
7. Past theories are true (MP 1,6)
8. Since (5,7) illustrate a contradiction, success of a theory is not a reliable test for truth. (Mizrahi 2013, p. 2)

The fact that the PMI argument can be formulated as a *reductio* suggests that the worry about base rate fallacies can be avoided. The base rate fallacy can only apply to inductive arguments that make probabilistic claims without established base rates, but the PMI presented as a *reductio ad absurdum* is not an

inductive argument. Moreover, each of the premises in the *reductio* does not rely on an inductive argument. For example, premise 1, “Success is a reliable test for truth,” is assumed by the anti-realist in order to derive the *reductio ad absurdum* and demonstrate that this claim is untenable. Premise 2 is a part of the realist thesis as well and is also assumed in the PMI *reductio*. Line 3 is derived from the conjunction of assumption 1 and premise 2. Premise 4 follows from the fact that current scientific theories and past scientific theories make different theoretical claims, and therefore must have different truth-values. Modus ponens from lines 3 and 4 gives us line 5. Premise 6 follows from the criteria of empirical success anti-realists posit (more on this later but for the present purposes it is important to note that premise 6 is entailed by the anti-realist’s criteria for empirical success and is not an inductive claim). Line 7 is a consequence of premise 1 and premise 6. Finally, lines 5 and 7 illustrate a contradiction, which supports the anti-realist conclusion that empirical success is not a reliable test for truth.

It is now clear that the PMI reformulated as a *reductio ad absurdum* is not an inductive argument and does not have premises that depend on inductive arguments. Although the PMI can be reformulated as a non-empirical argument, one may object that the arguments the realist and the anti-realist make concerning whether the entities posited by scientific theories exist are in fact empirical arguments, as these arguments hinge upon empirical data. So, for instance, anti-realists point to theories such as phlogiston theory to support their argument. Phlogiston theory postulated that when substances are burned, they release the element phlogiston into the atmosphere and it was empirically successful. Phlogiston theory was shown to be false because scientists discovered that when certain metals like magnesium burned, they increased in mass, contrary to the predictions of phlogiston theory. According to the theory, releasing phlogiston should cause the metal to lose mass. The anti-realist thus uses phlogiston theory to support the NMA because phlogiston theory enjoyed empirical success yet was strictly false. More importantly, if the anti-realist cites the results

of the experiment that proved phlogiston theory false, the anti-realist seems to be using empirical data to support her claim. Yet what is salient here is that although the truth or falsity of the premises relies in part on empirical data, the conclusions the realist and anti-realist derive from these premises are not empirical. The anti-realist concludes that, based on empirical data, phlogiston theory is false and phlogiston does not exist. This conclusion is not an empirical one. If the anti-realist wanted her claims to be strictly empirical, then the anti-realist could only conclude that when certain metals are burned they increase in mass. She could not make the further claim that phlogiston does not exist or that phlogiston theory is false. This example is meant to show that the anti-realist and realist claims cannot be argued against by means of probabilistic fallacies because they are not empirical arguments. Rather, they are philosophical claims about what are the constituents of reality and whether scientific inquiry leads to knowledge. Therefore Magnus and Callender are not justified in claiming that the wholesale realism debate is not substantive based on the fact that the NMA and PMI commit statistical fallacies.

If we refer back to the *reductio* we can see that the realist and anti-realist disagree about whether empirical success is a reliable test for truth. The realist believes that scientific theories are so empirically successful that it would be miraculous if they did not refer to actual constituents of reality. The anti-realist argues that empirical success is not a reliable indicator of truth, given that past theories were empirically successful and are now considered false. Yet as argued by Magnus and Callender, the realist and anti-realist are talking past each other because they refer to different scientific theories to support their claim—whereas the realist uses current successful theories, the anti-realist uses past false theories. Contrary to Magnus and Callender's claim, I contend that it is the case that the realist and anti-realist disagree on something substantive, namely premise 6, or whether it is the case that past theories exhibit the same type of empirical success that current theories do. The realist undermines the anti-realist PMI argument by claiming that past scientific theories do not illustrate the proper

empirical success necessary for a theory to be considered a good scientific theory, whereas anti-realists argue that past theories exhibit empirical success similar to current theories. Therefore, the realist and the anti-realist are not talking past one another when referring to empirically successful theories to advance their claim. Rather, they disagree on the epistemic criteria for what constitutes empirical success. Anti-realists such as Larry Laudan suggest that a theory is empirically successful if it “has functioned in a wide variety of explanatory contexts, has led to several confirmed predictions and has been of broad explanatory scope” (Laudan 1981, p. 22). However, realists such as Stathis Psillos argue that the criteria for empirical success should be set higher than simply fitting the facts. Instead, empirical success should be defined by a theory’s ability to make novel predictions that are testable (Psillos 1999, p. 105). This illustrates that the realist and anti-realist disagree on a substantive epistemological criteria for empirical success. The realist requires a higher standard for a theory to be considered empirically successful than the anti-realist.

WHY WHOLESALE REALISM IS THE CORRECT POSITION

Testable novel prediction is an important criterion for empirical success because otherwise theories could be designed arbitrarily to fit the empirical results. Therefore, current scientific theories are successful in a way that past scientific theories were not in that they are able to make predictions of phenomena that have not been previously observed. For example, consider Einstein’s relativity theory. Relativity theory made the prediction that two clocks, one accelerating around the Earth’s surface, and the other stationary, would record different times. This novel prediction was confirmed after scientists conducted the experiment. Now compare Einstein’s theory with Ptolemy’s geocentric model of planetary orbits. Ptolemy’s ‘scientific’ theory led to some confirmed predictions concerning the placement of the stars and had a general explanatory value, as it was able to explain why the

stars and planets seemed to revolve around the Earth. The anti-realist claims that Einstein's relativity theory and the geocentric model are both similarly empirically successful. Although the geocentric model led to confirmed predictions, the theory did not make predictions about features of reality that had yet to be observed as did the theory of relativity. Moreover, the geocentric model, like many of the past theories the anti-realists refer to in their PMI, is a scientific theory that simply fit the observations but did not accurately predict future states. Therefore, the past scientific theories anti-realists refer to do exhibit the same level of empirical success as current theories do.

One can also attribute the superior empirical success of current scientific theories to the technological advancements of our time. Current technology allows us to verify our theories with greater precision and has also enabled us to confirm the existence of unobservable particles posited by previous theories. For example, the geocentric model of planetary orbits was confirmed by observing how planets appeared to revolve around the Earth. This theory was 'well supported' because it seemed to be empirically verified by the means of observation available at the time. It was not until the advent of the telescope that Galileo was able to discover that the moons of Jupiter orbited around Jupiter, not the Earth.⁴ Therefore, it was the advent of the telescope that allowed us to empirically verify that the geocentric model of planetary orbits was false and the heliocentric model was correct.

Not only does technology allow us to disconfirm false theories and verify correct ones, it also allows us to observe constituents of reality that were previously unobservable. The existence of basic units of matter, or atoms, had long been posited by scientific theories, but until recently this could only be verified through indirect means.⁵ Yet the advent of the scanning tunneling microscope in the 1981 allowed scientists to image individual atoms and thus confirm their existence. As a result, current scientific theories are more empirically adequate than past scientific theories because technological developments allow us to both verify these theories with more precision and confirm the existence of previously

unobserved particles.

Yet in the same argumentative style of the PMI itself, the anti-realist may respond to this claim by pointing out that our current technology will undoubtedly be replaced in the future just as previous technological developments have been replaced. So, the anti-realist may argue, there is no way of substantiating the claim that improved technology leads to true theories. In fact, the history of science shows us that the technology currently used to confirm scientific theories will most likely become obsolete in the future. Therefore, it is likely the case that what we define as empirically successful today and the method by which we test a theory's empirical success will be considered inadequate in the future. The anti-realist can thus argue that advancements in our scientific methodologies do not make it the case that current theories are even approximately true. Rather, the future advancement of scientific methodologies and technology will probably lead to the disconfirmation of current theories.

The problem with this line of reasoning, which also underlies the motivation for the PMI in general, is that it is unclear why past theories and current scientific theories must differ in truth-values. The anti-realist assumes that current theories and past theories differ in significant enough ways such that the entities and explanations posited by past theories have been disconfirmed and replaced by current scientific theories. Although past theories have been replaced by current theories, this does not show that past theories are strictly false. Rather, there may be certain constituents of past theories that refer to real entities. Arguably, the correct way of looking at the pursuit of science is that current theories, in light of technological progress, more precisely describe what these entities are. For example, consider the developments in the empirical study of the mind. Freud's psychoanalytic theory that explains human psychology by means of unconscious mental processes influenced by an individual's libido, or sexual drives, is a past theory of human psychology. Freud's theory was deemed successful as it assisted individuals with psychological disorders specifically pertaining to depression and anxiety. Although

Freud's theory may seem strictly false, the notion that unconscious mental processes exert control over human behavior is a constituent of his theory that is still accepted by psychologists today. The reason psychoanalysis was successful for patients with depression and anxiety can be attributed to the specific claim that unconscious mental processes influence behavior. Recent studies have shown that depression and anxiety symptoms improve once patients know that unconscious processes cause their depression and anxiety. Yet technological developments have shown us that these unconscious mental processes are not due to the libido but rather, are caused by chemical imbalances in the brain. This example is meant to illustrate that when past theories are replaced by new theories, it is not always the case that past theories are now strictly false. Rather, the aspects that gave these past theories their empirical success are retained in current theories. Current theories simply describe the same constituents of reality more precisely.

CONCLUSION

Contrary to the claims of Magnus and Callender, the realism debate in the philosophy of science is a substantive argument. Magnus and Callender wish to deflate the realism debate by showing that the PMI and the NMA are bad inductive arguments. I have argued that this point is irrelevant because the PMI and the NMA are not in fact inductive arguments. This is evident when one sees the PMI argument reformulated as a *reductio ad absurdum*. When written as a *reductio*, it is clear that none of the premises of the PMI turns on an inductive argument. Although it is the case that the realist and anti-realist use empirical data in order to support their claims concerning the truth-value of scientific theories, the conclusions they draw are philosophical in nature. After arguing that the realism debate is substantive, I have attempted to clarify what issue the debate centers on. The PMI and NMA arguments reveal that the realist and anti-realist disagree on the proper criteria for empirical success. Given that this is the basis of their debate, I claim that the realist correctly sets a high standard of empirical

success for a theory to be considered scientific. I further claim that current theories are more empirically successful than past theories given that technology allows us to test theories more extensively and to observe previously unobservable entities. Although the realist may object that what we deem empirically successful now is subject to change once technology further develops in the future, I argue that this does not affect the realist thesis, for one can construe future scientific theories as more precise descriptions of objects described by current theories.

Notes

1. The false positive error is the same as the type I error in statistics. It occurs when there is an incorrect rejection of a true null hypothesis. The false positive error can be contrasted with the false negative error. The false negative error (or type II error) is a failure to reject a false null hypothesis.
2. In order to support the realist (or anti-realist) inductive claim, a thorough survey of all past theories would need to be made in order to establish the likelihood that a candidate theory will turn up true or false. If we could find an objective ratio of true to false theories, the way in which we found this ratio would require a way of establishing the truth or falsity of each theory. If we had such grounds to establish a theory as true or false, the PMI and NMA arguments would be unnecessary.
3. This is just one example in which implications of quantum mechanics and general relativity theory are logically inconsistent. Another inconsistency lies in the fact that quantum mechanics is a probabilistic system where as general relativity is deterministic. Further, one implication of quantum mechanics is non-locality or action at a distance, which also contradicts general relativity. What this is meant to illustrate is that the conjunction argument is flawed because empirically verified theories can be logically inconsistent.
4. The telescope was invented in 1609. Galileo Galilei used it to observe that Jupiter had moons. This violated the geocentric model, which required that all celestial entities orbit the Earth.
5. In the history of science, reference to atoms or basic units of matter date back to ancient India in the 6th century BCE and ancient Greece in the 5th century BCE.

Bibliography

Magnus, P.D. and Callender, Craig. (2004) "Realist Ennui and the Base Rate Fallacy," *Philosophy of Science* 71, pp. 1-40

- Mizrahi, Moti. (2013) "The Pessimistic Induction: A Bad Argument Gone Too Far," *Synthese* 109(15), pp. 1-19
- Laudan, Larry. (1981) "A Confutation of Convergent Realism," *Philosophy of Science* 48, pp. 19-49
- Psillos, Stathis. (1996) "How Science Tracks Truth," *Philosophy of Science* 21, pp. 101-114

DEFLATING DEFLATIONARY META-ONTOLOGY

Nicole Lavoie

I. BACKGROUND AND MOTIVATION

“What is there?” is traditionally understood as the ultimate ontological question, which has yet to be fully answered. Ontology traces back through the history of philosophy and is still provoking discussion among philosophers. More recently, the meta-ontological question, “what do we mean when we ask what there is?” has stimulated different debates regarding the status of our metaphysical questions. Claims that the problems metaphysicians are working on are not substantive, or can be reduced to a difference in language-use have developed as full-fledged meta-ontological views. I will call those who advocate this sort of view as deflationary meta-ontologists. I will call those who maintain that the work being done in metaphysics is substantive meta-ontological realists.¹

This paper will give a brief exegesis of two of the most well-known and convincing arguments for adopting a deflationary meta-ontological view. In Section II I address Eli Hirsch’s proposal for quantifier variance and the possibility of alternative languages within metaphysical debates. In Section III I address Amie Thomasson’s argument that the conceptual content of our terms generates analytic entailments that tell us which objects we are ontologically committed to. In sum, I argue on behalf of the meta-ontological realist that the methodology of each of these views are flawed, and neither of these deflationary views gives us substantial reason to believe that ontological debates are merely verbal or can be solved by amending our language-use.

II. QUANTIFIER VARIANCE AND ALTERNATIVE LANGUAGES

Eli Hirsch (2002) explores the doctrine² of quantifier variance and the consequential threat it poses to ontological realism. The doctrine of quantifier variance claims that there are possibly multiple ways to interpret “there exists an x” such that the existential quantifier can be interpreted in one way to make the sentence true and in another way to make the sentence false. Depending on how one interprets the existential quantifier, the truth conditions for the sentence “there exists an x” change and this results in disagreement about what objects exist. Hirsch elucidates this idea in terms of the debate about mereology.

The mereologist’s ontology includes not only ordinary objects from a common sense view, such as the Eiffel Tower and Obama’s nose, but also allows for the existence of objects that seem contrary to common sense, like extraordinary objects such as an object composed of Obama’s nose and the Eiffel Tower. It’s obvious that if one were to ask a person outside of the metaphysics room if such extraordinary objects (i.e. an object composed of Obama’s nose and the Eiffel tower) exist, ordinary fluent English speakers would say they don’t, or ask where they might find such a thing. Hirsch differentiates the language of fluent ordinary speakers of English as the A-language³ and the mereologist’s language as the M-Language because they are not using the term “existence” in the same way, so they are speaking different languages. Hirsch’s idea is that we should adopt “ordinary language philosophy... in which the disputed sentence admits of only one relevant meaning in plain English, and one of the disputants is saying something that—interpreted in plain English—is trivially absurd” (Hirsch 2010, p. 62).

In a later paper, Hirsch (2010) makes explicit that he believes some debates in metaphysics are substantive, but there are some disputes that are merely verbal such as the dispute between the mereologist and anti-mereologist. He defines a debate in ontology as “merely verbal” if the disputants are using language differ-

ently such as using “existence” in a nuanced way that allows for extraordinary objects to exist. He suggests that the correct way of nailing down the meanings of words is to determine how ordinary speakers of language use them, and if we accept the ordinary language, the “merely verbal” metaphysical disputes will dissolve.

There are two points I would like to bring up in response to Hirsch’s view.

Firstly, it is not at all obvious that the metaphysical parties involved would admit that the opposing view is correct in their own respective language. In fact, I don’t think they would say that they were speaking different “languages” at all. Ontologists are concerned with the truth of what there is. Surely the common sense mereologist and the anti-mereologist would not agree to disagree that the other party’s view would be right if they meant the same thing when they say extraordinary objects exist. The mereologist and the anti-mereologist are arguing about the way the world *really* is, and what objects *really* exist. It is not in the style of philosophers to grant that the advocates of the opposing view are correct on their way of interpreting language, especially when the central issue is theoretically trying to explain the nature of the world.

Secondly, I share the sentiment in Peter Van Inwagen’s (1998) paper that existence is univocal (and furthermore, this claim can be supported by the existential quantifier in logic, which does not allow for quantifier variance at all). He makes an analogy between the univocacy of numbers and the univocacy of existential claims:

No one would be inclined to suppose that number-words like ‘six’ or ‘forty-three’ mean different things when they are used to count different sorts of objects... To say that unicorns do not exist is something like saying that the number of unicorns is 0; to say that horses exist is to say that the number of horses is 1 or more. And to say that angels or ideas or prime numbers is greater than 0... (Van Inwagen 1998, p. 236)

In the spirit of Van Inwagen’s assessment of univocacy of

existence, I don't believe that there even are alternative languages within English that the mereologist and the anti-mereologist speak. The mereologist claiming extraordinary objects exist and the anti-mereologist claiming extraordinary objects don't exist means they have different senses of what "exists" means—they are using the same language and the same notion of existence but disagreeing about whether or not there are such things as extraordinary objects. If we reject the possibility of quantifier variance and accept that existence is univocal, to say something exists just means "there is some object composed of the Eiffel Tower and Obama's nose." Or, according to Van Inwagen's numerical analogy, either the number of objects composed of the Eiffel Tower and Obama's nose is 0 (if you're an anti-mereologist), or the number of objects composed of the Eiffel Tower and Obama's nose is 1 (if you're a mereologist). The mereologist and anti-mereologist are not disagreeing about what it means for an object composed of the Eiffel Tower and Obama's nose to exist; they are arguing about whether or not said object does or does not exist in the same language and on the same notion of existence.

III. REFERENCE AND CONCEPTUAL CONTENT

In the same vein as Hirsch's arguments, Amie Thomasson (2007, 2009) argues for a much stronger version of deflationary metaontology. Thomasson (2007) sets out to defuse many popular eliminativist arguments that deny the existence of ordinary objects, and in turn provides reasons for accepting the common sense view that objects like tables and chairs exist. She argues that the debates between the eliminativist and common sense ontologist about the existence of ordinary objects are really semantic disagreements about the content of sortal terms (i.e. disagreements in language about the appropriate employment of terms of a certain sort) as opposed to disagreements about the nature of objects themselves. She proposes a method for getting rid of the semantic issues and getting back to the actual metaphysical problem—whether ordinary objects exist.

On her view, if the disputants in metaphysical disagreements do not agree on the conceptual content of the sortal terms for objects in question, then they are talking past one another and therefore the debate is not substantive. Additionally, if the disagreement involves terms with indeterminate reference (where the conceptual content is not specified), then the debate is a result of the disputants trying to answer unanswerable questions, where "... [a] question is 'unanswerable' if no straightforward answer to it, stated in the same terms as the original question, is truth-evaluable—where this failing is in principle; not a reflection of mere epistemic shortcomings but of deficiencies in meaning" (Thomasson 2009, p.445), then again the debate is not substantive.

It's important to note that Thomasson's view and the importance of reference traces back to the Quinean criterion of ontological commitment. The main point of this criterion pertinent to this discussion is that we are ontologically committed to the objects referred to by nominal singular terms, if and only if we can successfully ground reference.⁴

Thomasson's suggestion for remedying deficiencies in meaning is as follows: when a speaker refers to an object, she should have in mind the general category or sort of object she is meaning to refer to (e.g. mass of matter, part of an object, an entire object), but she does not need to have any specific characteristics of the referent in mind beyond the exemplification of the sortal property (sortal properties would be something like an object exemplifying the property of table-ness). Each sortal term is conceptually associated with 'frame-level application conditions' and 'coapplication conditions' that prescribe its rules of use in language. Application conditions stipulate the criteria for initially applying a sortal term. Coapplication conditions stipulate under what conditions a term can be reapplied to the same object at a later time. To ground reference the speaker firstly must have in mind a general, conceptual sort of entity they intend to refer to (such as a part of an object, the particles composing the object or the object as a whole). Secondly, the speaker must be able to empirically determine whether the candidate referent satisfies the

associated conditions. If the conditions are satisfied, reference is grounded. If they are not, then the term fails to refer (Thomasson 2007, p. 42).⁵

There is an important distinction to be made between *application and coapplication conditions* and *existence, identity, and persistence conditions*. Application and coapplication conditions apply to language and employing the use of sortal terms. Existence, identity, and persistence conditions apply to actual physical things (i.e., the sortals themselves). Having conceptual content for a sortal term alone does not yield answers to metaphysical questions. However, there is a pivotal relationship between the criteria for reference and the criteria for approaching metaphysical questions:

The most important consequence of the hybrid approach to reference defended above is that the most basic conditions of existence, identity and persistence for the objects we refer to are discoverable by a kind of conceptual analysis, and the most basic claims about these conditions are analytic.... Criteria of application and coapplication are rules for the proper employment of terms of our language... nonetheless all metaphysical claims must be expressed using language, and these rules for the proper application and coapplication of our terms play a central role in establishing the truth conditions for metaphysical claims that use those terms in claims about when individuals, or objects of a given sort, exist, are identical, and persist. (Thomasson 2009, p. 54-55)

So, metaphysical questions such as “does x exist?” or “how many x’s are in the room?” cannot be answered without knowing what a speaker is intending “x” to refer to. This includes recognizing the appropriate categorical conception, relevant application and coapplication conditions, and empirically determining whether there is an object that satisfies the conditions. Thomasson concedes there can be controversy over conceptual content, as she states, “There may be differences in the conceptual analysis of

what the term, as ordinarily used in English, means, and so about what it would take for there to be, for example, tables, numbers...” (Thomasson, 2007, p. 199), and suggests that the new tasks of metaphysicians should be to clarify the appropriate conceptual content of our terms (Thomasson, 2007, p. 199).

Thomasson does not give a compelling argument as to why conceptual content so easily generates the analytic entailments needed to posit the existence of ordinary objects. Perhaps there are multiple, or “accepted” and “disputed” descriptions for ordinary objects, whereby the accepted description is one that the eliminativist and common sense ontologist would both accept, and the disputed description is one that only the common sense ontologist would accept.⁶ The accepted conceptual content of the term ‘table’, as Thomasson previously suggests, would be something like, ‘particles arranged tablewise’. If this condition were met empirically, the eliminativist would have to concede that tables do exist. This point is similar to Hirsch’s idea that if the disputants each believe the other is using existence in a different way than they are, then the parties would have to agree to disagree. But if the conceptual content of a “table” were something like ‘the fusion of particles arranged tablewise’ the eliminativist would deny that this condition is ever met and could maintain the position that tables do not exist.

One issue with Thomasson assuming the conceptual content of ‘table’ is the accepted description is that it is not at all obvious that a competent speaker would give ‘particles arranged tablewise’ as the application condition for ‘table’. In light of Thomasson’s stipulation that competent speakers need not be able to explicitly state what the conceptual content is; they only need to be able to recognize that the term is properly applied if certain empirical conditions are met. A competent speaker who knows the meaning of the term “table” would probably not infer that the criteria needed to properly employ the term “table” include “particles arranged tablewise”. In fact, a competent speaker might say (if there are such things as tables) the application condition would be something like, “a piece of furniture with one or more legs and a flat

surface”. Thomasson hasn’t given any good reason as to why one should think the conceptual contents of terms for ordinary objects come from their accepted descriptions, rather than their disputed (or alternative) descriptions. The conceptual content could reasonably come from something other than the accepted description; one might think the conceptual content of a term like ‘table’ is ‘particles arranged tablewise’, ‘the fusion of particles arranged table wise’, ‘an object with a flat surface and one or more legs’, etc. Consequently disagreements about conceptual content do not provide sufficient reason to regard debates in which the disputants have different application conditions as semantic or “talking past one another”, as opposed to reflecting differing views about the way the world is, which amounts to a substantive disagreement.

The response Thomasson gives to this anticipated objection is that conceptual content should come from the accepted description because the disputed description would imply that there is some object or thing in addition to particles arranged object-wise. An eliminativist wouldn’t deny that there are particles arranged tablewise, but they would deny that there is an object that the particles arranged tablewise compose. She says, “All hopes for reviving the debate between the eliminativists and common sense realists about ordinary objects, then, rely on the idea that there is some neutral use of ‘thing’ on which the eliminativist may deny, and the realist affirm, that there is some thing composed by the properly arranged particles” (Thomasson 2007, p. 158). Thomasson notes that Van Inwagen endorses this type of eliminativist view. He writes:

There are certain properties that a thing would have to have to be properly called a ‘table’ on anyone’s understanding of the word, and nothing has these properties. If anything did have them, it would be real, a true object, actually, a thing, a substance, a unified whole, and something more than a collection of particles. But there are no tables. (Van Inwagen 1990, p.100)

Thomasson’s issue with the neutral use of ‘thing’ or ‘object’

in questions such as “is there a thing here?” or “how many objects are in the room?” is that the questions are not truth-evaluable because the terms are not sortal in the sense that they don’t come with the application and coapplication conditions that allow us to disambiguate reference, insofar as we mean to talk about the object itself, the parts of the object, etc. This is directly related to Thomasson’s second deflationary claim, that metaphysicians are trying to respond to defective or unanswerable questions; because without determinate reference, questions involving the neutral use of ‘things’ or ‘objects’ are unanswerable, and any claims using the neutral-use of ‘thing’ or ‘object’ are not truth-evaluable. Denying the “sortal-neutral” use of ‘thing’ or ‘object’ does much of the work in her deflationary metaontology. Denial of a “sortal-neutral” use of ‘thing’ or ‘object’ should convince us that conceptual content should come from an accepted description (because the alternative requires the use of this sortal-neutral ‘thing’) and, by denying that general questions involving the terms ‘object’ or ‘thing’ are defective, the metaphysical disagreements are actually semantic issues. I propose that there is a way to understand this “sortal-neutral” use of ‘thing’ or ‘object’ and that making sense of the “sortal-neutral” ‘thing’ or ‘object’ gives us good reason to reject Thomasson’s deflationary metaontology.

In Thomasson’s explanation of how analytic entailments are supposed to work, she notes that there can be a hierarchy of sortal terms, such that:

The application of a sortal S1 to any entity x may analytically entail that another sortal, S2, also applies to x, as, for example, the application of “dog” to Fido guarantees the application of “animal” to Fido. Where this occurs, I will say that S2 is a “genus-sortal” with respect to S1, and S1 is a “species-sortal” with respect to S2. (Thomasson 2007, p.41)

On this view, it seems plausible to suggest that “sortal-neutral” terms like ‘thing’ and ‘object’ are in fact not sortal-neutral at all, but could be what Thomasson calls “high-level genus-sortals”

or “categorial terms.” The only issue remaining seems to be that Thomasson requires sortal terms in the hierarchy to be associated with application and coapplication conditions. However, Thomasson claims different sortal terms may be of the same category, “e.g. ‘dog’, ‘cat’, and ‘horse’ all are of the category ‘animal’; ‘fork’, ‘table’, and ‘shoe’ are all of the category ‘artifact’” (2007, p. 42), and furthermore, “...the application conditions for the categorial term are guaranteed to be fulfilled provided those for any of its species-sortals are” (2007, p.42). So, even if the higher-level categorials do not come with the type of conceptual content associated with sortals, then by virtue of any sub-sortals that fall under the categorial ‘object’ satisfying a more specific application condition and by Thomasson’s own notion of analytic entailments, it follows that if the criteria for “table” are met, then so are the criteria for “thing” or “object”. It seems plausible that “thing” and “object” are not neutral terms but can be treated as high-level genus-sortal terms that all ordinary objects would fall under. Thus, the meta-ontological realist can appeal to a neutral use of the terms “things” or “objects” acting as a categorial that does not need application and co-application conditions so long as when we proceed down the hierarchy there are some sub-sortals that do have application and co-application conditions. Then, we are back to the question as to why there is good reason to think the conceptual content of a term should come from the accepted description rather than the disputed or alternative description. Thomasson has yet to provide an adequate answer to this question.

IV. CONCLUSION

The discussion in this paper explicates two well-known deflationary meta-ontological views proposed by Eli Hirsch and Amie Thomasson who claim that certain debates in metaphysics can be solved or dissolved by clarifying or stipulating our language-use. I have argued on behalf of the meta-ontological realist that neither of these deflationary positions is successful in showing metaphysical debates to be non-substantive. Metaphysical questions about

the existence of various types of objects cannot be reduced to the way the disputants use language. Hirsch relies on the possibility of quantifier variance and alternative languages to dissolve the debate between mereologists and anti-mereologists on the existence of extraordinary objects. I have argued that the possibility of quantifier variance and alternative languages does not rid of the issue between the mereologist and anti-mereologist on the basis that existence is univocal and ascribing different meanings to our terms is not where the problem lies. Similarly, Thomasson develops a view that requires we use the agreed upon conceptual content of the terms for objects to determine whether or not an object exists, in virtue of arguing in favor of the common sense ontologist and deflating the eliminativist's position. I have argued that her solution of subscribing to the accepted conceptual content of our terms for ordinary objects does not have sufficient argumentation. There are multiple ways to construe the conceptual content of a term for an object, and this does not mean that the disputants are talking past one another. In both cases, the metaphysicians are not disagreeing because of the way they are using language; they disagree about the way the world is and what ontology they believe to be true.

Notes

1. I am indebted to David Chalmers for much of the terminology used in this paper. See: "Ontological Anti-Realism" in *Metametaphysics: New Essays on the Foundations of Ontology*.
2. Hirsch concedes that the idea of quantifier variance was a central thesis of Hilary Putnam's view. He makes clear that his paper is not an exposition of Putnam's view, but rather his own position of quantifier variance, regardless of the fact that for the most part Putnam and Hirsch's view share the same overall sentiment.
3. In Hirsch's later paper "Ordinary and Alternative Languages" (2009) he runs the same kind of argument using the debate between perdurantists and endurantists. He calls the language of the Perdurantist E-English and the language of the endurantist P-English. For the purpose of this paper, I will only discuss the earlier view about mereology and its A-Language and M-Language.
4. For a more elaborate discussion on the Criterion of Ontological Commitment and metaphysics, see Matthew Davidson and Tony Roy's paper "New

- Directions in Metaphysics” in the *Continuum Companion to Metaphysics* (2012).
5. An important feature of a term’s conceptual content is that competent speakers do not need to be able to explicitly state the application conditions of a term, they only need to be able to have a tacit understanding of when it is appropriate to use the term. “...for a term to have application conditions is for competent speakers to be able to evaluate, with respect to various hypothetical situations... whether or not the term would apply—not for competent speakers, or philosophers, to provide an explicit statement in other terms of what those conditions are” (Thomasson, 2007, p.44).
 6. The terminology of “accepted” and “disputed” descriptions come from J. Schaffer “The Deflationary Metaontology of Thomasson’s Ordinary Objects” *Philosophical Books* 50 (3):142-157 (2009)

Bibliography

- Chalmers, D. (2009). “Ontological Anti-Realism” in *Metametaphysics: New Essays on the Foundations of Ontology*. Oxford University Press.
- Davidson, M. and Roy, T. (2012) “New Directions in Metaphysics” in *The Continuum Companion to Metaphysics*. New York, NY: Continuum International Publishing Group.
- Hirsch, E. (2002). “Quantifier Variance and Realism” *Philosophical Issues* 12 (1):51-73
- Hirsch, E. (2009). “Ontology and Alternative Languages” in *Metametaphysics: New Essays on the Foundations of Ontology*. Oxford University Press.
- Quine, W. (1980). “On what there is” in *From a Logical Point of View*. Cambridge, Massachusetts: Harvard University Press
- Schaffer, J. 2009. “The Deflationary Metaontology of Thomasson’s Ordinary Objects” *Philosophical Books* 50 (3):142-157
- Thomasson, A. 2007. *Ordinary Objects*. Oxford University Press.
- Thomasson, A. 2009. “Answerable and Unanswerable Questions” in *Metametaphysics: New Essays on the Foundations of Ontology*. Oxford University Press.
- Van Inwagen, P. 1990. *Material Beings*. Ithaca, NY: Cornell University Press.
- Van Inwagen, P. 1998. “Meta-Ontology” *Erkenntnis* 48 (2/3) 233-50

THE NORMATIVE CRITIQUE OF LAW'S CLAIM TO AUTHORITY: A BAD CONCEPT OF LAW

Pedro Viramontes, Jr.

The thesis that law claims authority (LCA) has been the subject of much discussion in the philosophy of law, specifically within the legal positivist camp. Joseph Raz (1985) originally proposed LCA to capture a distinctive feature of law: legal regimes purport to have the status of a moral authority. LCA, as Raz presented it, left legal scholars scratching their heads, trying to make sense of it: it is apparent, if we reflect upon the legal regimes that populate history, that many legal regimes do not possess the authority being claimed. It is clear that, at very minimum, there is tension between LCA and this observation. Philip Soper (2002) clarified this apparent tension and argued it should lead one to reject LCA—this is the so-called normative critique of LCA. The normative critique argues that if our concepts suffer from some sort of internal incoherence caused by tension either between descriptive facts or normative facts, then one of these facts must be excluded from the concept in order to retain some coherence. Bas Van Der Vossen (2011) has argued in defense of LCA and shown the objection raised by normative critique to be non-threatening. In this paper, I will argue that Van Der Vossen has failed to deflate the normative critique.

Before I present my arguments, I must clarify the important aspects of the debate relevant to this talk. Section I will be devoted to explicating what LCA means. I will focus on the authority being claimed and the nature of the obligations that legal regimes purport to create. Section II will present the role that LCA has in our understanding of law. Namely, that LCA describes the practices of legal regimes—necessarily law claims authority—and this

points towards a necessary feature of law. In section III, I will give an account of the normative critique of LCA and show the role philosophical anarchism (PA) has within it. PA brings forth a normative fact that has the consequence of denying that can ever make a justified claim to authority. As such, all claims to authority are necessarily false. In section IV, I will run through Van Der Vossen's second interpretation of the normative critique—the so-called '*bad*' concept of law—and show that the arguments given against it fail. He presents two cases in which one wouldn't want to reconstruct concepts, and therefore the normative critique lacks teeth. I will show that the two purportedly analogous cases are not actually so; and as such, do not threaten the conclusions imposed by the normative critique. Finally in section V, I wish to make some remarks about where this leaves one if the normative critique remains standing.

I. WHAT DOES 'LAW CLAIMS AUTHORITY' MEAN?¹

A legal regime is a form of practical authority--its "commands or pronouncements give us distinctive reasons to act in accordance with them" (Estlund 2012, p. 23). If a legal regime issues a directive to *x*, its subjects have reasons to *x*, where these reasons "are both preemptory and content independent" (Ibid, p. 23). More precisely, to say that a legal regime, L, is an authority is to say: L has the ability to impose on its subjects, S, obligations that preempt other obligations held by S and that are to be observed regardless of their content, when L dictates it. From this conception of authority, we can define LCA in terms of its corollary: Law claims authority if and only if "law claims that its subjects are obligated to comply with its requirements when it holds them so obligated" (Van Der Vossen, p. 485).

There is an important distinction to be made between *de facto* authority and *de jure* authority. De facto authority is held by any regime if and only if it *actually* exercises authority over

its subjects. *De jure* authority is held by a regime if and only if it is *justified* in exercising authority over its subjects—*de jure* authority is legitimate authority. Legal regimes undoubtedly are *de facto* authorities: they actually have authority over us. There is no debate regarding this issue. But we want legal regimes to be more than merely *de facto* authorities, for this implies that they are merely coercive systems—i.e. legal regimes only have authority because if one were to do contrary to the directive of the regime, one would be subjected to penalization by the state. If law can be a *de jure* authority, then it would distinguish itself from such systems because it would attain a certain moral character. The type of authority being claimed in LCA is *de jure* authority—legal regimes purport to be justified authorities. (From hereon, I will simply refer to *de jure* authority as authority.)

Returning to the issue of obligations, it is still unclear what is the nature of the obligations imposed by law. If they are *legal* obligations, this would be a trivial fact as this form of obligation is analogous to one created by a game. When one plays Monopoly, for instance, one acquires certain *Monopoly* obligations while playing the game—e.g. don't steal from the bank, don't cheat, collect \$200 upon passing 'Go', etc. This obligation ceases to exist when game ceases to be played because it is held so long as one is in the game. Rather, law claims to impose obligations that are binding even when one is outside the game; that is, the obligations imposed on subjects are *moral* obligations. That is, the obligations generated by law are on par with other obligations generated by other normative systems, and affect the overall balance of a subject's practical reasons in the same manner. With this in mind, we must now turn to the content of the obligation: law's obligation is an "obligation to obey legal directives because they have the status of law" (Himma 2001, p. 9).

To say that the obligations imposed by law are content independent is to say that an authority qua authority can issue a directive to perform *x*, and the subjects have an obligation to perform *x*, regardless of the content of *x*. The holder of the obligation is moved to observe the legal directive without analyzing the merit

of the directive itself. To illustrate, imagine that you and a friend have agreed to perform *x* whenever it is asked by the other, and neither of you have bothered to define by what this performance will be. In this instance each of you have acquire a content independent obligation to perform however the other person desires. Legal regimes purport to issue the same type of obligations to its subjects. Another aspect of this obligation is that the obligation preempts all other held obligations. To say that the obligations imposed by law are preemptory is to say that an authority qua authority can issue a directive to perform *x*, and if the subject holds an obligation to do not-*x*, the subject must perform according the directives of the authority—i.e. perform *x*.²

Though by no means an exhaustive account of authority and the nature of obligations, this should suffice for current purposes. Taken altogether, LCA is a claim made on the behalf of law that it is an authority: it purports to be a justified moral authority capable of creating within its subjects moral obligations that are content independent and preemptive. In §III I will retake the issue of the obligation created by authority, and show that this form of authority creates a form of blind obedience that infringes upon a person's autonomy. Taken together with PA, the normative critique shows there to be a sort of incoherence if both LCA and PA have a space within the concept of law, and as a result LCA must be excluded. However, I would now like to turn to the role that LCA plays in the concept of law.

II. LCA'S ROLE IN LAW

LCA describes the manner in which legal regimes actually act—i.e. legal regimes issue directives as if they were authoritative. Raz writes,

The law's claim to authority is manifested by the fact that legal institutions are officially designated as "authorities," by the fact that they regard themselves as having the right to impose obligations on their subjects, by their claims that their subjects owe them allegiance, and that their subjects

ought to obey the law as it requires to be obeyed. (Raz 1985, p. 300)

To illustrate, think of the Constitution of the United States. It grants powers, procedural rules, etc. to the various branches of government. But, nowhere is it mentioned that it has the right to create these powers, rules, etc. Granted, it purports to give reasons why it should be treated as an authority, but this is no different from what Raz is claiming regimes do. It merely assumes that it is an authority, and acts accordingly. I do not think that this is a limited instance—this seems to be representative of all regimes, including legal regimes. Thus, we may conclude that necessarily law claims authority.

The fact that necessarily all legal regimes purport to be authorities signifies something important about the nature of law. Raz writes, “To claim authority it [law] must be capable of having it, it must be a system of a kind which is capable in principle of possessing the requisite moral properties of authority” (Ibid, p. 199). In other words, if law claims authority, then it is structured in such a way that is capable of creating obligations within subjects that are content independent and preemptive. It is through LCA that we uncover a defining feature of law.

LCA is informative and tells us something about the nature of law, for this reason it is given a defining role in our conception of law. Further, it may well be the case that LCA is often false—that the authority being claimed by the legal regimes is outside its grasp—but because it has explanatory power, we grant it the central role that it does. It may well be the case that law does not have authority in certain instances, but this does not necessitate that law is such that it is incapable of being an authority. It simply means that in this instance, it failed to be one. The question I want to engage in now is: is the claim to authority more than sometimes false? If it can be shown that there are facts that make the claim to be necessarily false, then we may have stronger reasons to push LCA aside.

III. THE NORMATIVE CRITIQUE OF LCA

The normative critique maintains that if our concepts suffer from some sort of internal incoherence caused by tension either between descriptive facts or normative facts, then one of these facts must be excluded from the concept in order to retain some coherence. To inform the normative critique and better force the conclusion, I will work in the thesis of philosophical anarchism (PA). Understood in terms of PA, the normative critique maintains that the tension is that legal regimes are *incapable* of being authorities, and to maintain a coherent conception of law, LCA must be discharged.

The normative critique can be surmised as follows. Van Der Vossen interprets Soper as saying with respect to LCA:

Why would we continue to accept a concept of law that commits law to claiming authority if, in fact, we do not believe such a claim is defensible as a matter of political theory? (Van Der Vossen, p. 482)

In other words, if we believe that normative facts deny the possibility of law having authority, then we have no reason to continue granting LCA a central role within law. But what are these normative facts that play such a decisive role in making LCA indefensible? Simply, it is the fact that our conception of man is such that he is autonomous and holds a preemptive duty to autonomy.

Robert Paul Wolff (1970) argues that man, insofar as he is a moral being, is an autonomous being. In order for man to be truly autonomous, he must be free to take responsibility “to determine what one ought to do” (Wolff 1970, p. 12). From this responsibility to determine one’s action it follows that the agent is to engage her rational faculties and decide according to the most rational choice. Paraphrasing Kant, Wolff writes, that the autonomous person “gives laws to himself” (Ibid, p. 14). In other words, Wolff is arguing that in light of being an autonomous being, one has a duty to preserve this autonomy. An important consequence of this duty to autonomy is that,

The autonomous man, insofar as he is autonomous, is not subject to the will of another. He may do what another tells him, but not because he has been told to do it (my emphasis). (Ibid, p. 14)

The duty to autonomy prohibits acting when the reasons to do so are not the agent's own—to do so would be to violate one's duty to autonomy. More specifically, the agent must be able to give her own reasons to act in the way she did. Thus, this denies that any obligation created by an authority (i.e. content independent and preemptive) is ever justified. If the agent observes such an obligation, her autonomy is diminished as these create a condition of blind obedience to the authority. From this, PA can be defined as the thesis that an agent has a duty to autonomy and denies that agents have obligations to do x if she holds no reasons of her own to do x —i.e. one cannot do as one is told merely because one is told to do so.

PA denies the corollary of LCA—i.e. it denies that subjects have a content independent and preemptive obligation to act in accordance with the directives of the legal regime. This leads one to form the following conclusion:

The legal theory claim [LCA] is false; it is not the case that a correct account must portray law as claiming authority in the sense of an obligation to obey directives just because they are the law (my brackets). (Soper 1982, p. 5)

Law must be capable of being an authority if it is to claim authority. But, if PA is true, then it is false that law is capable of being authority—necessarily, law cannot have authority. Law cannot impose this form of obligation upon its subjects because its subjects are such that they cannot have this form of obligation. If this is so, then it makes no sense to grant LCA a central role within our concept of law. It simply does not capture the nature of law, and should be jettisoned.

Himma notes, “if there is no conceptually possible legal system in which law's claim to authority is true, then it is false

that law is capable of legitimate authority” (Himma 2001, p. 11). He acknowledged the weight of PA, if it were true, but quickly dismisses it as a fringe theory. Similarly, in presenting his argument Raz assumes the falsehood of PA. Neither sees the necessity of proving PA false—it is such a fringe theory that it doesn’t merit an argument. Surely, it cannot be denied that the claims made by PA require one to have complete and unrestricted autonomy (i.e. a preemptive duty to autonomy), and because of this perhaps it is true that PA is an implausible moral claim. PA requires much argumentation in order to prove that its central tenets are plausible, but this is outside the scope of the paper. In his essay, Van Der Vossen’s goal is to show that PA, *even if true*, is non-threatening to LCA. Because I am merely arguing against Van Der Vossen, the paper will succeed if it can be shown that the truth of PA entails that LCA is false. I will now argue contrary to this claim and show that if we grant the truth of PA, the normative critique deals a fateful blow to LCA, and thus forces the conclusion of the normative critique. In §VI I will return to the matter of PA and make some brief comments about it.

IV. VAN DER VOSSEN’S ATTACK OF THE NORMATIVE CRITIQUE

In his paper, Van Der Vossen seeks to show that the normative critique poses no significant threat to LCA. To this end, he provides possible interpretations of the normative critique in which the force of its conclusions is apparent, finding in each possible interpretation that the purported conclusion—that LCA should be jettisoned from our conception of law—is not well founded. He concludes by interpreting the normative critique as an *oddness thesis*, finding this interpretation as the most plausible and showing it to be non-threatening to LCA.³ However, instead of arguing against this interpretation, I want to argue against the *‘bad’ concept of law* and show that the buck stops here.

Before presenting his arguments against this interpretation and my response to it, I want to make the case for my strategy—

that is, I want to make explicit why my argument is directed specifically at the second interpretation and not also the remaining interpretations that Van Der Vossen provides. The goal of the paper is to show that the normative critique poses a significant threat to LCA, and if I am to succeed in this respect, it must be shown that, at the moment, there is no good argument to show that LCA can be retained within any understanding of law. At the moment, Van Der Vossen seems to have provided the clearest objections against the normative critique. Before reaching the *best possible* interpretation of the normative critique—where the best possible interpretation means the interpretation in which the normative critique poses the greatest threat to LCA—he presents other interpretations that fail to dent LCA. If I succeed in showing that the *'bad' concept of law* interpretation is the best possible interpretation, then I can block the move to interpret the normative critique as an *oddness thesis*; there would be no reason to re-interpret the normative critique if we have already found the damaging interpretation of it. Having established the reasons for my strategy, I will now present the argument and follow it with my response.

On the *'bad' concept of law* interpretation, Van Der Vossen takes the presentation of the normative critique as presented in §III. His goal here is to show that even if “law imposes content-independent obligations on no subject,” that LCA remains a useful thesis (Van Der Vossen 2011, p. 489). Having assumed this interpretation of the normative critique, he concludes:

This is where the current reading of the normative critique finds its foothold: we are to reject or reformulate our understanding of concepts when (elements of) these are in tension or contradiction with (elements of) other concepts to which we give credence. LCA holds that law makes a claim that is denied by PA. This leads to a gap between what law claims and what law is, and thus, assuming the truth of PA, LCA has to go. (Ibid, p. 489-90)

In other words, our conceptual understanding of law involves two elements: a normative element (PA) the denial that law may

impose certain obligations on its subjects; and a descriptive element (LCA) the claim that law claims authority. PA and LCA *at very least* exist in tension with one another because the former denies that law can ever be authoritative and the latter suggests that law must be capable of being authoritative. In order to have a coherent conceptual understanding of law, one of these theses must be rejected. However, PA was assumed at the outset to be true, so we cannot reject this thesis. Thus, we must jettison LCA from our conceptual understanding of law.

Van Der Vossen argues that this conclusion is not forced upon us for two reasons: first, it isn't clear what the problem is about concepts that are unjustifiable; and second, it isn't clear what the problem is about concepts that make false claims. Regarding the first reason, he writes:

Often it seems perfectly acceptable, and sometimes unavoidable, to adopt a conceptual understanding of a practice as one that fails to stand up to moral scrutiny. Indeed, it is not difficult to come up with cases where such conflicts occur in even starker form. Theft, rape and slavery are all unjust, even conceptually so, but this does not mean that we should redefine our understanding of those practices. (Ibid, p. 490)

Undoubtedly we have concepts that include claims that are unjustified, but we do not want to reconstruct these same concepts for that reason. We can find various examples of such unjustified concepts—Van Der Vossen offers the examples of theft, rape, and slavery, but there are many more that can occur to us. Slavery, for instance, includes a descriptive claim—“the ownership of one individual by another”—that fails to stand to moral scrutiny. That is, we hold a certain conception of man (a normative fact) that shows slavery to be unjust. But, it would be foolish to argue that one must revise our conceptual understanding of slavery—and any other such moral concept—in light of this contained claim that is unjust. These claims help us better understand the concept and redress any harm caused by instances of these concepts.

Law is similar in this respect: it contains a descriptive claim that fails to stand to our moral standards. More specifically, LCA is unjust because we hold a conception of man such that he is autonomous—i.e. its fails to be morally defensible, in light of PA. Here we can observe a certain parallel between law and moral concepts. Because there is no reason to reformulate moral concepts to remove unjust claims, we may conclude that the conclusions of the normative critique do not necessarily follow.

Regarding the second reason he provides an analogy to a liar. A liar, by definition, is a person who makes false claims: “All liars, at least on some understandings, necessarily make claims that are false” (Ibid, p. 490). However, again, we wouldn’t reconstruct our conceptual understanding of liars to remove this falsity from the concept, for it is an essential part of what it means for some one to be a liar. But law also makes a claim that is false—LCA is necessarily false in light of PA. Therefore, it isn’t clear that one must reformulate concepts to remove claims that are false. Here again, the conclusions of the normative critique are not forceful.

These two arguments are sufficient to show that this interpretation of the normative critique must not be the intended interpretation. Van Der Vossen concludes: “Whatever kind of (in)consistency at stake in this case is neither here nor there for the question whether law makes certain claims” (Ibid, p. 490). In other words, it isn’t clear what the threat of the normative critique is—it does not force upon one to reconstruct our conceptual understanding of law such that LCA is not given a central role. However, I am not so convinced that Van Der Vossen has succeeded in showing that this interpretation of the normative critique is ‘harmless’ to LCA. I will now argue that the arguments given against this reading of the normative critique are faulty.

Van Der Vossen argues that liars necessarily “make claims that are false. But surely we are not then to conclude that we should revise our concept of a liar” (Ibid, p. 490). Similarly, that law claims authority is false as well, in light of PA; and thus, our concept of law shouldn’t be revised. The problem with this argument is that the truth-values of the claims are different. In the case

of the liar, our conceptual understanding of it is composed of many different claims, the most distinctive that necessarily liars intentionally make false claims. But this claim is informative about what a liar is: ‘liars make false claims’ tells us something about the way liars are, namely, that one cannot trust the truth-value that the liar purports that sentences have because occasionally the assigned truth-value will be incorrect. In light of this, the claim that ‘liars make false claims’ is true.

In the case of law, our conceptual understanding is composed most notably of LCA. In light of the assumed truth of PA, it is apparent that it is not an informative claim in the same way as in the case of the liar: ‘Law claims authority’ is false because in order to be capable of claiming authority, law must be the kind of thing capable of having it, but PA denies this. Therefore, it is no longer clear that the claims made about the liar and the claims made about law are similar. Having distinguished between the two, the analogy no longer stands and Van Der Vossen’s argument here falls apart—the normative critique retains its force.

But, Van Der Vossen also provides an analogy to moral concepts, arguing, “it seems perfectly acceptable, and sometimes unavoidable, to adopt a conceptual understanding of a practice as one that fails to stand up to moral scrutiny” (Ibid, p. 490). Both law and certain moral concepts fail to stand to moral scrutiny, but we would not restructure our understanding of them because of this, as they serve an important role in understanding these concepts. Here too I find that there is an important sense in which LCA differs from claims in moral concepts. The problem is not that the descriptive claims contained within such concepts are unjust; rather it is that we have no reason to support the role of LCA within the concept of law. It is in this sense that LCA is an unjustified claim; Van Der Vossen equivocates the word ‘unjustifiable’ and misses the point of the argument.

The use of a descriptive claim within a concept must serve to inform it by capturing its distinctive features and clarify the concept itself. If it does so, then we can say that we are justified in including it within the concept—we have reasons to include

such claims. For instance, “the ownership of one individual by another” describes accurately what is meant by the term slavery. It informs others as to the nature of the concept. The same goes with all moral concepts picked out by Van Der Vossen and still others. In light of this clarification, we can see that the descriptive claims of moral concepts, such as slavery, should not be jettisoned or restructured, for we have good reasons to maintain their central role. However, in the case of law, it isn’t apparent how LCA captures the nature of law.

LCA does not inform us or clarify the concept of law, in light of the assumed truth of PA. The claim is useful insofar as it is an aid in analyzing legal regimes and bringing forth their essential nature—that they are capable of authority. But, PA has shown that this is not true of legal regimes: they are incapable of being authorities because their subjects have a duty to autonomy that is infringed if they observe the directives of the authority. So, LCA doesn’t tell us anything about the nature of law. (Undoubtedly, it describes how legal regimes act, but it doesn’t go up a level and uncover the nature of law.) LCA is simply not defensible claim if its role is to capture the nature of legal regimes. It is in this sense, that the normative critique maintains that LCA is an unjustified claim: there is simply no good reason to give this claim the centrality it has been given in the concept of law.

If the term ‘justification’ is understood in terms of reasons, then it can be seen that including certain descriptive claims in moral concepts are done so with good reason—that is, it is justified. But, turning to LCA, we find that there is no reason to include it within the concept of law—that is, it is unjustified. The fact that in each case the descriptive claims describe an act that is unjust is beside the point. Having clarified the sense in which the term ‘justification’ is used it becomes apparent that Van Der Vossen’s analogy quickly falls apart once more.

The arguments that I have provided show that Van Der Vossen fails to give truly analogous cases to LCA that show the force of the normative critique to be neither here nor there. He has not successfully shown that interpreted as a *‘bad’ concept of*

law that the normative critique lacks force. Thus, the conclusions raised by it remain threatening to those who wish to give LCA a *central role* in describing the nature of law. This consequently blocks any move to interpret the normative critique as an *oddness thesis*—in which Van Der Vossen finds that it gives support to a concept of law that includes LCA! I want to now make some brief remarks regarding the normative critique itself and alternatives present to the legal positivist if the critique is successful.

V. CONCLUDING REMARKS

There are two important points to be made regarding PA and an explanation of the duty we feel that we owe to law. Regarding PA, I find that we have no good reason to believe that it is a tenable moral theory—it simply requires too much. However, I do not think this weakens the normative critique of LCA. Defining law in terms of LCA shows that one acquires an obligation to obey law because it is law, and thus requires blind obedience to an extent. Though there may be no such duty as a duty to autonomy, it is likely that many individuals would find an obligation to blindly obey the directives of legal regimes to be bad, because autonomy seems to hold more value than an obligation to obedience. The role of PA in the normative critique of LCA was to show more forcefully this conclusion.

Turning to the second point: if a concept of law that includes LCA is a bad one because one concludes that law is incapable of being an authority, there seems to be a worry about what legal officials are doing when performing their legal duties. In other words, if law isn't authoritative, does it mean that legislatures and judges are simply creating laws for fun? This worry seems to be without foundation. For one thing, legal regimes are nevertheless *de facto* authorities: legal regimes issue directives and expect that their subjects act accordingly, and are fully capable of coercing subjects to do so if they act defiantly (by imposing sanctions, imprisonment, etc.). In such an instance the legal regime would lack any justified authority and maybe considered a purely coer-

cive system—“gunman situation writ large”, to use H.L.A Hart’s phrase (Hart 1958, p. 63). Though undesirable, this may ultimately be the reality of legal regimes if we fail to show that they do not have any moral standing (i.e. *de jure* authority).

Further, there is recent research that suggests that perhaps our reasons to act according to law are not founded on any claims to authority, but instead on the more social aspects of law like its ability to better organize our social lives. Jules L. Coleman (2009) claims, “There is, something essentially coordinative and coordinating at the center of law” (Coleman 2009, p. 386). If law serves such an end, the worry would be settled here too, because observing the law would benefit society as a whole. Such an account of law would allow the legal positivist to capture the normativity of law without incurring the baggage created by LCA, and thus distinguish legal regimes from merely coercive systems.

Would an account of law founded upon this violate an individual’s autonomy in the same manner as one that contains LCA? There is good reason to believe that it doesn’t: a reasonable person could observe the directive issued by a legal regime and determine that there is good reason to observe those directive because it creates a situation in which she can pursue her private ends—provided they are in accordance with law. There would be no content independent and preemptive obligations under such a conception of law. And I find that this account of law better captures our intuitions about law—that we have certain to obligation to obey law because it intends—though not in all cases—to better our social lives.

Notes

1. The term ‘claim’ is being used in the linguistic sense; it is equivalent to the claiming being done when one says, “So you claim to have seen a unicorn?”
2. David Enoch (2011) spends a considerable amount of time regarding this form of obligation by studying the reasons that it gives subjects to act. He argues that this form of obligation gives one “robust reasons” to act. For a fuller account see his paper “Reason-Giving and the Law”.
3. A theory of law is *prima facie* implausible if it holds that law essentially claims something is the case that is unjustifiable” (Van Der Vossen, p.

492.). Under this interpretation of the normative critique, if our concept of law contains elements that are unjustifiable or false (such as LCA), then it is *undesirable*. This interpretation does not force one to abandon such a conception of law altogether. Instead, a concept of law may be revised only if there exists a better conception of law—one that does not suffer from a form of internal incoherence. In light of the apparent absence of any theory of law that is “better” than one centered around LCA, he concludes that we have no reason to revise it, at the moment.

Bibliography

- Coleman, Jules L. (2001) *The Postscript: Essays on the Postscript to the Concept of Law* (Oxford: Oxford University Press)
- Coleman, Jules L. (2009) “Beyond Inclusive Legal Positivism” *Ratio Juris* 22(3), pp. 359-94
- Enoch, David (2011) “Reason-Giving and the Law” *Oxford Studies in Philosophy of Law* 1, pp. 1-38.
- Estlund, David (Ed). (2012) *The Oxford Handbook of Political Philosophy* (Oxford: Oxford University Press)
- Gardner, John. (2009) “How Law Claims, What Law Claims” *Legal Research Paper Series*
- Hart, H.L.A. (1958) “Positivism and the Separation of Law and Morals” *Harvard Law Review*, pp. 593-629
- Himma, Kenneth Einar. (2001) “Laws Claim of Legitimate Authority,” in: Jules L. Coleman (Ed), *The Postscript: Essays On the Postscript to the Concept of Law* (Oxford: Oxford University Press)
- Raz, Joseph. (1985) “Authority, Law and Morality” *The Monist* 68(3), pp. 295-324
- Raz, Joseph. (1979) *The Authority of Law* (Oxford: Oxford University Press)
- Soper, Philip. (2002) *The Ethics of Deference* (Cambridge: Cambridge University Press)
- Van Der Vossen, Bass. (2011) “Assessing Law’s Claim to Authority” *Oxford Journal of Legal Studies* 31(3), pp. 481-501
- Wolff, Robert P. (1970) *In Defense of Anarchism* (Berkeley: University of California Press)

REASONS AND NORMATIVITY

Mark Gaynor

INTRODUCTION

Externally sourced reasons have come to play a significant role in several contemporary versions of moral realism—in particular, these kinds of reasons (which I will variably call either external or object-given) are employed in the moral epistemology of Chang (2012), Parfit (2011), and, to a lesser extent, Singer (2012). However, the exact nature of reasons remains controversial, and some philosophers (Williams, among others) have argued that reasons cannot fulfill the role that these moral realists want them to, citing either metaphysical or epistemological problems with object-given reasons. In this paper, I will argue that the problems associated with object-given reasons can plausibly be dissolved, but that the particularity in which moral realists want reasons to be cast may be unattainable.

In framing my argument, I will set up the debate concerning the nature of reasons as being between source internalists and source externalists. In developing the position that I will defend, I will consider Bernard Williams' arguments against the plausibility of external reasons and argue that these challenges can be plausibly overcome without recourse to any metaphysically or epistemologically questionable entities. Specifically, the conception of reasons that I will advocate will take reasons to be encoded by norms, where norms are conceived of as sets of rules that are expressed through psychological dispositions to promote certain values and outcomes. Importantly, this conception of reasons will lack the particularity of the aforementioned moral realist conception of reasons—rather than there being only those reasons generated by moral truths, there will be reasons possessing every possible content. I will argue that this consequence also follows

from Parfit's non-metaphysical cognitivism. In my concluding remarks, I will suggest that moral realists do not need to take recourse to reasons as fundamental, but instead need to argue for principles that individuate and designate the reasons they intend.

SUBJECT- AND OBJECT-GIVEN REASONS

Reasons are given by facts, or purported facts, and are taken to *count for* or *against* a particular action—that is, they hold justificatory weight, or force, in considerations of what course of action a person should pursue. Reasons play a crucial role in explaining conscious behavior in everyday life—that is, much of the time, we act on the basis of reasons. Reasons are also taken to be inextricably related to rationality—our beliefs and actions can be evaluated in terms of how they count as responses to the reasons we have. If, for instance, I loathe bubblegum ice cream, and I am offered some, I have a fairly strong reason to decline the offer. Barring the existence of any countervailing reasons, if I choose to eat it anyway, then my doing so would be at least a bit irrational. If, on the other hand, I am offered a large amount of money to begrudgingly eat the bubblegum ice cream, and I am in financial dire straits, I might find myself with a stronger (or perhaps merely countervailing, depending on the force of my hate for the ice cream compared to the extent of my financial need) reason to just eat the stuff.

Now, some philosophers, following David Hume in the belief that “reason is the slave of the passions”, have argued that only certain kinds of things can give us reasons—namely, our desires.¹ Let's call these internal, or subject-given, reasons. Proponents of the subject-given reasons view take rationality to be inherently about reaction to and advancement of one's own, internally stipulated, aims.² Perhaps the most powerful motivation for this view is provided by the fact that it can be easily incorporated into a larger naturalistic picture of the mind and its relation to the world.

In contrast, others, perhaps the most prominent of which being Derek Parfit, believe that in addition to there being reasons

that stem from our desires, the nature of certain phenomena can give us reasons to desire or act to bring about those phenomena—that some possible action would result in an experience of pain, for instance, generates a (not inherently decisive) reason to avoid performing that particular action. Let's call these external, or object-given, reasons. Philosophers who have argued for object-given reasons have thought that the rationality of our actions depends at least in part upon what consequences those actions are expected to bear upon the wellbeing of ourselves and others. Put another way, proponents of this view believe that independent from considerations of a person's aims, there are facts concerning whether some goals are worth desiring. To better assess the plausibility of these frameworks, it will be necessary to make a few distinctions.

REASONS AND EXPLANATION

First, taking reasons as purely explanatory entities, when proponents of subject-given reasons claim that all human actions are motivated by desires, thus forcing the conclusion that the only thing that can count as a reason is a desire, this claim is advanced either as an unfalsifiable platitude or a highly controversial empirical thesis—one that, on the face of it, seems obviously false. First, the platitude: one can claim that whatever one does just is what one most wanted to do, because the only motivating powers inherent in the mind are desires. On this model, if a person considers her options, finds that she would really like to go to the movies, and would far prefer that to doing laundry, but then decides to do her laundry anyway, it turns out, according to the internal reasons proponent, that really, that person most wanted to do her laundry. One point that's worth noticing here is that the concept of desire is being employed in two functionally distinct ways—desire is being used in the first place to denote an psychological state of wanting something, whereas in the second place it is being used to denote having chosen to pursue some particular course of action. It is not at all obvious that whatever one chooses

to do is wholly determined by one's antecedent states of wanting. If we move away from a simplistic picture of the mind that admits only of desires as having causal efficacy toward one that grants the possibility of acting on the basis of forces other than desire, then it becomes an open empirical question whether a person's reasons stem from her desires.

What is required to vindicate the external reasons model as having explanatory power is (i) for a person's desires to sometimes fail to fully determine the course of action she does pursue, (ii) for it to be possible for a person to assign deliberative weight to certain expectations about an action or its consequences, and (iii) that this assigned weight can go part of the way toward determining the particular course of action a person does pursue. Now, (i) has already been argued for, (ii) is introspectively obvious, and (iii) can be argued for as follows: it is possible to act on the basis of criteria that don't make reference to our desires. A simple case of this occurs when a person flips a coin to make a decision—for this to be useful to proponents of object-given reasons, it needs to be the case that we can develop criteria internal to a decision-making process that are reactive to what can be expected to result from the pursuit of a particular course of action. If a person decides that she wants to pursue some particular course of action, then reflects upon what consequences that course of action is likely to produce, and realizes that it may cause significant harm to others, it needs to be possible that she can revise her chosen course of action in light of recognizing that fact. If it is possible for her to revise her course of action in light of recognizing the harm her original course of action might cause, then object-given reasons would seem to have at least the same explanatory power as subject-given reasons within this dimension of explanation. Williams, however, argues that this sort of explanation cannot proceed straightforwardly, as it fails to sufficiently attend to the relation that holds between the source of a reason, a person's motivations, and her subsequent actions.

On this view, appeals to external reasons are problematic because these reasons are not falsifiable in the same way that

internal reasons are (Williams 1981, p. 109). A claim about internal reasons is falsified “by the absence of some appropriate element from [a person’s motivational set]” (Ibid, p.102). Because external reasons are not given by elements in a person’s motivational set, Williams claims that it is not possible to explain how these reasons can give rise to action. I believe that this problem, however, is not insoluble. Williams assumes that it is, and he thinks that the problem grounds the claim one’s response to external reasons, if there are any such things, does not affect the status of one’s rationality, whereas response to one’s motivations does fundamentally affect the status of one’s rationality. To articulate how that this problem can be overcome, I will follow the analysis of Jollimore (2005).

Williams’ claim is that adherence and response to subject-given reasons, in the form of instrumental rationality, just is what rationality amounts to. This is either because this is just what we mean when we claim that a person is acting rationally, or because there is some deep fact of the matter about the nature of rationality.

This view might be expressed along the following lines:

- (1) It is necessarily irrational to acknowledge that an action will contribute to the achievement of one’s goals, and yet fail to recognize a reason to perform it. (Jollimore 2005, p. 291)

and

- (2) It is not necessarily irrational to acknowledge that an action is morally required, and yet fail to recognize a reason to perform it. (Ibid, p. 291)

Proponents of the object-given reasons model of rationality will reject (2), as it is just another way of saying that it is possible to recognize that you have an object-given reason and fail to see that as giving you a reason to act—they will respond by saying that anyone claiming (2) is begging the question and that a person *would* be irrational to acknowledge some act as morally required but not recognize that as a providing reason to perform it. Let’s suppress this for the moment, as it is not yet clear whether this

amounts to anything more than a semantic dispute about the application conditions for the term ‘rational’. The more pertinent question here is whether there are any facts that we can take recourse to as a means of solving the quibble between the two parties.

I think (1) and (2) track some of our pre-theoretical intuitions fairly well, and these assumptions can serve as a starting place for showing that instrumental rationality cannot plausibly lay a monopolistic claim upon our deliberations. First, as a proviso about what this can show: I think we ought to reject the claim that our ordinary language use and our intuitions can relevantly arbitrate the debate concerning how we should reason. That we do apply terms and generally think in a certain way cannot count as an argument for adopting a practice, or for preserving that practice when it comes under scrutiny. To emphasize this point, consider that within some sects of Islam, women are treated as subservient, in terms of both their autonomy and overall worth, compared to men. I assume that people within these societies have intuitions that confirm these practices. The presence of these intuitions, by themselves, does nothing to aid in the resolution of the dispute whether the associated practices have merit or should be preserved. If intuitions are to do any real work here, it needs to be because there is something factual underlying them, or because there is an argument that supports the same conclusion; otherwise, they play a role in a circular explanation. If there is something underlying them, or some structure that intuitions respond appropriately to, then presumably that can be explicated in terms of some non-intuition based justification. If there is any fact about what justifies an action in a normative sense, then it’s not going to be found by inspecting our ordinary language practices. So, we should reject the ordinary language approach, assume for the moment that there is some fact of the matter, and just look for an argument.

The kind of argument needed to settle this dispute is going to be normative—it will be concerning the value of a practice, and it is going to have to take shape as an argument for the acceptance of (1). The argument runs as follows: (1) is presumably based upon

and derived from the following, more fundamental instrumental principle:

(IP): The fact that an agent endorses *e* as one of her ends guarantees that she has (at least some) reason to pursue *e*.
(Ibid, p. 293)

In order to see the plausibility of this claim, consider a person who is absolutely unmoved to act in accord with it, let's call her Ann: Ann has some goals, and she recognizes that she can accomplish them by pursuing some course of action *a*, and that *a* would be instrumental to satisfying all of her present goals, but she refuses to accept that she has reason to pursue that course of action (Ibid, p. 294). We can suppose, to illustrate the important relation of dependence here, that the goals she endorses and thinks are worth pursuing are at least partially constituted by aims that a proponent of an object-given reasons account would endorse. This allows for a powerful claim: if there are any goals worth pursuing, then it is necessary for a person to act as (IP) recommends. This means that even if externalists are right about there being object-given reasons and those reasons bearing normative weight, internalists seem to have their hands on the more fundamental kind of rationality in light of the fact that acting in accordance with the dictates of (IP) is necessary for any sort of valued action.

This internalist view of rationality, however, is not without its limits, and it runs into problems if one attempts to systematically endorse it. In particular, it almost immediately loses traction if one attempts to employ it in the realm of theoretical reason. To see how this is so, we will need to translate the former argument into one pertaining to theoretical reason. (IP) will need to be translated to an inferential principle:

(IT): An agent has (at least some) reason to accept those claims that follow logically from beliefs she accepts. (Ibid, p. 296)

And the theoretical corollary claims will be:

(1t) It is necessarily irrational to acknowledge that a claim is logically implied by one's current beliefs, and yet fail to recognize a reason to believe it. (Ibid, p. 297)

and

(2t) It is not necessarily irrational to acknowledge that a claim is supported by the best currently available scientific evidence, and yet fail to recognize a reason to believe it. (Ibid, p. 297)

(1t) and (2t) together entail that while it would necessarily be irrational to reject some beliefs that are logically entailed by one's currently held beliefs, it would not necessarily be irrational to reject well vetted claims from the empirical sciences—a person should reject the best contemporary scientific accounts of physics, or biology, or whatever, if they conflict with her beliefs in astrology and the efficacy of witchcraft. This claim about the primacy of *logically entailed beliefs* over *beliefs acquired from scientific investigation* is functionally equivalent to the primacy given to instrumental rationality over object-given reason rationality by (1) and (2).

One might think it possible to save (1) and (2) from the insidious consequences of (1t) and (2t) by claiming that it is possible to draw a sharp distinction between practical and theoretical reason, and simply grant that theoretical reason works differently than practical reason. I think this attempt is doomed: there simply is no such sharp distinction between the formation of our aims and our changing our beliefs in response to new information.

To see this in a fairly benign case, suppose that you see a snake and want to pick it up. After deciding that you want to do this, somebody points out to you that the snake is poisonous and in a bad mood. If you accept that this gives you a reason to not pick the snake up (presumably because having to deal with the resulting medical issues would be burdensome and in conflict with your long term goals), then you are accepting that facts about what would likely occur if you were to pick up the snake can give

you a reason. In this case, the object-given reason is derived from facts about your goals and how the world is, and so it might still seem that the internalist is at least mostly right—reasons still seem to have their ultimate source in a person’s aims in this case. But perhaps object-given reasons can erode subject-given reasons.

Let’s suppose that there is a person Ralph, that he is a deeply committed racist, and that his beliefs form a coherent whole. Let’s further suppose that Ralph has come in contact with a biologist who explains to Ralph that there are absolutely no biological grounds for believing that any such thing as race even exists; that skin color is the product of adaptation to a particular climate over time, and that race as we know it is a social construct. If (1t) and (2t) are true, then Ralph would be irrational to even call into question his racist beliefs on these grounds. As far as the acquisition of knowledge is concerned, these claims about rationality seem obviously false. What dooms the attempt to preserve the distinction between theoretical and practical reason is the fact that, according to proponents of object-given reasons, rationality sometimes just consists in a person’s moving from the recognition of some information (and here the distinction between theoretical and practical is superficial)—that racism is unfounded—to the abandonment of previously endorsed aims stemming from subject-given reasons.

What all of this suggests is that when scrutinized by our reflective lights, instrumental rationality is not secure as a complete account of how we should react to reasons. If we grant the merits of epistemological projects and of acting on the basis of beliefs that are grounded by the way the world is, then some of our subject-given reasons can be complemented, modified, and invalidated by object-given reasons. We now have the means to respond to Williams’ argument against external reasons.

Object-given reasons, in order to give rise to action, must connect with a person’s motivational set. These reasons connect with the motivations of a person in virtue of that person recognizing their significance. In recognizing the significance of an object-given reason, a person incorporates that reason into her aims, allowing for the reason to be run through her instrumental

rationality and to give rise to action. In this way, she is capable of responding to information in ways that can be evaluated apart from her aims. I want to turn now to considerations of the source of object-given reasons.

OBJECT-GIVEN REASONS AND THEIR SOURCE

Of those who believe that object-given reasons can substantially inform deliberation, most are non-naturalists, and that they take this position is seemingly necessary: these philosophers claim that some reasons are abstract normative entities—i.e., non-physical, non-mental things that can direct action. Central to this view is the claim that there are certain irreducibly normative truths—that a particular, limited set of reasons partially constitute the fundamental furniture of the world, and that these reasons come, in some sense, pre-packaged—their content and force determined by the world, or by laws applying over all possible worlds. As Parfit writes, “In any possible world, pain would be in itself bad, and *prima facie* to be relieved rather than perpetuated” (Parfit 2011, p. 489).

These claims are built on the backbone of what Parfit refers to as “non-metaphysical cognitivism”. This view, as Parfit explains it, allows for there one to be committed to the truth of certain statements without being committed to any ontologically prior truth-makers—in other words, we can evaluate the truth of some statements based purely on their semantics. Parfit believes that claims concerning reasons are of this kind, and claims:

There are some claims that are irreducibly normative in the reason-involving sense, and are in the strongest sense true. But these truths have no ontological implications. For such claims to be true, these reason-involving properties not need exist either as natural properties in the spatio-temporal world, or in some non-spatio-temporal part of reality. (Ibid, p. 486)

If it is the case that we are to evaluate these claims merely

on the basis of their meanings, then I see no means by which we can block the entrance of reasons possessing every possible kind of content—there will be plenty of reasons the content of which any competent speaker of English would recognize as crazy. Parfit claims that we can determine what reasons we do have by reacting appropriately to the situations we find ourselves in and by valid reasoning, so long as our reasoning is not affected by any distorting influences—however, it is by no means apparent that we can attain anything like precise content by means of this method. I take this to significantly reduce the apparent viability of this approach.

Parfit makes similar claims about the nature of mathematics, and I believe we might draw a parallel here between the development of mathematical systems and systems of reasons in order to disavow ourselves of the problems that are generated by the entrance of reasons bearing every possible kind of content. Mathematicians, in developing systems of mathematics, stipulate certain axioms, and then draw entailments from those axioms to conclusions about the systems. As a parallel, I suggest that moral realists ought to stipulate certain principles, and then subsequently derive normative claims from those principles. In order for projects in mathematics to be undertaken successfully, mathematicians need to take the semantics of mathematics seriously—I suggest the same holds true for moral theory.

In order for a mathematical theory to be useful, it needs to be appropriately non-arbitrary and have conditions under which we can employ that theory in our projects. Mathematicians debate the usefulness of certain axioms and the viability of systems relative to one another. Relative to the status of normative claims, there are theories that are more rigorously structured and thought out, whose foundational principles are more thoroughly defensible, and the models of which are capable of preserving and maintaining many things which are in humanity's interests and which promote its wellbeing when considered in the context of long term consequences. In the same way we don't need there to be one true system of mathematics, we don't need it to be the case that there

exists a singular necessary set of reasons. What makes a physical theory good is its being able to simplify an explanation of physical events down to something that can be used to predict future events. What we can reasonably take a moral theory's worth to be determined by its being capable of guiding our practice in a way that can be endorsed from outside our particular experiences. If we distance ourselves from our own actions and the events within which those occur, we have a better chance of determining whether that course of action can be integrated into its context as a meaningful and beneficial in an impartial sense to those parties affected.

In terms of metaphysical commitments, all we need for this to be the case is for there to be identifiable norms. This requires nothing philosophically controversial. In terms of how this relates to the theory developed by Parfit, I suggest that this does not force any dramatic changes. It allows for much of the normative content to go untouched, while allowing for us to disavow reasons as being in any sense fundamental. In taking the semantics seriously, but reframing the matter in terms of principles, we have a better interface for inquiry and revision of moral theory.

Notes

1. Williams notes that Hume's own views are more complicated than the straightforward internalist interpretation of reasons (102). As my aim is to show that there is a robust conception of reasons available that doesn't require any undue metaphysical commitments, and because I will argue that there is no metaphysical fact of the matter concerning the nature of rationality, I will not spend time worrying about more nuanced versions of the two models.
2. It should be noted that some philosophers think that a person's reasons are generated by a combination of one's desires in addition to some structural constraints given by considerations of logic or efficacy. As I am conceiving of it here, this position is actually a hybrid of the internalist and the externalist position—the structural constraints cannot plausibly be written off as anything other than external considerations imposed upon a deliberation process.

Bibliography

- Jollimore, Troy. (2005) "Why is Instrumental Rationality Rational?," *Canadian Journal of Philosophy* 35.2, pp. 289-307
- Parfit, Derek. (2011) *On What Matters* (New York: Oxford University Press)
- Williams, Bernard. (1981) *Moral Luck: Internal and External Reasons* (Cambridge: Cambridge University Press)

PHENOMENOLOGY AND TIME: HUSSERL, DERRIDA, ZAHAVI

Jared Gee

PHENOMENOLOGY AS PRESUPPOSITIONLESS

The phenomenological turn in early 20th century philosophy made radical anti-metaphysical claims. Phenomenology is a practical experiential engagement with the world, not a set of a priori claims about nature and reason. Edmund Husserl claims that phenomenology should be presuppositionless, rejecting previous assumptions about the nature of the world and existence. Husserl's focus on experiential engagement led to a complex elaboration of the experience of consciousness that changed the path of philosophy.

Since phenomenology claims to be presuppositionless, it becomes necessary to scrutinize its own foundations in order to ground phenomenology on something other than metaphysical claims. This scrutiny seeks to hold phenomenology to the standards it claims for itself. This paper will address the implications of Husserl's attempt to develop a presuppositionless philosophy of temporal experience. In doing so I will turn to Jacques Derrida, Dan Zahavi, and Martin Hagglund to show that Derrida has gone beyond this metaphysics while Zahavi still stands shakily upon it. This discussion aims to hold phenomenology accountable to its own goals and will not only clarify what is at stake, but will allow us to re-evaluate the status and use of phenomenology today.

Husserl's work on intuitive interaction with the world necessitates a new and complex theory of time. This task, however, is difficult, and later we will see that Husserl makes different claims about the nature of internal time-consciousness that show that he ultimately lost his battle against metaphysics. Later interpreters of Husserl's theory of internal time-consciousness also grapple with this problem. While Jacques Derrida formulates a radical critique

of the metaphysics in which Husserl gets caught, Dan Zahavi seems to reinterpret inner time-consciousness but, like Husserl, cannot elaborate a specific non-metaphysical theory.

Derrida concerns himself with the possibility of a non-metaphysical theory of time-consciousness by addressing the self-constitution of the absolute flow of consciousness, the flow upon which internal time-consciousness is based. As we will see, Husserl's theory is that the absolute flow constitutes itself, and although Husserl fluctuated on the role of temporality regarding the absolute flow, he ultimately delivered a metaphysical premise upon which to ground his philosophy. Dan Zahavi's reading of Husserl, although astute, does not ultimately take a stance on how to address the constitution of the flow of consciousness without metaphysical premises. Instead he focuses on the way that the absolute flow and temporal perceptual experience work together so that self-awareness can be pre-reflective. The self-constitution of the absolute flow raises large problems for phenomenology, relegating it to metaphysics. Meaning, knowledge and identity rest upon this self-constitution. As a result, phenomenology cannot provide a sufficient account of the most foundational aspect of the project.

While Zahavi's reading of Husserl follows a pre-reflective stance toward consciousness, Derrida's view posits reflection and representation as the ground for the possibility of meaning. Zahavi's view of pre-reflective consciousness allows for the appearance of the self to itself within the flow of time at the same exact moment. Such a view presupposes the unity of self and consciousness without any temporal delay, such that no reflection upon oneself in the now is necessary. Experiential consciousness presupposes self-awareness. If we remove reflection from self-awareness then we remove any temporal delay. Zahavi's view negates the representational structure of experience that Derrida claims, on which any self-awareness requires a reflective temporal delay, disrupting any self-presence. For Zahavi the subject is given to itself in unity with no mediation.

Zahavi's view fails to account for the founding metaphysical

principle of Husserl's thought, the constitution of the absolute flow of consciousness. Following Derrida and Martin Hagglund's work on Derrida and Zahavi, I will argue that Zahavi's view of pre-reflexive self-awareness still rests on a metaphysical foundation and, as a result, his theory of pre-reflective self-consciousness cannot ground itself temporally. Further, Zahavi's theory requires that the absolute flow be either atemporal or metaphysically structured in order to maintain pre-reflective self-awareness. Once the constitution of the absolute flow is temporalized, Zahavi's theory of pre-reflective self-awareness falls apart.

INTERNAL TIME-CONSCIOUSNESS: AN OVERVIEW

Husserl's theory of internal time consciousness characterizes experience as having three elements: protention, retention, and primal impression. Primal impression is the now of consciousness, protention is the anticipation of what is to come, and retention is the holding onto of what has just passed. This view holds that consciousness is not just a series of now-moments, one after another, but a connected flow in which what has just passed and what is next are necessary in order to make sense of any experience of objects in the world. It is clear that our experience of songs, for example, requires a connection between the moments in time to establish the coherent unity that is the song.

I do not hear the melody but only the particular tone which is actually present. That expired part of the melody is objective to me is due—one is inclined to say—to memory, and it is due to expectation which looks ahead that, on encountering the tone actually sounding, I do not assume that that is all. (Husserl 1964, p. 43)

This basic experiential structure of temporal objects outlines the way in which the tone of the now passes away and is still held in perceptive experience while the anticipation of what is to come is also a part of that experience.

When the tonal now, the primal impression, passes over into retention, this retention is itself again a now, an actual existent. While it itself is actual (but not an actual sound), it is the retention of a sound that has been. (Ibid, p. 50)

The retention of the sound in experience, Husserl says, is internally real even though the external sound itself no longer exists. The sound is actual internally, yet externally non-existent.

Although Husserl describes retention as memory in this quote, he also attempts to differentiate retention from recollection. He makes a distinction between primary memory, which is retention, and secondary memory, which is the reflective act of recalling a memory from the past. This differentiation, however, is tenuous. “We do not really hear and have not really heard when in memory or phantasy...in the former case [retention] we really hear; the temporal Object itself is perceived; the melody itself is the object of perception” (Ibid, p. 58). The difference for Husserl between primary and secondary memory rests on the actuality of hearing in perception, as opposed to recalling a perception, which for Husserl is not really hearing. However, as stated above, Husserl establishes retention as present experience, even though the tone itself is no longer present objectively in the world. Yet if the tone itself no longer exists in the world, then we no longer can hear it now. Here we begin to see the difficulties in Husserl’s distinction between primary and secondary memory. Later Husserl makes a clear contradiction: “After the melody has sounded, we no longer perceive it as present although we still have it in consciousness” (Ibid, p. 58). Differentiating between perception and experience, Husserl posits that one can experience the object as not present, yet still perceive it as now in retention. If the perception is over, and retention relies on perception, then recollection and retention are no longer distinct. On the same page Husserl states both that retention is still the now and retention is no longer perceived as present. If the melody is still in consciousness as retained now, yet no longer perceived as present in consciousness, then we have a theory on which the now and the not now are simultaneous.

The difference between primary memory and secondary memory begins to unravel. If this difference doesn't hold, then our ability to perceive a flow of experience rests on our ability to remember something non-present. Recollection requires representation. A theory of signs is now necessary for the constitution of meaning in Husserl's phenomenology. This is the starting point of Derrida's critique of Husserl.

DERRIDA'S CRITIQUE OF INTERNAL TIME CONSCIOUSNESS

Derrida holds that if the non-present constitutes the possibility for temporal perception in the now, then representation or signification must be at the core of all experience. Memory must constitute the possibility of presence and therefore meaning, and a representative structure or a substitute would ground the possibility of any presence and meaning. As a result there is no possible experiential self-presence of meaning as now. Any now is really a retention re-presenting what's just past while what's to come, protention, brings the possibility of its erasure in the movement of time. For Derrida, there is no now.

Derrida also takes issue with internal time-consciousness because it rests on the self-constitution of the absolute flow of consciousness, a metaphysical premise. Derrida's critique of Husserl's time-consciousness goes right to the heart of Husserl's project. Derrida states that for Husserl, "Self-presence must be produced in the undivided unity of a temporal present so as to have nothing to reveal to itself by the agency of signs" (Derrida 1973, p. 60). This undivided unity without signs founds the metaphysics of presence that Derrida critiques, yet for Husserl it grounds the possibility of any meaning. If self-presence lacks the need for signification, and experience of objects in the now allows for the possibility of meaning, then meaning serves as its own evidence through presence. Such a position metaphysically grounds all possible meaning, and the function of language or representation is to fulfill the need to relay this self-evident meaning. This meta-

physics is the crux of Husserl's phenomenology. If it is undermined, then phenomenology must be drastically revised.

In *Speech and Phenomena*, Derrida begins, "no now can be isolated as a pure instant, a pure punctuality" (Ibid, p. 61). Husserl holds the source point, the now, to be primary since none of the aspects of time consciousness alone could make sense of our experience. If internal time-consciousness is the protention of what is about to become the now, which is then retained for connection with the next now point, then there must be a past now point that is still in consciousness. If a past now that is no longer present is required for the stream of consciousness to make sense, then a non-present must be at the heart of all perception and all meaning. "One then sees quickly that the presence of the perceived present can appear as such only inasmuch as it is continuously compounded with a non-presence and non-perception" (Ibid, p. 64). Retention and protention as non-presence allow for the possibility of the now. "These non-perceptions are neither added to, nor do they occasionally accompany, the actually perceived now; they are essentially and indispensably involved in its possibility" (Ibid, p. 64).

By introducing alterity, or non-presence, into the now and linking the necessity of presence to a metaphysical foundation, Derrida attempts to reformulate meaning outside of metaphysics. For Husserl, presence establishes self-evident meaning. This then leads to the secondary necessity for signification as the passing on of meaning. If meaning is produced through experience based on presence, then signification becomes an exterior aspect of meaning as presence. Meaning for Husserl is the self-evidence of the interiority of self, and language is the exteriorization of meaning. Derrida states, "Such a perception or intuition of self by self in presence would not only be the case where 'signification' in general could not occur, but also would assure the general possibility of a primordial perception or intuition, i.e., of nonsignification as the 'principle of principles' " (Ibid, p. 60). Meaning as presence in the now of self-unity is the production of evidence. It produces meaning and justifies it simultaneously. "Within philos-

ophy there is no possible objection concerning this privilege of the present-now; it defines the very element of philosophical thought, it is evidence itself, conscious thought itself, it governs every possible concept of truth and sense” (Ibid, p. 62).

Derrida goes on to target Husserl’s distinction between expression and indication, where the linguistic expression of the object coincides with the experience of the object. Indicative speech, however, refers to things that are not present in the now. Indication and expression are related, but indication ultimately stands for non-evidence, non-presence, and representation. An expressive sign, however, carries a sense or meaning as evidence due to the experience of the referent in consciousness. In Derrida’s project, if no pure presence in the experience of the now can be established, then interior grounding of meaning as presence becomes impossible. All meaning would be the function of indicative language, in which signs stand in for referents. We can now see why internal time-consciousness is so important for Derrida in his investigation of the metaphysical premises that constitute meaning. Derrida’s critique of internal time-consciousness replaces self-evident meaning in presence with a representative structure of meaning whereby all speech is indicative speech. Expressive speech cannot exist without the self-evident meaning established in the now of experience. Since all experience is formed by retention and protention, meaning can no longer be grounded on self-evident presence. Signification is necessary for the possibility of any meaning.

Derrida’s key point in his reading of Husserl is that Husserl’s primary arguments come unraveled when these distinctions are closely analyzed. He states that, for Husserl, the “source of certitude in general is the primordial character of the living now; it is necessary therefore to keep retention in the sphere of primordial certitude and to shift the frontier between the primordial and the non-primordial” (Ibid, p. 67). Since Husserl’s theory depends on the present for experiential self-evident meaning, and he knows that retention is a necessary part of our experience of objects, then he must make retention part of the present and keep secondary

memory or re-presentation separate from the primary. This distinction allows his theory to continue to function.

For Derrida the appearance of self to self, unmediated as self-awareness, is no longer possible. Although Husserl, and later Zahavi, attempts to make a distinction between re-presentation and retention, the only way this distinction can be maintained is by calling retention perception. Husserl's project is threatened at its core and Zahavi's pre-reflective theory becomes a recapitulation of Husserl's metaphysics.

ZAHAVI ON HUSSERL AND DERRIDA

Husserl calls the possibility of time consciousness the absolute flow. Husserl, however, is not clear regarding the temporality of the absolute flow. He only states that internal time-consciousness is based on it. Husserl claims not to have words for how the absolute flow itself is constituted.

[The absolute flow is] the primal source-point, that from which springs the 'now', and so on. In the lived experience of actuality, we have the primal source-point and a continuity of moments of reverberation. For all this, names are lacking. (Husserl 1964, p. 100)

The absolute flow, however, must itself be explained. If it is self-constituting and not subject to temporality, then it rests on a metaphysical premise. While Derrida's theory shows the inability of internal time-consciousness to explain itself outside of metaphysics, Zahavi's theory can't move away from metaphysics in order to maintain a pre-reflective self-awareness.

Zahavi takes an extended view of consciousness, according to which primal impression, retention, and protention together constitute the now. Retention and protention, together with primal impression, are part of the living present. For Zahavi retention is not just past and protention is not just to come but instead both are part of the primal impression.

The retention and protention are not past or future in regard

to the primal impression. They are ‘together’ with it, and the self-manifestation of stretched consciousness consequently possesses the full structure primal impression-retention-protention. (Zahavi 1999, p. 85)

Zahavi claims that this allows him to maintain the now and presence, and therefore a view of pre-reflective self-awareness as the coincidence of self-manifestation of the absolute flow and pre-reflective self-awareness. In order for Zahavi’s theory to be grounded, the constitution of the flow of consciousness must be explained. Even if Zahavi is correct in positing retention together with primal impression in the now, he must still explain how the constitution of the absolute flow constitutes itself and brings awareness to the act of experience.

Zahavi critiques Derrida by calling into question his interpretation of retention as delay and absence, as well as its relation to primal impression. Zahavi elaborates his extended view, “Thus it is not the retention, but that which is given in it, namely, the retained, which is past and absent” (Ibid, p. 85). Zahavi asks how Derrida can account for the fact that experience tells us there is a now, rather than a just past moment we perceive as now. Zahavi argues that this places unconsciousness at the heart of consciousness and that all the problems of the reflection-theory surface again. “The retention retains that which has just appeared, and if nothing appears, there is nothing to retain. Thus, retention presupposes self-awareness” (Ibid, p. 86). Zahavi characterizes Husserl’s inner time consciousness as the point where the pre-reflective self-awareness of the act and the self-manifestation of the absolute flow are one and the same.

Zahavi’s extended view allows him to posit a pre-reflective self-awareness, but it does not account for what makes up the absolute flow itself. By collapsing the absolute flow into the givenness of the act, he attempts to bypass an account of how the absolute flow comes to be. “The pre-reflective self-awareness of the experience is nothing but the perpetual self-manifestation of the flow. They are one and the same” (Ibid, p. 80). Here Zahavi’s

metaphysics of the absolute flow begin to show, and his stand on the temporality of the absolute flow wavers. He posits the absolute flow as constituting time, but not of that time. “The structure of this field of experiencing is not temporally extended” (Ibid, p. 81). He goes on to say that only in secondary reflection does the flow become temporalized. Therefore, “the structure of constituting time-consciousness cannot be adequately grasped using temporal concepts derived from that which it constitutes” (Ibid, p. 82). This is an attempt to escape the issue. If the flow cannot be grasped using temporal concepts derived from that which it constitutes (primal impression, retention, and protention), then Zahavi must provide another temporal possibility for its constitution, rather than merely claiming that it is the constitution of the experience that constitutes the absolute flow.

For Zahavi the absolute flow presences and absences. It is streaming and allows for temporal self-awareness and reflection, but he claims that because it makes up the temporal dimension of act-consciousness, it cannot itself be constituted by this same temporality. “Time-constituting consciousness is not in time, but it is not merely a consciousness of time; it is itself a form of temporality” (Ibid, p. 82). The absolute flow grounds the possibility of temporality, yet it does not conform to the same temporality. It is a form of temporality. It is not subjected to the same rules of temporality as experience, even though its constitution is the self-awareness of experience. The absolute flow still does not have a clear temporality or reason for its constitution. It is self-constituting as a form and not relegated to the temporal structure of experience.

In Zahavi and Gallagher (2014) the authors state, “Temporal experience...is not an object occurring in time, but neither is it merely a consciousness of time; rather it is itself a form of temporality, and ultimately the question to ask is whether it makes sense to ascribe temporal predicates to time itself” (p. 4). Zahavi has argued that the flow cannot be of the same temporality as objects of the flow. He has also argued that it may not make sense to ascribe temporality to time itself. Further, the constitutive elements of time consciousness, he says, do not match the protention-retention-

tion-primal impression structure of experience. “Rather it is their very conjunction [the object with internal time constitution] which makes possible the sense of present, past, and future” (Ibid, p. 5). Given Zahavi’s claims, we still do not have a solid argument for the constitution of the flow. Zahavi is not firm in his views on the flow and does not offer to justify its status. Instead we are left with an absolute flow that is not subjected to the temporality it makes possible. Zahavi, like Husserl, gives a special status to the absolute flow. For Zahavi the absolute flow is self-constituting and not subjected to the same temporality as experience.

IN DEFENSE OF DERRIDA: HAGGLUND ON ZAHAVI

In his book, *Radical Atheism: Derrida and the Time of Life*, Martin Hagglund argues that Zahavi’s stance on internal time consciousness fails to address the most important implications of Derrida’s work. The problem with the absolute flow is that Husserl wants it to end the infinite regress of reflection by positing it as constituted already in experience. If the absolute flow maintains the temporality of internal time-consciousness, then its constitution is subject to the same division of time, and therefore it cannot give a presence to itself in the now. Husserl’s attempt to place self-presence in the absolute flow seeks to preserve the self-presence of identity and end the infinite regress of reflection. Further, the implications for the constitution of the absolute flow and Derrida’s critique of the metaphysics of presence would threaten the possibility for ideality in Husserl. Zahavi, by not calling into question the absolute flow, reinforces a metaphysical structure. Hagglund states, “It turns out that the problems inherent to reflexivity question Husserl’s metaphysical postulates, whereas the concept of a pre-reflexive subjectivity is introduced in order to save the idea of a fundamental presence” (Hagglund 2008, p. 55). This attempt to save a fundamental presence rests on a constitution of the absolute flow that allows for the primal impression-protection-retention structure to occur. Following Hagglund’s argument, a synthesized

temporal experience cannot be atemporal. It must be the result of another temporal synthesis. The problem of reflection arises. If the synthesis of experience must be temporal, then an infinite regress arises for each synthesis that must be synthesized by another act. For Husserl the structure of temporal consciousness always faces this problem. Since protention and retention are necessary for any experience, the now cannot ground its own temporal experience and must depend on another instance for its own appearance. This is where the absolute flow is required to constitute a presence upon which temporal experience can happen.

Following Hagglund, an analysis of Zahavi's view shows that the extended view of presence still allows for a self-constituting absolute flow that maintains a metaphysical foundation within the analysis.

The basic problem here is that self-presence and temporal extension are mutually exclusive attributes. Nothing that is temporally extended can ever be present in itself. Rather, it is necessarily divided between the past and the future.
(Ibid, p. 61)

The problem at hand is that the absolute flow cannot hold together the temporal experience of objects without itself being temporal.

Hagglund takes issue with Zahavi's pre-reflective, extended view of consciousness precisely because Zahavi posits the self-manifestation of the absolute flow as potentially atemporal and as necessary for a pre-reflective stance. Hagglund states, the "absolute flow is supposed to put an end to the threat of an infinite regress by being "self-constituting" and thereby safeguarding a primordial unity in the temporal flow" (Ibid, p. 69). Further, "This solution requires that the subject appears to itself through a longitudinal intentionality that is not subjected to the constraints of a dyadic and temporal reflexivity" (Ibid, p. 69). On such a view the givenness of the flow creates a unity between it and temporal experience so as to ground temporal experience and avoid the infinite regress. Self-reflection would appear atemporally. It would not be reflection at all but pre-reflective unity of self to self in

experience. Hagglund argues that although Husserl does state that the flow constitutes itself atemporally, he also claims that the phases of the flow do not coincide with themselves. If the flow cannot coincide with itself then it must be temporal and divided.

No phase of consciousness can intend itself. It is always intended by another phase that in turn must be intended by another phase, in a chain of references that neither has an ulterior instance nor an absolute origin. (Ibid, p. 69)

Self-presence, then, must be divided by temporality at the level of the constitution of the flow. Derrida's point, which Hagglund reiterates, is that time constitutes the subject rather than the subject constituting time.

In conclusion, it is easy to see both the promise and limitations of phenomenology. Its aim to be a presuppositionless philosophy is undermined from within. Both Husserl and later Zahavi grapple with the same issue of the metaphysics of the absolute flow. For Derrida the most radical implications for Husserl's thought are erased, with the result that phenomenology undermines itself. The implications of Derrida's critique allow us to reformulate the constitution of meaning, to place representation and language at the core of all experience, and to rethink any notion of identity. Since every now is the re-presentation of a memory for the future, it is impossible to hold on to what has just passed, and is no longer present, in the now of the present. Temporality divides the self all the way down to the possibility of the absolute flow and the relation of self to self. It is through such a non-metaphysical theory of the time-consciousness that phenomenology may achieve its own goals as a presuppositionless philosophy.

Bibliography

- Derrida, Jacques. (1973) *Speech and Phenomena and Other Essays on Husserl's Theory of Signs* (Evanston: Northwestern University Press)
- Gallagher, Shaun and Zahavi, Dan. (2014) "Primal Impression and Enactive Perception," in: D. Lloyd and V. Arstilla (Eds), *Subjective Time: The Philosophy, Psychology, and Neuroscience of Temporality*. (Cambridge: MIT Press)

- Hagglund, Martin. (2008) *Radical Atheism: Derrida and the Time of Life* (Stanford: Stanford University Press)
- Husserl, Edmund. (1964) *The Phenomenology of Internal Time-Consciousness* (Bloomington: Indiana University Press)
- Zahavi, Dan. (1999) *Self-Awareness and Alterity: A Phenomenological Investigation* (Evanston: Northwestern University Press)

WHY THERE STILL MAY BE SUCH A THING AS TIME-TRAVEL

Samuel Batzdorff

1

Among science fiction tropes, perhaps none is better loved than the time-travel story. In such stories (with some notable exceptions) there is usually a mutability to the past and future—part of the folk understanding of time-travel is an opportunity to go back and change the past, or change the present to have an effect on ‘possible futures.’ We’ve all seen it in movies or read it in books—the hero goes back and changes what has happened hoping to right the wrongs of the past. This may be successful, but upon returning to the future, things have now changed, perhaps in unexpected ways. Assuming the past can be changed, the changes to the future seem unavoidable, of course—one can hardly expect to go back in time, save Socrates from his execution (or some such major thing) and return to the exact same “present” they left. This can create (sometimes in the context of story, sometimes just in the mind of the astute reader) paradoxes, inconsistencies, and other conceptual issues. This is not too big of a problem for fiction, but is generally best avoided in philosophy, where the ironing out of paradoxes is more often our fare than giving rise to them.

Rupert Read (2012) argues that philosophers espousing the possibility of time-travel, such as David Lewis, are mistaken, having been misled as to the nature of time. I will be arguing that Read’s position, rather than being a new one, boils down to a presentist view of the nature of time. I will attempt to show that while presentism may not allow for time-travel, we have compelling reason to accept his view in favor of Lewis’ eternalist position, on which time-travel need not be a problem.

I will begin by going over some basic worries about para-

doxes arising from time-travel and give a quick overview of presentism and eternalism. In section two I'll outline David Lewis' position, which gives a good account of why an eternalist shouldn't fear time-travel. In the third section I will consider Read's argument, and then in the fourth, I will attempt to show that he ends up arguing for a kind of presentism, with the impossibility of time-travel conditional upon that presentism. Finally, in the fifth section, I will argue that with this conditional structure explicitly shown, Read is unsuccessful in establishing the incoherency of time-travel.

In this paper, I will not be concerned with the *physical* possibility of time-travel (still hotly contested by physicists) but only with time-travel's *logical* possibility. Time-travel may not seem, at first glance, to be logically impossible. No obvious paradoxes arise from going back in time to 399 BCE and witnessing the death of Socrates (provided you keep a low profile, of course). But the idea of this kind of changing past events does seem to be problematic: once the past is changed, it is no longer *the* past. By definition, your past is the stuff that happened before now, and if that is changed, *different stuff happened*, and the events no longer meet that definition. In addition to this conceptual issue, logical paradoxes seem to come about as well. "Socrates died in 399 BCE, not 390 BCE" is true. If you go back in time to 399 BCE and smack the hemlock out of his hand, causing him to live nine more years (and perhaps add five or six more chapters to Plato's *Republic*) then "Socrates died in 390 BCE" *will* be true, and we'll either have a contradiction (if history does not "rewrite" the present, and the original fact is still true) or the past and present will no longer be *the* past, and *the* present (i.e. the past we wanted to journey into, and the present we left from). Consider as well that unless we can work out this problem, our earlier intuitions about observation of the past is also under attack: if we say that one hundred people viewed the death of Socrates, then the time traveler goes back in time and gets involved, then won't there now be one hundred and one viewers? Even if no history book records this, the contradictory fact about the past is still a contradiction. After considering

all of this, one is left with the distinct feeling that changing the past *would* be logically impossible, and thus any sort of backwards time-travel should be equally impossible.

The mutability of the past and future is where the problems and paradoxes seem to arise, so the nature of time will play a part in the discussion. In particular, the theories of presentism and eternalism will be relevant.

For the presentist, the current temporal position (i.e. the present) is all that exists. Therefore the presentist universe is ontologically made up only of objects that exist at the current time. A list of everything that exists, for the presentist, would contain Mars, the pyramids, and you (the current reader). The list would not contain Socrates, or Michael Jackson, or ‘the first child born next year.’

Eternalism says more or less the opposite of presentism. For the eternalist, past, present, and future times exist at all points, and it is only our temporal position that changes. Thus, temporality has no effect on ontology, and the eternalist list of all existent objects *will* contain Socrates and Michael Jackson, along with ‘the first child born next year.’ Socrates may not be present to me now, but that is only because his temporal points are disconnected from mine now. That is to say, none of the temporal points at which Socrates is alive overlap with any ones in which I’m alive (assuming I’m not a time-traveler, of course). Eternalists can draw an analogy to three-dimensional physical space—objects in Las Vegas or India are not spatially present to me, while I’m sitting at a computer in California. But the objects in Las Vegas still exist, even if I have no access to them under the current circumstances. Eternalism says that the same is true of objects regarding their temporal position.

2

It is from an eternalist standpoint that David Lewis (1976) tells the story of Tim, a time-traveler who hates his grandfather, and desperately wants to kill him. Before going into the details of Tim's

story, however, Lewis gives us some helpful thoughts on referring to the exploits of a time-traveler. When we speak of time-travel, it will help to consider a “personal time” for the traveler, separate from “external time” that the non-time-traveler experiences. A traveler’s personal time is “roughly, that which is measured by his wristwatch” (Lewis 1976, p. 69). So, if he takes a trip back one hundred years and the trip takes ten minutes, then we can say that it is ten minutes from now in his personal time, and one hundred years before now in external time. Also, if a forty year old time-traveler travels thirty years into the past, there may be two of her and we can say that in 1985 there is, at one time, both a her who is ten years old in personal time (which is how old she is in external time) and a her that is 40 years old in personal time.

Back to the story of Tim. He hates his grandfather, and goes back in time to 1921 (when his grandfather was a young man) with the full intention of shooting him. What happens? Well, one thing we know is that assuming the man is actually the grandfather Tim remembers surviving long enough to hate so much, he didn’t die in 1921, or Tim wouldn’t have known him to hate him. Nobody, time-traveler or otherwise, can kill a person prior to the time that he actually died. Tim cannot kill his grandfather. Tim’s *attempting* to kill his grandfather, however, should be perfectly consistent with the facts of Tim’s grandfather surviving. So he can go back in time (provided he has the technology) and try to kill his grandfather. But we know that Tim fails, so something goes wrong (his gun jams, he slips on a banana peel, etc.). Whatever it is, Tim will not/did not succeed. The important point is this: Tim, upon arriving in the past, changes nothing. He “participates” in the past, creating it but not recreating it. In his personal time, the events are new and he can do whatever he wants. But those events are not new in external time, and whether Tim knows it, whatever he happens to do *is* what happened and caused the present that he left from.

Either the events of 1921 timelessly do include Tim's killing of Grandfather, or else they timelessly don't. We

may be tempted to speak of the “original” 1921 that lies in Tim's personal past, many years before his birth, in which Grandfather lived; and of the “new” 1921 in which Tim now finds himself waiting in ambush to kill Grandfather. But if we do speak so, we merely confer two names on one thing. (Ibid, p. 76)

It should be immediately noted that there is a wrong way to read the above discussion of Tim's misadventures, and we need to make sure we avoid it. This wrong way to read it is something like, “If Tim were to go back in time, gun jamming, and banana peel slipping, etc. would happen to keep him from killing his grandfather.” While no part of a statement like that is false according to this account, it has the wrong implication: that there is, actually, a past without the time-traveling Tim, that he time-travels to, which then (somehow) makes gun-jams and the like happen, to stop him from changing the past. This is *not* Lewis' story. Instead, we should say that, if it comes to pass at some future point, that Tim goes back in time to 1921 then he must have been in 1921 (even if no-one knew this until now). But, in spite of Tim's murderous intent and any possible opportunities for grandfather killing, his grandfather *did* in fact live, so we know that something went wrong with Tim's plan. Somehow or another Tim's grandfather did not die, and whatever led to this true fact about 1921 is *what happened*. We may say that the things which prevent Tim's killing of his grandfather, along with the fact that everything he chooses to do fits with what happened in the “original” 1921 is a series of strange coincidences. But doing so is not actually the right way of thinking about it, akin to the old joke of marveling at the absurd improbability of all of one's ancestors, thousands of people going back thousands of years, living long enough to be able to produce offspring, when so many people do not. This is to say, it is only improbable in retrospect—at the time it is merely happenstance.

To reiterate, the important point of Tim's time-travel is this: no change happens, at least not from any actualized state of affairs. There is, in a sense, a change from how things would

have been had Tim not traveled back in time (a person he bumped in the street may not have been bumped, there would have been one less person around at that point, no one may have tried to kill his grandfather that day, etc.) but this is a counter-factual claim, with no relevant differences from imagining what would have happened had John Kerry won the presidential election in 2004. This is just as possible about the present: I can ask myself “what would I be doing right now had I studied computer science instead of philosophy?” or “how would it be if I were to dye my hair green?” But the change Tim makes in going back in time, to try to kill his grandfather, is no different than the change I made by not dying my hair green.

As Lewis puts it,

You cannot change a present or future event from what it was originally to what it is after you change it. What you can do is to change the present or the future from the unactualized way they would have been without some action of yours to the way they actually are. But that is not an actual change: not a difference between two successive actualities. And Tim can certainly do as much; he changes the past from the unactualized way it would have been without him to the one and only way it actually is. To “change” the past in this way, Tim need not do anything momentous; it is enough just to be there, however unobtrusively. (Ibid, p. 76)

Keep in mind the eternalist viewpoint: that all times exist and only the temporal position of objects changes. It is not a problem then, that while at a 1921 temporal position Tim is epically failing to kill his grandfather, Tim is also an angry teen hating his grandfather at a 1970 temporal position. Tim can also be at a 1980 temporal position building his time machine, and in a 1985 temporal position stepping into the time machine. Tim’s actions and choices throughout time are already existent, and his personal time being out of sync with objective time is strange, but not conceptually problematic.

Tim himself may be surprised, if he has not read Lewis, that

he will fail every time he tries to change the past, perhaps even in ways that seem oddly convenient. Considering this, he may say something wrong-headed like “everything seems to conspire against me changing the past. Nature must attempt to avoid a paradox, and ‘the universe’ won’t let me kill my grandfather.” But this is just Tim’s problem, not nature’s—he has made the mistake of believing there was a non-Tim-affected history that was some way before he got involved.

In this way, Lewis has shown us that, at least for an eternalist, the outlook for time-travel’s logical possibility need not be so grim as it may *prima facie* appear. The idea of participating in the past, and having been a part of it all along, seems to give us the non-standard movement through time without the paradoxes. But not everyone agrees.

3

Read (2012) gives a solidly built argument for the impossibility of non-standard movement through time, particularly movement into the past. Backwards time-travel is important to Read because he feels forward time-travel is uninteresting and it is in moving to the past that we see all the paradoxes show themselves.

Read begins by determining that in order for time-travel to really feel like what we talk about when we say “time-travel,” we’ll need to be able to go backwards. Traveling forward in time is, after all, something we are all currently doing and does not seem immediately relevant to the question. If traveling with a one-to-one correspondence isn’t enough like time-travel for you, then consider sleep or cryogenic freezing. These are ways in which we may seem to “jump” forward in time, yet they are not the sort of “time-travel” that we are interested in discussing here. So, it is not enough to just go forward, we must be able to come back—backwards time-travel should be the goal, as it is what actually interests us (Read 2012, pp. 139-140).

Having established that backwards time-travel is necessary for the concept to really be time-travel, Read focuses his

argument specifically towards that along with the metaphysical and conceptual impossibility of changing the past.

But in order not to have changed the past, and made it something other than the very thing that you wanted to voyage into, you cannot have had any impact at all, not even one so slight that it evaded all records and notice. You cannot have affected the energetics of the atmosphere, the trajectories of light-beams, etc. You must have been entirely subtle. (Ibid, p. 141)

Read's point fits well with the exact intuitive issues we saw come up before: if by traveling to the past we change it, then traveling *to the past* is impossible. No simple rewording or concept shift will get us out of the problem, since we are then still vulnerable to the additional paradoxes that break the causal chain. Following Read's thoughts along these lines, we can easily see his point: the only "safe" travel to the past is an entirely ethereal one: tantamount to watching a video or other record of the past. This seems to be all time-travel can amount to, and, like the sleeping to move forward in time, will leave us unsatisfied (Ibid, pp.141-142). With this clear, there is nothing for Read that can actually deserve to be called time travel since all we can get are mundane things: sleeping to go forward and watching video footage to go back. Anything more than this is logically impossible.

Considerations such as these raised by Read, though *prima facie* worrisome to the would-be time-traveler, seem an easy target for Lewis' argument: if we have already participated in the past, then no paradox arises in going to it. Nothing is "changed," the past remains "our past," and all should be well. Read's clear disagreement with Lewis on these points may seem puzzling but I believe that stems from an unstated assumption throughout the paper: namely, that only the present exists (i.e. presentism is true). While Read does not identify as a presentist in the paper, I will be attempting to show, in the next section, that his argument is a presentist one.

Read's stated diagnosis of the issue is that belief in time-travel stems from our incorrect (though understandable) application of spatial talk and reasoning to time. By using space as a metaphor for time, we have ended up confusing ourselves. He hopes, admirably, to not take an official stand on time metaphysics as can be seen here:

Metaphysics and ontology of time (and similarly of other "things") are rash over-generalisations of or reifications of fragments of the "metaphorical patchwork" that time is. ...these would-be metaphysics of time take one particular temporal conceptual metaphor (invariably, with a spatial basis: such as the concept of having parts, or slices), and unwisely project it to "capture" the essence of temporality itself. (Ibid, p. 146)

At no point does Read bring up "presentism" or "eternalism" directly. However, by not taking a stance on this issue, Read leaves us to extrapolate what he feels about the metaphysics of time. I believe that Read does take a stance implicitly, as can be seen in his talk of non-present times:

I stressed an asymmetry between our desires vis-a-vis "travel into the future" on the one hand and "travel into the past" on the other (namely: the desires, respectively, to fix what is fluid, and to fluidise what is fixed)... The desire to visit the future is the fantasy that there is something definite going on there that is already visitable... (Ibid, p.149)

This seems to show disbelief in an existent future and, by extension, existent future objects—if nothing is "going on" there (in the ontological sense of nothing, as I assume Read means here), then there is no object "there." So, Read does not seem to believe in an existent future:

"Time-travel" supposes that past events, "events in the past," are still somehow there now. The past allegedly

exists now, because you can go to it now (or could, if you had a time-machine). But: those events are over. You want to be present at events that are over... That is the impression that one needs to overcome: that, in the sense in which we speak of tourism in space (or indeed space-tourism!), there can be time-tourism. The past is not an undiscovered or unexperienced country. 1900 is no more truly “out there (somewhere)” than is that elusive room that we see “inside” the mirror. Do not let deflationary, potentially-therapeutically-useful expressions such as “The past is fixed” mislead you into thinking that the past is still there waiting to be visited. It is precisely because there is (now) no there there that the past is fixed, unalterable. (Ibid, pp.149-150)

So the past, like the future, is not really “out there somewhere.” The past is “fixed” only in that it does not, strictly speaking, exist. This is my interpretation of “there is (now) no there there...”—that in the present moment, which we all exist in, there is not a past which still exists.

Through Read’s statements, his view on the nature of time starts to become clear. His concept of time, freed from “spatial metaphors” is one in which, rather than being currently existent places that we could visit, the past is gone and the future has not come into being yet. The future is rendered unknowable and the past is rendered fixed due to there being no future and no past “now.” Thus, there is a present but no existent future and past. Stated this way, I believe that what Read describes above are some of the central tenets of presentism and that Read’s argument against time-travel takes a presentist stance.

Considering this stance, we can clearly see Read’s argument, and it is a strong one—that time-travel is incoherent as a concept, since we live in a presentist world. Presentism *must* be true for his argument to go through. There is a weaker argument within his broader one: if presentism is true, then time-travel is incoherent. This is by no-means an uncommon position. Non-standard movement through time is clearly a sketchy proposal within presentism.

Since there is no future and no past currently existent, it is *prima facie* incoherent to think something has “come from the future” or will be “going to the past” or, at any point, to think “something in the future will go to the past.” While there has been some argument against this position (such as Kellor and Nelson (2001)) I have no problem with granting it, in order to focus on the stronger argument. It is quite straightforward:

1. If presentism is true, then time-travel is incoherent as a concept.
2. Only the present exists (i.e. presentism is true).

Therefore,

3. Time-travel is incoherent.

The argument is clearly valid and I’m granting the first premise. So, all the pressure is on premise 2. Read’s argument for this premise, as I will show in the next section, is insufficient.

5

If my view of Read is correct, then we can consider his worries about changing the past to be a *reductio* argument for presentism:

- i. If non-present times exist (i.e. if presentism is false), then one could, in principle, travel to the past.
- ii. Upon doing so, the person would be free to act as they chose, and in acting would undoubtedly change the present and future.
- iii. This leads to paradoxes (i.e. logical contradictions).

Therefore,

- iv. There are no non-present times (i.e. presentism is true).

If this argument is Read’s intention, it is countered fairly well by Lewis’ story, a view that allows for non-present times and avoids paradoxes. Read comments briefly on Lewis, pointing

out that he “abjectly fails to explain *why* it is impossible for a time-traveller to kill their grandfather” (Ibid, p. 148). This objection alone, I believe, shows the tension between the eternalist and presentist camps. If Read is arguing from a presentist position, then the worry is understandable: if Tim has somehow managed to find his way back to 1921 (despite the seeming impossibility of that trip within presentism) then 1921 would be “the present,” and the only time which exists. Tim is currently in the only existent time, with as much freedom of action as any person ever has, and seemingly free to shoot his grandfather, and only an outside force should be able to stop him. To an eternalist, on the other hand, the objection itself is incoherent. Tim does not kill his grandfather, and we know this because his grandfather did not die, and he didn’t die because no one killed him, and that’s just the way time *is*. There is a problem in this for the presentist, but not for the eternalist. As such it is question begging as an argument for presentism.

If, as I have argued, Read’s stance collapses into presentism, then his position against time-travel is far from novel. Time-travel’s being possible within eternalism and impossible within presentism has long been the “standard” view on the matter. The arguer against time-travel can bring up many paradoxical consequences of mucking about in the space-time continuum, just as Read does here. But (as with the grandfather paradox) many of these logical problems lose their teeth when given an eternalist explanation, as Lewis does.

All of which is not to say that Lewis’ picture does not have some problems. Backwards and circular causation are issues for time-travel and, while Lewis tackles these to his own satisfaction, many still feel that paradoxes arise. Even without paradox related worries, one could worry that the causal connection between Tim’s arrival and departure is too weak to call it “travel” and to call the “departing” and “arriving” Tims a single entity.

It may well be that presentism is true and that may entail the impossibility of time-travel. But I believe that we do not yet have a good reason to take a stand either way and I don’t feel that

Read's argument has given us any need to accept the impossibility or incoherence of the idea of time-travel.

Bibliography

- Keller, Simon and Nelson, Michael. (2001) "Presentists Should Believe in Time-travel," *Australasian Journal of Philosophy* 79: 333–345.
- Lewis, David. (1976) "The Paradoxes of Time-travel" in: *Philosophical Papers Volume II*, pp. 67-80 New York: Oxford University Press.
- Read, Rupert, "Why There Can be No Such Thing as Time Travel" (2012) *Philosophical Investigations* 35:2 138-153

RECONSIDERING TWO NOTIONS OF ANALYTICITY

Neil Sanchez

INTRODUCTION

Consider two sentences: “All bachelors are unmarried men” and “Two halves make a whole”. Most philosophers hold such sentences to be of a special class of “analytic” sentences. Analytic sentences are classically defined as true in virtue of their meaning alone. Such sentences are typically held to vindicate the possibility of a priori truths that are justifiable without any empirical investigation. These are contrasted with synthetic sentences that require empirical investigation in order to determine their truth. The analytic-synthetic distinction is a semantic notion that has been a part of western philosophy’s foundation since at least the Modern era. However, in “Two Dogmas of Empiricism” (TD), W. V. Quine refutes the notion of analyticity by showing it to fundamentally rely on metaphysical commitments that run contrary to an empiricist research program. In other words, analyticity is committed to things like meanings and synonymy, which assume each other in their explanation. In the aftermath, philosophers were left with the choice of either buying into the rejection of analyticity or precariously employing the distinction with no adequate response to Quine’s objections.

Recently, an attempt to rescue the notion of analyticity has been made by Paul Boghossian in “Analyticity Reconsidered” (1996). In order to do this, he articulates a distinction between two notions of analyticity—metaphysical and epistemic analyticity. Expanding on the classical definition, Boghossian asserts that a statement is analytic in a *metaphysical* sense if “in some appropriate sense, it *owes its truth value completely to its meaning*, and not at all to the facts” (1996, p. 366). Alternatively, a sentence

is analytic in an epistemic sense if “grasp of its meaning alone suffices for justified belief in its truth” (Ibid.). With this distinction in place, he argues for three claims: First, he argues that Quine’s arguments against analyticity only apply to the metaphysical notion and not the epistemological notion. Second, he claims that epistemological analyticity explains the apriority of our logical terms like “and”, “if”, and “then”, therefore making possible the apriority of conceptual truths. Finally he argues that the rejection of epistemic analyticity entails the acceptance of skepticism about meaning. This entire project hinges entirely on the distinction between the two notions of analyticity, for separating the epistemic notion from the apparently inferior metaphysical notion allows the notion of analyticity to escape Quine’s attacks. Moreover, it characterizes Quine’s work as a source of insight for understanding analyticity itself.

I argue that Boghossian’s distinction is incorrect. Specifically, I claim that his understanding of conventionalism and the inherent problems it causes for metaphysical analyticity are incorrect and that analytic truth is plausible in the context of conventionalism, despite the challenges he identifies. Furthermore, I claim that metaphysical and epistemic analyticity are different ways to understand the classical definition of analyticity, but they do not represent two different ways that a sentence can be analytic. The two are inseparable, contrary to what Boghossian suggests. This is significant because the inseparability of the two notions of analyticity provides a complete account of analytic truth. We can know how these statements are true, their commitments, and the justification for their truth. Most importantly, this shows that epistemic analyticity is not isolated from metaphysical analyticity in a way that escapes the attacks made by Quine.

In order to argue this view, I will begin with a defense of the metaphysical notion against the accusations of incoherence entailed by conventionalism. In this section I argue that Boghossian conflates the manners in which truth and necessary truth are determined according to conventional stipulation. In the section that follows, I argue that metaphysical and epistemological analyt-

icity are two parts of a complete account of analyticity that is not accomplished by the definition ‘truth in virtue of meaning’. In order to do so, I argue for two claims: All metaphysically analytic sentences are epistemologically analytic and all epistemologically analytic sentences are metaphysically analytic. The basis for substantiating these two claims relies on the role of the logical words in Quine’s distinction between logical truths and hard cases. In the last section I entertain some objections and worries.

SECTION 1: METAPHYSICAL ANALYTICITY REVISITED

According to Boghossian, metaphysical analyticity is a dubious, incoherent notion and therefore is not worth pursuing. To this extent he is in agreement with Quine’s refutation of analyticity in TD. More specifically, the idea that knowledge of meaning is somehow sufficient for knowledge of truth appears implausible without any appeal to ‘the facts’ or the obtainment of some state of affairs within the world. He holds that analytic statements possess a mystery about them and that metaphysical analyticity make this obvious by making it explicit that if the truth of a sentence is determined completely by its meaning, then its truth does not rely on any appeal to ‘the facts’. Furthermore, this apparent mysteriousness is complicated by conventionalism about meaning—the claim that the truth of what a sentence expresses depends on its being expressed by the sentence.

Conventionalism about meaning is the view that the meaning of our terms is grounded in agreements between people in some implicit or explicit way. This would mean that “bachelor” means “unmarried man” by convention and the sentence (R), “All bachelors are unmarried men”, is true according to convention and knowledge of its truth is justified by its meaning. This directly contrasts with the idea that facts in the external world ground our truth claims and our knowledge claims. For Boghossian, the problems of metaphysical analyticity arise from the way in which the truth of a sentence is determined according to convention-

alism. According to conventionalism, knowledge of meaning is knowledge of truth without justification by facts. A second look at metaphysical analyticity, however, will show that the notion is plausible within the bounds of conventionalism. Moreover, it will show that Boghossian conflates the way in which truth and necessary truth are determined within conventionalism.

Metaphysical analyticity appears incoherent according to conventionalism.

[Metaphysical analyticity entails] that the truth of what the sentence expresses depends on the fact that it is expressed by that sentence so that we can say that what is expressed wouldn't have been true at all had it not been for the fact that it is expressed by that sentence. (Boghossian 1996, p. 365)

Boghossian identifies the following meaning-truth truism as an explanation for conventional stipulation: for any statement *S*, *S is true iff for some p, S means that p and p*. Conventional stipulation is just how a term gets its meaning by conventional agreement. It is arbitrarily decided that a given term means what it does according to both the act of stipulation and the continued use of the term according to its intended meaning. For example, “bachelor” means what it does because it was conventionally stipulated and there has been consistent use of it in agreement with its stipulated meaning. Consequently, (R) is true because the term “bachelor” means “unmarried man” and none of the terms within the sentence deviate from their stipulated meanings.

The claim that metaphysical analyticity is incoherent is motivated by the way that conventionalism determines the truth of a sentence. Boghossian presses the problem further by granting two features of the meaning-truth truism as the normal dependence of truth on meaning. First he grants that the meaning of component terms within a given sentence fix what is expressed by that sentence. Second, he grants that our ordinary use of the terms within the sentence determines its truth-value. The main problem with this approach, according to Boghossian, is that it

can't explain how the mere fact that S means that p makes it the case S is true, or that what is expressed wouldn't have been true at all had it not been for the fact that it is expressed by that sentence. He contends that in order for the sentence to be true, it also needs to be the case that p. In other words, (R) is true because the words mean what they do and because it is the case that unmarried men are bachelors in some appropriate sense. How else would "bachelor" mean "unmarried man" if there were no unmarried men that were bachelors? It would seem that for a term like "bachelor" to possess a specific stipulated meaning ultimately requires some sort of appeal to the facts—in this case that there exist unmarried men. This is where Boghossian makes a mistake.

The problem is generated by a misunderstanding about how the truth of a sentence is determined and how necessary truths are established, according to conventionalism. Boghossian notes that conventionalism about meaning developed in the context of logical positivism.

[The logical positivists] attempted to show that that all necessities could be understood to consist in linguistic necessities, in the shadows cast by conventional decisions concerning the meanings of words. Conventional linguistic meaning, by itself, was supposed to generate necessary truth; a fortiori, conventional linguistic meaning, by itself, was supposed to generate truth. (Boghossian 1996, p. 365)

Take the sentence (S): "Either snow is white or it isn't". Boghossian asserts that snow's being white is true prior to our stipulating the meanings of the terms in (S). In other words, he is arguing that we do not need an act of conventional stipulation for snow to be white. Snow would be white regardless of whether anyone should express it. This directly opposes the conventionalist claim that what is expressed wouldn't have been true at all had it not been for the fact that it is expressed by that sentence.

However, there is a critical difference between sentences like (R) and (S). The second is a logical truth or tautology whereas the first requires the translation of its synonyms in order to be reduced

to a logical truth. Kathrin Gluer exploits a vulnerable point in Boghossian's analysis of (S). She contends that

...conventionalism about truth does *not* follow from conventionalism about necessary truth; what is generated by linguistic meaning, according to conventionalism, might be the *necessity* of a statement but not its *truth*. Truth may be given independently, that is, and necessity added. (Gluer 2003, p. 4)

She substantiates this claim by distinguishing between conventionalism about truth and conventionalism about necessary truth. According to the former, truth is generated in the way Boghossian describes for analytic sentences, but also for all sentences. Alternatively, the latter accounts for the necessity of stipulated terms independently of the sentence's truth-value. Conventionalism about necessary truth can tell us that the terms in (R) and (S) mean what they do and that their meanings should hold necessarily. It does not, however, commit itself to a view that some combination of component terms should suffice to determine the truth-value for every well-formed sentence of a language. Conventionalism about necessary truth only accounts for the necessity of meanings and that said meanings contribute to the truth of a sentence. It does not, however, commit to how the truth-value of a sentence is determined.

Boghossian refers to conventionalism and the linguistic doctrine of necessary truth synonymously. However, there is a difference between conventionalism about truth and conventionalism about necessary truth and he does not adequately distinguish between these notions. Consequently, metaphysical analyticity is plausible despite Boghossian's analysis of the meaning-truth truism because conventional stipulation establishes the necessity of a given term.

For any statement S, there can be a corresponding fact, no matter how this fact is created, no matter whether it is metaphysically dependent on our meaning p by S in any

sense. The idea simply is that, for some statements, this is a necessary fact, a fact that, so to speak, obtains in every possible world. Therefore, the facts cannot possibly make a difference to the truth-value of such a statement. (Gluer 2003, p. 4)

In other words, ‘bachelor’ will always mean ‘unmarried man’ so long as the user of the term doesn’t deviate in her use of it. Moreover, its extension in every possible world is unmarried men because in order for it to not refer to unmarried men, a user has to deviate from its intended use. It follows that metaphysically analytic statements are not as mysterious as Boghossian paints them, because the necessity of conventional stipulation commits to that fact obtaining. The ordinary use of our terms does not need to rely on the facts because the facts are already built into them through stipulated meanings.

SECTION 2: THE INSEPARABILITY OF METAPHYSICAL AND EPISTEMOLOGICAL ANALYTICITY

Having established that metaphysical analyticity is not a mysterious notion that warrants rejection, we can turn to how it is inseparable from the epistemological notion. Consider the following two sentences in TD (Quine 1951, p. 23):

- (1) No unmarried man is married.
- (2) No bachelor is married.

The contrast between (1) and (2) explain why (R) is not a logical truth like (S). These two sentences exemplify two kinds of analytic statements. The first is an example of a logical truth. Quine explains that statements like (1) are not only true as they stand; they are also true under all reinterpretations of their components. For example, we can replace equivalent terms for “man” or “married” and the sentence will always be true. The truth of the sentence turns on the logical particles. Logical particles include any of the constant logical terms such as ‘no’, ‘if’, ‘then’, ‘and’,

etc... The second sentence is an example of what Quine identifies as hard cases. These statements are complicated by attempts to explain translation between extralogical synonym-pairs. Examples of this are 'bachelor' and 'unmarried man'. Quine simply defines these synonym-pairs as the type that give rise to the second class of analytic statements.

First, if Quine had not categorically rejected the notions of meaning or synonymy, logically true statements would be perfectly acceptable analytic statements because their meaning is transparent and their translation is uncontroversial. These can be thought of as properly formed analytic sentences because their truth-value is determined by their logical particles rather than from their constituent terms; i.e. terms like 'bachelor'. The truth-value of a sentence of this kind depends upon its logical particles rather than its constituent terms, and so in such cases truth-value is determined by structure rather than by meaning.

The existence of extra-logical synonym-pairs motivates the difficulty of explaining the hard cases. In contrast to the first class of sentences, the truth-value of the second class of sentences is not determined by its logical particles. Instead, the role of the singular terms shifts truth determination away from the statement's structure. The synonymy of pairs of words like 'bachelor' and 'unmarried man' rests exclusively on the meanings of these terms. What makes us hold (2) to be reducible to (1) is our command of the meanings of the singular terms.

The crucial difference between these two classes of statements is that, on one hand, class-one statements turn on the meanings of the logical words, i.e. the inventory of logical particles that we take to have a fixed meaning through all translations. On the other hand, class-two statements rely on their singular terms because extra-logical synonym-pairs like 'bachelor' and 'unmarried man' aren't transparent in the same way as terms in (1). The upshot of this contrast is that class-one statements seem to suppose some semblance of meaning that applies to the logical words despite Quine's general rejection of meaning. According to a conventionalist framework, it would be the case that the logical

particles Quine identifies get their meanings from conventional stipulation. This explains how the necessity generally attributed to them as logical terms is simply the same necessity that follows from stipulation. Through conventional stipulation, sentences like (1) and (2) are metaphysically analytic because their truth-value is determined by their meaning. The missing part of this story, however, is an explanation for how we can be justified in accepting the truth of sentences like (2) in the same way as (1). This is where epistemic analyticity is invoked.

I propose, instead, that the two notions of analyticity can be taken to address different questions about analyticity. Metaphysical analyticity addresses questions such as what an analytic truth is and what differentiates it from merely logical truth. Alternatively, the epistemological notion addresses how we can accept hard cases such as (2) to be reducible into logical truths. According to this view, metaphysical analyticity accounts for how a statement's truth-value is determined and epistemic analyticity is invoked in order to explain our justification for accepting such truths. I claim that metaphysical analyticity is necessary and sufficient for epistemic analyticity. In other words, metaphysical and epistemic analyticity entail each other.

Metaphysical analyticity explains what an analytic statement is. Specifically, it explains how an analytic statement is structured and how its truth-value is determined. For example, logical truths like (1) provide a template for how the truth-values of analytic statements are determined. However, such statements do so trivially by turning to their logical particles rather than their singular terms, since the terms are identical. Logical truths are true under every interpretation because the same truth-value is generated regardless of what term is used to replace "married." Moreover, our entitlement to believe the truth of a class-one statement is established in an uncontroversial way because the meanings of the logical terms are taken for granted. This is due to their fixed meanings.

Compared to the metaphysical notion, epistemic analyticity appears more useful in instances where the transparency between

subject and predicate terms isn't as obvious. For example, class-two statements aren't as straightforwardly interpreted as logical truths. Epistemic analyticity is invoked where metaphysical analyticity isn't obvious because synonym-pairs like 'bachelor' and 'unmarried man' appear to rely on translation for their logical truth. The synonymy of 'bachelor' and 'unmarried man' isn't as obvious as the fact that 'unmarried' is the negation of 'married'. Interpreting how a statement like (2) reduces into a class-one statement becomes elusive when trying to explain translation according to meaning or synonymy in the face of Quine's attacks. Rather, it seems as though we can accept the truth of a class-two statement if we are given compelling testimony or other appropriate means of believing that translation into a logical truth can be achieved. Ultimately, knowing that sentences like (2) are true because they transform into sentences like (1) entails that there is some account for how (1) gets its truth-value. In other words, metaphysical analyticity is entailed as an explanation for how truth is generated. Accordingly, all epistemic analyticities are metaphysically analytic and all metaphysical analyticities are epistemologically analytic.

All metaphysically analytic sentences are epistemologically analytic. Metaphysically analytic sentences possess a certain triviality. However, this is because triviality is characteristic of analytic truths and, a fortiori, the a priori truths they purport to vindicate. Metaphysical analyticity simply explains how the truth-value of a sentence is determined. This is made obvious in the case of class-one analytic statements.

When examining (1) we can see that its truth is determined by an agreement between subject and predicate terms. However, this obtains without controversy since they employ the same term "married." Once we know that this sentence must be true, since it is a logical truth, we are also justified in believing its truth. Our entitlement comes from the fact that the sentence is true under every translation in virtue of the logical terms rather than the singular terms. Furthermore, we know that the logical terms have necessary truth because of conventional stipulation or some other

account that explains their necessity. In other words, we can only be sure that the truth of (1) is owed completely to its meaning and not at all to the facts only if our grasp of what it means suffices for justified belief in its truth. This latter portion, however, is achieved unconvincingly because we rely on the logical particles of the sentence rather than its singular terms.

Similarly, all epistemologically analytic sentences are metaphysically analytic. The difference here is that epistemological analyticity seems to have more utility than its metaphysical counterpart. However, it ultimately relies on some notion of metaphysical analyticity in order to give a full explanation of how the sentence is analytic in the first place. Metaphysical analyticity follows from epistemic analyticity. For sentences such as (2), we have to turn our attention to the singular terms because the sentence's translation into a logical truth isn't made obvious by the component logical terms within the sentence. This is precisely the difficulty posed by extra-logical synonym-pairs that Quine identifies in class-two statements. Metaphysical analyticity is entailed by (2)'s being transformable into (1). Grasp of the meaning of (2) alone suffices for justified belief in its truth only if it owes its truth completely to its meaning and not at all to the facts.

SECTION 3: OBJECTIONS

Having established the inseparability of the two notions, I will entertain two objections: first a general objection, and then an objection that Boghossian might make.

The first objection is to insist that the account of analyticity outlined in the previous sections isn't substantially different from Boghossian's. Conceiving of the two notions as alternative interpretations of "true in virtue of meaning alone" might be thought to be identical to Boghossian's characterization of them as different ways sentences can be analytic. He says that they're two distinct notions, or two different ways analyticity obtains. It follows that what Boghossian suggests is that we ought to abandon trying to understand analyticity in the metaphysical sense because it is

riddled with problems. The metaphysical notion is how Quine understands analyticity and thinking of it in the metaphysical sense opens analyticity to the attacks made in TD. Instead, conceiving of analyticity in the epistemic sense allows us to have a notion of analyticity that is separate from the problematic metaphysical sense and therefore not refuted by Quine. There are two responses to this objection.

First, recall that Quine rejects analyticity categorically. This is a crucial part of his entire project. For Quine, the very idea of truth in virtue of meaning is incoherent because of his rejection of meaning and synonymy in TD. Moreover, Boghossian is committed to a stronger claim than that analyticity can simply be interpreted in two ways. He is claiming that metaphysical analyticity and epistemic analyticity are two different ways in which a sentence can be analytic. They're separate because they are indicative of two distinct kinds of analytic sentences. This point is addressed further in the second objection.

Second, the view espoused in the previous sections is not an attempt to rescue the doctrine of analyticity from Quine. Instead, all that is accomplished is to reinforce the difficulty of overcoming the attacks made in TD. Articulating epistemic analyticity as an alternate way of understanding sentences that are also metaphysically analytic does not amount to an account of analyticity similar to Boghossian's. That portion of his project occurs in his account of "implicit definition" and is reliant on the truth of his distinction between metaphysical and epistemic analyticity. That portion is outside the scope of this paper. The primary aim of my argument is to organize how we think of the same core definition of analyticity.

The second objection considers Boghossian's defense of epistemic analyticity as a separate and distinct notion from its metaphysical counterpart. In fact, claiming that all epistemologically analytic sentences are metaphysically analytic is directly at odds with Boghossian's claim that "there could be epistemologically analytic sentences that are not metaphysically analytic" (2003, p. 2). To fully grasp this objection, we need to see what

Boghossian is responding to. He makes this claim in response to an objection posed by Margolis and Laurence, who claim that metaphysical and epistemological analyticity are on a par.

Recall the meaning-truth truism that for any statement *S*: *S is true iff for some p, S means that p and p*. The on-a-par objection is as follows:

If *p* really is an independent fact that makes *S* true, then just knowing that *S* means that *p* couldn't suffice for the needed justification; one would also need to be justified in believing that *p*. In other words, so long as the truth of *S* isn't merely a matter of what it means, then grasping its meaning can only be (at best) part of the story about why one is justified in holding it to be true. The other part—and by far the more important part—concerns one's epistemic access to *p* itself and why one is justified in believing *p*. (Margolis and Laurence 2001, p. 294)

This objection denies that epistemic analyticity enjoys some privilege that does not extend to metaphysical analyticity. In other words, we do not have a complete account of analyticity by accepting that if what makes a statement *S* true is some extra-linguistic fact *p*, then just knowing that *p* gives us warrant for believing *S*'s truth. Rather, justification for believing that *p* is required as well.

Boghossian takes the challenge to propose an unsound principle:

So long as the truth of *S* isn't merely a matter of *F*, but is also a function of *G*, then being justified in believing *F* can only be (at best) part of the story about why one is justified in holding *S* to be true one would also need to be justified in believing *G*. (2003 p. 3)

To be clear, “a matter of *F*” is “knowing *p*” in the above objection. Alternatively, “a function of *G*” appears to explain our justification for knowing *p* in the same way as the on-a-par objection. The problem of what makes justification for knowing *p* (i.e. a function

of G) an additional requirement to simply knowing p (i.e. a function of F) for accepting the truth of S echoes the problem of the meaning-truth truism generated by conventionalism.

Having established the preliminaries, the core of the second objection is Boghossian's use of the statement 'This is water' as an example of an epistemically analytic sentence that is not metaphysically analytic:

The truth of the sentence "This is water" isn't merely a matter of how some substance looks or feels it is also a matter of being H₂O. However, it doesn't follow that I could be justified in holding some stuff to be water without first being justified in believing it to be H₂O. (Boghossian 2003, p.3)

For 'This is water' to be epistemologically analytic, grasp of its meaning alone should suffice for justified belief in its truth. And, since the sentence isn't a straightforward logical truth like a class-one statement, grasp of its meaning would have to come from the singular terms. Knowing that 'water' refers to something that appropriately looks and feels like what we ordinarily understand to be water is granted to the speaker. What the speaker indicates to be 'water' is the matter of F whereas the water actually being H₂O is the function of G according to the unsound principle. Such an explanation appeals to the facts by means of the function of G as a source of epistemic justification, so that when the speaker indicates that water is being referred to, the referent is in fact water.

The drawback of this example is that it doesn't provide an explanation of what distinguishes epistemic analyticity without falling back on the metaphysical notion. 'This is water' can be transformed into a logical truth if 'water' could be substituted for 'this'. The modified sentence, 'Water is water' is a logical truth in the same way as (1) but it is not entirely clear how 'This is water' is a hard case in the same way as (2). Moreover, why water actually being H₂O is even relevant to the truth of this sentence is somewhat of a mystery. If the speaker is using the term in a way that intends to preserve its conventionally stipulated meaning,

then water's being H_2O is a feature of the term 'water'. There is no advantage in knowing that the speaker refers to something that is both appropriately water-like and is H_2O . The only role it could serve is to confirm that the speaker's indication of 'this' is merely accurate. It seems that a demonstrative is not a good way to explain how an epistemologically analytic sentence can exist in a way that is not metaphysically analytic as well.

CONCLUSION

Defending the metaphysical notion of analyticity is not meant to take away from the contributions of epistemic analyticity as an effort to better understand analytic truth. Rather, all that is offered here is a view that treats the two notions as alternative ways to read what is meant by analytic truth. The metaphysical notion accounts for what an analytic truth is and is not complicated by conventionalism. This is a characteristic of all analytic statements and it is identified through Quine's analysis of (1) and (2), which states that analytic statements ultimately reduce to logical truths. The feature of reducibility to logical truth explains the structure of analytic truth and how truth-value can be determined without an appeal to the facts. Moreover, the analysis of the metaphysical notion within a conventionalist framework also shows how the facts could not possibly make a difference to the truth of the statement. Lastly, it is not entirely clear how a sentence could be epistemically analytic without being metaphysically analytic as well.

Bibliography

- Boghossian, Paul. (1996) "Analyticity Reconsidered," *Nous* 30, pp. 360-91
- Boghossian, Paul. (2003) "Epistemic Analyticity: A Defense," *Grazer Philosophische Studien* 66, pp. 15-35
- Gluer, Kathrin. (2003) "Analyticity and Implicit Definition," *Grazer Philosophische Studien* 66, pp. 37-60
- Margolis, Eric; Laurence, Stephen. (2001) "Boghossian on Analyticity," *Analysis* 61, pp. 293-302
- Quine, W.V.O. (1951) "Two Dogmas of Empiricism," *The Philosophical Review* 60, pp. 20-43.

THE INCOMPATIBILITY OF COMPOSITIONAL NIHILISM AND DIRECT REFERENCE THEORY

Lena Becerra

INTRODUCTION

In this paper I will argue that a particular fictionalist resolution to the problem of composition is incompatible with Direct Reference Theory (DRT). The problem with composition arises because if an object is composed of smaller objects, then there must be not only the number of smaller parts, but another object in addition to them, that is the whole, occupying the same space. This results in too many things in the same space. While common sense assumes the existence of composite objects (i.e., objects composed of smaller objects), it is not clear that common sense is reflective. The fictionalist solution, which suggests another way of interpreting what our common sense language actually commits to, entails that there are no composite objects. Fictionalism is in many ways plausible, however I will argue, that the solution is incompatible with DRT, in terms of how language refers.

THE COMPOSITION PROBLEM

In this paper, perfectly simple entities (i.e., entities without proper parts) will be called *simples*. Objects that are not perfectly simple are referred to as *composite objects*. That is, they are composed of proper parts. The question this paper is concerned with is the question of whether there are any composite objects. The question of existence I intend to address is that of existence *simpliciter*. What I mean by existence simpliciter is a state of simply existing unconditionally. I will also not consider possible objects in other possible worlds. The question addressed in this paper is specific to

actual, physical objects in this world only.

Rosen and Dorr (2002) introduce the problem of composition best, in my opinion. Suppose that in an otherwise empty space, there are three “point particles” (a, b, and c) that are perfectly simple. Suppose further that two of the three simples (a and b) are stuck together and act as a unit. Then the question arises, *how many material objects are in the specified space* (Rosen and Dorr 2002, p. 151)?

Common sense allows for four objects (a, b, c and ab as a unit). The object that is ab is a composite object, (i.e. a single object composed of smaller objects). Universalism allows for seven objects (a, b, c, a+b, b+c, a+c, a+b+c) based on unrestricted mereology where any set of physical objects always form a composite object. Compositional nihilism allows only for a, b and c individually and denies that composite objects exist.

Universalism	Common Sense	Compositional Nihilism
a	a	a
b	b	b
c	c	c
a+b	ab	
b+c		
c+a		
a+b+c		
(7 objects)	(4 objects)	(3 objects)

When we apply a consistent rule to composition, like either universalism or compositional nihilism, the result is either too many or too few objects, respectively, in relation to what we intuitively believe exists. While universalism may be too open to accepting any combination of parts as an object (even things that do not seem to have any relation to each other), nihilism appears too restrictive since nothing will have grounds to be an object if

it has any part that is smaller than the whole. Universalism will allow for things to be considered objects that we would never intuitively accept, and nihilism will deny that objects we are familiar with exist at all. The only problem with relying solely on common sense for these answers, though, is that there is no consistent rule for when composite objects exist and when they don't, except to say that they seem to exist or not exist. Our intuition has nothing, then, to be measured against, and no way to be proven. Therefore, it seems arbitrary.

It may seem obvious that there are composite objects like tables, and cats. To account for what we might mean by our common sense perspective, where smaller objects together can compose a single object, we might adopt David Lewis's (1991) view that composition in regard to the relation between the parts and the whole, can act like identity (p. 82). In this view the existence of the whole is not additional to the existence of the composite parts. The only problem with this view is that composition and identity are logically very different relationships. Composition is a many-to-one relation, whereas identity is a one-to-one relation. It is logically impossible to say that many things are *equal* to one thing. Therefore, identity and composition cannot be used interchangeably. This can be reasoned as follows:

(I1) Composition is a many-to-one relation.

(I2) Identity is a one-to-one relation.

Therefore,

(I3) The relation of composition is dissimilar to identity in exactly the way needed to resolve the composition problem.

Composition results in another object over and above the composite parts. So this fails to solve the problem of *additional* objects such as ab or $a+b+c$ over and above the three simples a , b and c in the original question. Lewis's 'composition as identity' fails to solve the problem of too many versus too few *things* in the

original question.

Another way to justify the commonsense view is to adopt Ned Markosian's (1998) brutal composition view, which holds that for any object composed of proper parts, it is simply a brute fact that there is a composite object. This argument is appealing because it enables us to claim that the composite objects that actually exist are precisely the ones that our intuitions tell us exist. However, brutal composition is argued for simply by proving that there is no acceptable answer to the special composition question (i.e., 'What necessary and jointly sufficient conditions must any parts (xs) satisfy in order for it to be the case that there is an object composed of those parts (xs)?').¹ Markosian offers a convincing argument for the claim that there is no true non-trivial answer to the special composition question. He argues that in order to provide an answer to the question, the answer would simply enumerate each individual case of composition. He argues the use cases for composition are so numerous that they render the answer infinitely long. Markosian concludes then, that since there is no useful answer to the special composition question, we must accept composition as a brute fact. However, his argument by elimination is not actually *a proof* of brutal composition; it is only proof of the lack of an acceptable answer the question. Another problem with Markosian's view is that it suggests that an object appears when we see one, and is thus arbitrary and anthropocentric. The commonsense view does not seem to be well motivated. I will now argue that compositional nihilism is a better view.

COMPOSITIONAL NIHILISM

In my opinion, compositional nihilism is the only answer to the question at the beginning of the paper that is tenable. If in a specific region of otherwise empty space, there are three simples, then there are three material objects present. If simples a , b and c are present, then there are not four or seven objects—there are only three. I will illustrate with an example. If on my plate there were two apples and one orange, and the two apples got stuck

together, which of the following would you say is true?

- (a) There are three things on my plate (apple, apple, orange).
- (b) There are four things on my plate (apple, apple, orange, apple – apple).
- (c) There are seven things on my plate (apple, apple, orange, apple – apple, apple – orange, orange – apple, and apple – apple – orange).

Now the common sense answer is that there are three things on my plate. This example is not exactly consistent with my argument since apples and oranges, themselves, are composite objects and would not exist if compositional nihilism were true. However, the example serves to illustrate how our intuition about things we *do* perceive in our daily lives would lead us to the logical conclusion that only a, b and c exist in the initial question on composition. It is my opinion that if we could see at the sub-atomic level, we would count those objects in this manner as well. The only reason we count objects the way we do is because we see at a specific level of size. But this is arbitrary. Therefore, it seems to me that any belief in composite objects will be arbitrary in identifying what actually exists. The only way to refer to what material objects actually exist, without requiring an arbitrary definition or double counting, is to acknowledge only the objects that have no parts.

Common sense rules of composition are vague at best. The same logic which leads us to conclude that three pieces of fruit are a total of three objects, and not four, should also lead to the view that, say, a grain of salt that's composed of 1.2×10^{18} atoms is actually 1.2×10^{18} objects, and not 1.2×10^{18} objects plus an additional object that is a grain of salt. But common sense doesn't endorse that view. So it is arbitrary. Rosen and Dorr point to the fact that this is because this common sense is *unreflective*, and I think they are right. Common sense may appear to commit to the existence of composite objects, but it does so without ever having

actually considered the plausibility of compositional nihilism or the logical problems associated with composition at all. Once presented explicitly, the common sense answer would likely reach no definitive conclusion on the question of composition and further have no real authority to do so.²

This obviously is not a knockdown argument for compositional nihilism, but in what follows, we will not need one.

A FICTIONALIST SOLUTION TO COMPOSITION

Given that we accept nihilism, do we have to say that our common-sense beliefs about things like tables are entirely without merit? Not necessarily. We can endorse the view that Rosen and Dorr call *fictionalism*.

Suppose there are simples arranged phone-wise in front of me and consider the sentence ‘There is a phone in front of me.’ Fictionalists claim that this sentence is strictly speaking false. But it’s not entirely without merit. We can say that it’s *virtually true*, because there are simples arranged phone-wise in front of me.

In the case of fictionalism, all statements about composite objects may be strictly speaking false, however, they are virtually true in that they can be verified by how we experience the world. The argument is basically that even if there are truly no composite objects, we can navigate the world with the useful fiction of composite objects. All that is required then to turn false statements true, is to specify that we are talking about the simples that are the material stuff of the objects we think we perceive. Therefore, a phone may not exist but simples may exist arranged phone-wise. To speak of the phone would be false, but to speak of the simples arranged phone-wise would be true, however inconvenient.

The result of this solution is a gradated truth. This kind of truth is presented here as three-tiered, but I would like to suggest that over time there may be a potential for even more tiers. Since I have not exhausted the consideration of all use-cases, I prefer to leave open the possibility by thinking of a truth spectrum. For the time being there are clearly three levels of truth. There are three

sentences below to demonstrate these levels.

- (1) There are simples.
- (2) There is a phone.
- (3) There are unicorns.

The truth claims of the sentences are:

- (1) strictly-speaking true and virtually true
- (2) strictly-speaking false and virtually true
- (3) strictly-speaking false and not virtually true

If one wanted to transform virtually true sentences to strictly true ones, one could replace “phone” with “simples arranged phone-wise.” This would yield a strictly-true sentence since it accounts for the fiction of composition (Rosen and Dorr 2002, p. 169). For the fictionalist this preserves the proposition, enabling speech to make sense of daily life. This would, according to them, satisfy all but the ambitious metaphysician who wants to know things simply for the sake of knowing (Rosen and Dorr 2002, p. 171). However, to claim that statements about composite objects are strictly-speaking false assumes that compositional nihilism is true—which will render this solution ineffective if Direct Reference Theory holds true.

THE INCOMPATIBILITY OF FICTIONALISM AND DIRECT REFERENCE THEORY

I will now argue that the fictionalist view is inconsistent with Direct Reference Theory (DRT). DRT, as developed by Saul Kripke (1972), is widely accepted as the standard view on how names refer to objects. According to DRT, names refer *directly* to their referents, without being mediated by meanings or senses.³ So on Kripke’s view, the only semantic value of a name is its referent. If the object that a name is supposed to refer to does not actually exist, then the name has no semantic value at all. It is meaningless.

Keeping in mind how reference operates in DRT, I return to an example of the fictionalist solution to the composition problem. In referring to a person, one may say, “Robert is over there.” In the fictionalist view, this sentence may be strictly-speaking false but virtually true. In order to render the statement strictly-speaking true, we may simply rephrase the sentence, “Simples arranged Robert-wise are over there.” Still, in the rephrased “strictly-true” sentence, the name ‘Robert’ appears. But if nihilism is true, then ‘Robert’ has no reference. And if DRT is also true, then ‘Robert’ has no semantic value or meaning. So fictionalists can’t say that the above sentence is strictly true, because they can’t even claim that it’s meaningful. The argument can be summed up as follows:

1. If compositional nihilism is true, then Robert does not exist.
2. If DRT is also true, then the name “Robert” has no referent.

Therefore,

3. If nihilism and DRT are both true, then the phrase “simples arranged Robert-wise” has no meaning.

Therefore,

4. Nihilists can’t save common sense by endorsing fictionalism.

In this case, fictionalism is problematic because it relies on the reference to a composite object, which in turn relies on the existence of the referent. Therefore, we arrive at the disjunction where we must give up either DRT or fictionalism. If compositional nihilism is true, then we cannot name anything except objects that are perfectly simple.

CONCLUSION

The problem of composition remains unsolved in my opinion. I further believe it is a genuine problem, rather than a pseudo problem as Rosen and Dorr suggest. I agree that common sense and our general belief in objects as we perceive them are indeed

unreflective. However, when faced with the explicit question of composition, the dilemma presented by a lack of an acceptable answer points to a legitimate inadequacy in our understanding of the nature of things. Without knowing what necessary and jointly sufficient conditions must be satisfied for a single object to be composed of proper parts, we cannot determine what objects exist independent of our anthropocentric bias. Both DRT and fictionalism are not only plausible, but credible and compelling theories. The fact that they are incompatible, then, points to a significant gap in our understanding—one that serves as an indication of where we may continue to pursue answers.

Notes

1. See Peter Van Inwagen (1990) and Ned Markosian (1998) for their respective ‘Special Composition Questions.’ They vary slightly, but not in a significant way according to Markosian.
2. For this explanation in greater detail, see Rosen and Dorr 2002, p.158.
3. Kripke’s view differs from Frege. While for Frege names have semantic value beyond their referent, specifically they have a meaning and a sense, for Kripke a name’s only semantic value is that of directly referring to the referent. That is, for Kripke a name only has semantic value if it specifically designates the person or thing intended. See Gottlob Frege (1892) “On sense and nominatum,” for further reading.

Bibliography

- Balaguer, Mark. (2011) “Is There A Fact of the Matter Between Direct Reference Theory and (Neo-)Fregeanism?,” *Philosophical Studies* 154, pp. 53-78
- Kripke, Saul. (1972) *Naming and Necessity* (Cambridge: Harvard University Press)
- Lewis, David. (1991) *Parts of Classes* (Oxford: Basil Blackwell)
- Markosian, Ned. (1998) “Brutal Composition,” *Philosophical Studies* 92, pp. 211-248
- Rosen, Gideon and Cian Dorr. (2002) “Composition as Fiction,” in: R.M. Gale (Ed), *The Blackwell Guide to Metaphysics*, pp. 151-174 (Oxford: Basil Blackwell)
- Van Inwagen, Peter. (1990) *Material Beings* (Ithaca: Cornell University Press)

A DEFENSE OF BARE PARTICULARS

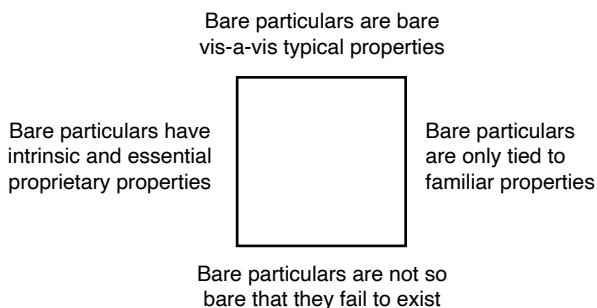
Ric Saenz

INTRODUCTION

How are ordinary objects composed? Some philosophers say that they're just bundles of properties. Others say that, besides their properties, objects have a bare particular—an underlying substratum—that is the bearer of properties.¹ I agree with the latter group. I say that familiar objects like apples, dogs, and tables can be analyzed in terms of properties and its bare particular. Moreover, I believe that bare particulars are that which accounts for the individuation of objects.² Thus, I'm committed to the following claim:

BP: bare particulars (1) are the bearer of properties and
(2) they account for individuation.

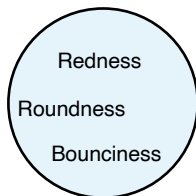
This claim, however, is not uncontroversial. In what follows, I will defend it against six objections. These objections are: (a) Loux's worry that BP is inappropriate since it just explains the phenomenon of individuation, (b) the Classic objection to BP, (c) the New Objection to BP, (d) the worry that if bare particulars are absolutely bare, then they slip into non-existence, (e) Loux's "Individuating the Individuator" objection, and finally (f) Mertz's "Only one Bare Particular" objection. What emerges is a version of bare particularism, **BP***, that adheres to the following rubric:



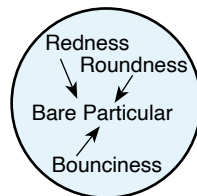
IS BP INAPPROPRIATE?

Suppose we have two qualitatively identical balls, B_1 and B_2 , sitting on a table. B_1 and B_2 have every property in common: both are Red, Round, Bouncy, Physical, etc.... But if B_1 and B_2 have every property in common, and if objects just are bundles of properties, then it follows that B_1 and B_2 are identical. In other words, we should have one ball rather than two. But the fact is that we do have two balls. So, then it must be false that B_1 and B_2 have every property in common or it must be false that objects are just bundles of properties. Now if we are realists about properties like Red and Roundness, then surely it is possible for two objects to exemplify the same exact properties; hence it is possible that B_1 and B_2 have every property in common and so it follows that it is false that objects are just bundles of properties. Many philosophers consider the possibility of qualitatively identical yet numerically distinct objects a fatal blow to bundle theories.

But if objects are not just bundles of properties, then what else might they include? Bare particularists argue that in addition to the properties of an object there is also a bare particular, an underlying substratum which is both the bearer of properties as well the thing which accounts for the individuation of qualitatively identical objects. For example, B_1 is composed of Redness, Roundness, Bounciness, as well as a bare particular, BP_1 . And so while B_2 has Redness, Roundness, and Bounciness just like B_1 , it does not have BP_1 —instead it has its own unique bare particular, viz. BP_2 . And since B_1 and B_2 have constituents beyond their universal properties, viz. BP_1 and BP_2 , we thus have the ontological resources to account for the individuation of objects that share all the same properties. Diagrams 1 and 2 show the difference between bundle theories and BP:



D1 Bundle Theory



D2 Bare Particularism

So the fact that we can have numerically distinct objects that are qualitatively identical has led some philosophers to posit a theoretic entity—the bare particular—to explain this phenomenon. However, some complain that it is somehow improper to posit an entity just to explain some phenomena. Michael Loux is representative:

Suppose that there is some phenomenon *x*; now, the occurrence of *x*'s is an unquestioned fact. Unfortunately, it is also an unexplained fact. Suppose that some philosopher comes along; and after reflecting on *x*'s, declares, "There are *y*'s." When asked what *y*'s are, he tells us, "They are the things that responsible for the occurrence of *x*'s." Pressed to amplify this cryptic characterization, our philosopher tells us that *y*'s exemplify all the transcendental properties and all those properties entailed by their being *y*'s. Pressed further, however, he denies that a more complete characterization of *y*'s is possible. "*y*'s are the things which are responsible for the occurrence of *x*'s, and that is the end of the matter!" he insists. *Now, we are likely to have doubts about the explanatory value of his account because it fails to provide us with any characterization of the entities it postulates which makes it clear how it is that they play the explanatory role the account attributes to them.* (Loux 1978, pp. 149-150, emphasis mine)

Loux takes bare particularism to be in the same boat. There is a phenomena—the possibility of two qualitatively identical objects—and then we posit some theoretical entity to explain the phenomena—the bare particular. But we aren't told anything more about our theoretical entity; we're just told that they account for individuation. And this is supposed to count against bare particularism.

I say that Loux's objection carries no weight, for there is nothing inappropriate about postulating a purely theoretical entity *in philosophy*—this seems to be part of the territory.³ In the sciences, we definitely would not be content with a mere theo-

retical entity, for we would expect empirical data to tell us more about our theoretical entity. In philosophy, however, it is standard fare to postulate a theoretic entity that only explains the relevant phenomena. For example, bundle theorists appeal to a special power—“compresence”—to explain why certain properties tend to hang together, to explain the phenomena of libertarian free-will, philosophers posit agent causation which is a unique power that allows one to break free of the sequence of purely physical causal connections, and scientific realists posit electrons to explain field distortions and observable effects on nearby objects. Thus, in each case, a theoretic entity is put forth and nothing more is said about it. And so long as we are comfortable with arm chair philosophy, then we should be able to tolerate purely theoretic entities.⁴

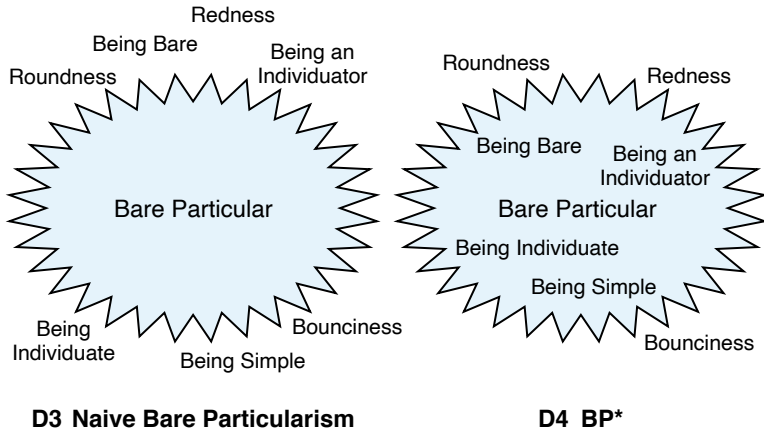
THE CLASSIC OBJECTION TO BP

Let’s move on to the second objection to BP.⁵ The Classic Objection entails that BP is contradictory, for BP supposedly asserts that bare particulars are both bare and not bare.⁶ So, *ex hypothesi*, the bare particular of a ball exemplifies all the “ball” properties: it exemplifies Red, Round, Bouncy, etc.... But on the other hand the bare particular is *bare*; i.e., as the thing that stands “under” the properties of the ball, the bare particular *lacks* Red, Round, and Bouncy. So bare particulars allegedly have and don’t have properties, they are both bare and not bare.

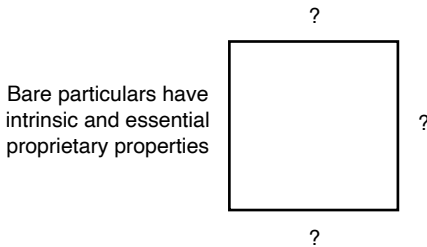
But according to my view, *BP**, bare particulars are not *absolutely* bare; to the contrary, they have a great many intrinsic and essential properties, viz., proprietary properties like Being Individuate, Being the Individuator, Being the Bearer of Properties, Being Simple, etc, so it is simply not true that bare particulars are absolutely bare, and, hence, the Classic Objection is not a problem for *BP**.

It will be useful to compare my view to what I call *naive bare particularism*. According to this view, bare particulars lack all intrinsic and essential properties. They are absolutely bare. Any relation that bare particulars have to properties like Redness and

Being an Individuator is purely *external*. But as we saw above, this leads to the Classic Objection, and so this makes it untenable. For this reason I submit that bare particulars are not absolutely bare; I say that they do have intrinsic and essential properties. The following diagrams show how my view (BP*) is distinct from naive bare particularism:



This, then, is the first side of the rubric mentioned above: bare particulars have intrinsic and essential proprietary properties.



THE NEW OBJECTION TO BP

Andrew Bailey (2012) has raised the following objection to BP which he calls the New Objection:

- (1) Bare particulars either do or do not have the properties of their host substance.

- (2) If they don't, then we've eliminated the central intuition of bare particularism.
- (3) If they do, then we have an overcrowding problem.
- (4) Either way, bare particularism is in trouble.

(1) is straightforward. (2) asserts that if bare particulars lack the properties associated with its host, then we've just abandoned bare particularism. For example, suppose BP1 is the bare particular of a ball, B1. But if BP1 doesn't have the properties of the ball then in what sense is BP1 the bearer of properties? Thus, we've abandoned bare particularism. (3) asserts that if bare particulars do have the properties of their host, then we have an overcrowding problem. For example, if we have a red ball which has a bare particular that is also red, then this one object has two instances of red—both the ball and the bare particular are red. And so Bailey concludes that (4) bare particularism is in trouble for either we abandon the central intuition of bare particularism or we face the overcrowding problem.

But I say that bare particulars lack *typical properties* like Redness and Roundness, and if this is the case, then premise (3) is false. Suppose we have a ball, B1, that is red. According to my view, B1 has a bare particular, BP1, which is not in itself red. In fact, BP1 lacks all typical properties, i.e., it lacks all the properties that we normally associate with the host substance, e.g. Redness, Roundness, Bounciness, Hardness, Being Concrete, etc, And so if BP1 is bare vis-a-vis typical properties, then there is no overcrowding problem, i.e., it is false that there are two instances of, say, Redness. And so Bailey's New Objection breaks down. Note also that I am only claiming that bare particulars lack *typical* properties, for, as was argued in the previous section, they do have a great many *atypical* properties like Being an Individuator and Being the Bearer of Properties. This then is the second side of the rubric:

Bare Particulars are bare
vis-a-vis typical properties

Bare particulars have
intrinsic and essential
proprietary properties



?

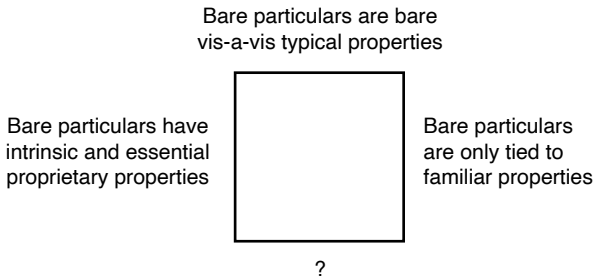
But if we deny (3), doesn't BP* fall prey to (2)? If bare particulars lack typical properties like Red and Roundness, then haven't we just abandoned the central intuition of BP? That is, bare particulars are supposed to be the ultimate bearers of typical properties like Redness and Roundness, but now we appear to be saying that bare particulars don't bear these properties. So haven't we abandoned the initial motivation for BP?

Not exactly. J.P. Moreland has introduced the following distinction: he says that a property can be "rooted in" in a subject or it can be "tied to" a subject (1997, p. 257). Regarding the former, Moreland says that a dog has the property Brownness and Being Mammal, and these properties are *caused* by the dog, i.e., the dog has the capacity to exemplify the properties Brownness and Being Mammal—there is something intrinsic to dogs that *predisposes* them to manifest Brownness and Being Mammal. Hence, *if a property, P, is caused by an object, x, then P is rooted in x.*

Bare particulars, on the other hand, are not the root of properties like Brownness or Being Mammal. Instead, they are only *tied to* such properties. The idea is that the relationship between bare particulars and typical properties is that they are part of the same set: x and P are not causally related, but they are both members of the same set. Analogously, all the even numbers are related to each by being part of the set of even numbers, but they are causally isolated from one another. And so there might be a set which includes a bare particular, BP_1 , and a property, P , in which case the two are "tied." This is what it means for a bare particular to be "tied to" a property.

The upshot, then, is that there is a sense in which bare partic-

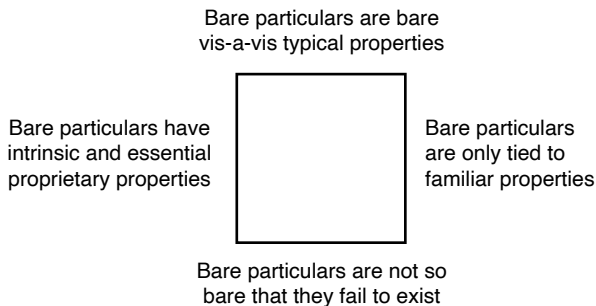
ulars do bear familiar properties: properties are not “rooted in” bare particulars but are “tied to” them, which is to say that the bare particular and property are causally isolated from each other and that they form a set. And so we preserve the central intuition of bare particularism. This then is the third side of the rubric: bare particulars are only “tied to” familiar properties.



BARE PARTICULARS AND NON-EXISTENCE

Bare particulars are often characterized as being absolutely bare. But, in addition to the problems mentioned above, there is the worry that if bare particulars are absolutely bare, then they fail to exist, for—quite plausibly—to exist just is to bear some property.

However, as I’ve already said, bare particulars do have some intrinsic and essential properties, albeit they aren’t the normal everyday properties that we are acquainted with. They have properties like Being the Bearer of Properties and Being Individuate and, in light of this fact, they are safe guarded from slipping into non-existence. So I’m committed to the claim that bare particulars do have properties, and this completes the four-sided rubric:



But by conceding that bare particulars have intrinsic and essential properties we are moving through dangerous waters (which is why some philosophers have been eager to keep them *absolutely bare*). There are two serious problems that the bare particularist must address.

First, if bare particulars have a nature or a set of essential and intrinsic properties, then we've reintroduced the problem that bare particulars were meant to solve. Remember, if two bundles or properties have every constituent in common, then it's hard to see why there is two rather than one bundle of properties. And so philosophers were led to posit an entity—the bare particular—to explain the numerical diversity of objects that are qualitatively identical. But now bare particulars appear to be in the same boat as bundles of properties for if bare particulars are qualitatively identical, then what accounts for *their* numeric diversity? Opponents of BP* will claim that an infinite regress is looming, for the bare particularist must posit more and more bare particulars to individuate the previous bare particular; e.g., BP₁ individuates a bundle of properties, B; BP₁* individuates BP₁; BP₁** individuates BP₁*; etc. ad infinitum.

The second problem is that if bare particulars have a common nature, then we have a serious identity problem; viz., if bare particulars are qualitatively identical, then—says my opponent—there is in fact only one bare particular in existence. Below I will offer rejoinders to each objection.

INDIVIDUATING THE INDIVIDUATOR

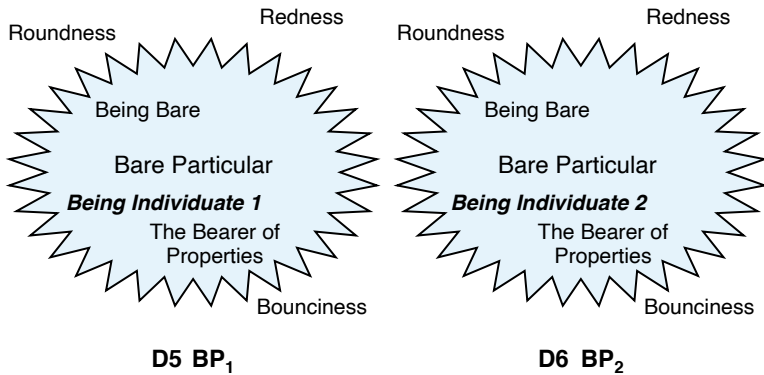
Loux argues that if bare particulars have an identical set of essential properties—e.g., all have the property Being the Bearer of Properties and Being that which Individuates—then bare particulars are qualitatively identical, and, hence, are in need of something to individuate them (Loux 1978). So suppose we have two bare particulars that are qualitatively identical. Both have the property Being the Bearer of Properties and Being that which Individuates, etc. but if they are identical vis-a-vis all of their essential proper-

ties, then bare particulars—the thing that was supposed to individuate macro objects like balls and trees—are in need of something to individuate them. And so a regress is looming.

The friend of BP* has at least two available rejoinders. First, he can concede that there is a problem in saying that bare particulars have essential properties. He can admit that if bare particulars have essential properties, then we will need an individuator for our individuator. But even if we can't explain the structure of micro objects like bare particulars (for the reasons Loux gives), at least we've put forth a satisfying explanation of everyday macro objects like apples and trees. Remember, no philosopher has to explain everything.

The second response available to the friend of BP* is that bare particulars—and *only bare particulars*—come packaged with the property Being Individuate, which is the property that individuates qualitatively identical bare particulars. So the idea is that there is a property, Being Individuate, whose sole function is to generate numeric diversity; and, if bare particulars have this property, then we have the means by which qualitatively identical bare particulars can be numerically distinct.

It is important to mention that this move is not ad hoc but is in the spirit of bare particularism, for bare particulars were introduced to—among other things—individuate qualitatively identical objects. And so it is no surprise that bare particulars have properties like Being Individuate that are responsible for individuating qualitatively identical bare particulars. Thus, if bare particulars have a unique and individuating property, Being Individuate, then Loux's objection doesn't threaten BP*, since it is false that all bare particulars are composed entirely of the same set of essential properties. Notice in the diagrams below that bare particulars have most properties in common, and that each has a unique, non-repeatable property, viz., Being Individuate 1 and Being Individuate 2.



Now my opponent will not go down without a fight. He will argue that if we can just stipulate that bare particulars have the property Being Individuate, then we ought to predicate this property of either bundles of properties or Aristotelian substances, which, in either case, avoids the bloated ontology of BP*.

But I say this move is unavailable to my opponent because I don't see how a bundle of properties or an Aristotelian substance can exemplify *any* property. Consider the following argument:

- (P1) Properties are predicated of either (a) bare particulars, (b) bundles of properties, or (c) Aristotelian substances.
- (P2) Properties can't be predicated of (b) bundles of properties or (c) Aristotelian substances.
- (C) Thus, properties are predicated of (a) bare particulars.

(P1) asserts that there are only three plausible candidates for how properties are instantiated. (P2) asserts that bundles of properties or Aristotelian substances are unsuitable candidates for property exemplification. And so it follows that (C) properties are predicated of bare particulars.

Let me say a little bit about (P1) and (P2). Beginning with (P1), it seems that these theories exhaust the alternatives, and so probably only one of these is true. Of course, someone might look at these options and conclude that that the whole enterprise

is suspect.⁷ Nevertheless, we will proceed as if these are the only theories in town.

(P2) asserts that bundles of properties and Aristotelian substances are ill-equipped for the task of predication. Now, the primary reason for thinking that we can't predicate Being Individuate of bundles of properties or Aristotelian substance is because *the predicating abilities of each are highly suspect*. So, even if one of these theories identifies that right constituents for objects, neither a bundle of properties or an Aristotelian substance seem able to exemplify a property. Put differently, bundles of properties or Aristotelian substance are not the types of things capable of predication; and, so it might be a simple category error to claim that bundle of properties or Aristotelian substance bears a property. Now this may seem like a mere rejection of either theory. However, it is not. What it is, instead, is an indictment on the ability of both bundles of properties and Aristotelian substances to predicate properties.

Allow me to develop this line of reasoning. Consider, first, a bundle of properties. There are four reasons to think that bundle of properties cannot exemplify properties, which entails that they cannot exemplify the property Being Individuate. First, it seems completely ad hoc to claim that a bundle of properties just so happens to have the property Being Individuate; it would be surprising if a bundle of properties that was composed of, say, Red, Round, and Bouncy, *coincidentally* possessed the property Being Individuate. Of the millions and millions of properties that exist, it would be surprising if a bundle just so happened to have this property.

Second, how can a property or bundle of properties bear another property? How, for example, can Redness bear Roundness? How can Redness be the thing in which Roundness subsists? In general, how can one property subsist in another property? The intuition is that properties, no matter how we conceive them, can't bear properties.⁸ For this reason I say that a bundle of properties cannot bear the property Being Individuate.

Third, positing Being Individuate to a bundle of properties

completely ignores the intuition that properties must belong to a subject. I assume that bundle theories fail to account for the *prima facie* fact that properties subsist in a subject. Instead, the theory entails that properties merely hang together by means of compresence. Of course, we can't hang all our philosophical beliefs on common sense intuitions but, all things being equal, it is better if our theories resonate with our experiences rather than not. And so I think that bundle theories go wrong in ignoring the intuition that properties subsist in subjects, and this ought to count against it.

Fourth, and most importantly, Being Individuate cannot belong to a bundle of properties unless that bundle is *antecedently* individuated.⁹ For example, suppose we have a bundle of properties, B_1 , which is composed of Red, Round, and Bouncy. Now ex hypothesis, this bundle is not numerically distinct from another bundle, B_2 , which is composed of Red, Round, and Bouncy, since both have the same exact properties. What is needed is the special property Being Individuate. But this property cannot be attributed to either B_1 or B_2 because there is no B_1 and there is no B_2 , i.e., B_1 and B_2 have not been individuated from some other bundle, B_3 (and of course B_3 hasn't been individuated from B_4 , etc...). Thus, Being Individuate cannot be exemplified or predicated of any bundle because that bundle would have to be antecedently individuated; but then the bundle theorist would be forced to beg the question by asserting that a bundle, B_1 , is individuated and is thus "ready to receive" the property Being Individuate.

Now let's turn our attention to Aristotelian substances. Roughly speaking, Aristotelian substances are normal, everyday objects like dogs and trees. These are said to be the ontologically basic entities in which properties subsist. So, according to this view, the property Brownness and Having Hair subsist in an individual dog—say Fido. But what is a Fido over and above its parts? We know of its properties and parts like Brownness, Hairiness, Being a Mammal, Having Teeth, having such and such weight and height, etc., but what is "Fido" over and above all of this, and how could this thing—if there is a thing at all—be the bearer of properties?¹⁰ One might say that Aristotelian substances are

instances of a kind. For example, Fido is an instance of Dog, and it is Fido which bears all the relevant properties. But, again, it is hard to see how an instance of a kind, whatever that is, can bear properties. In fact, what reason is there to think that an instance is something other than a bundle of properties? Thus, I say that Aristotelian substances like a Dog or Tree are not capable of predicating properties, since it is highly questionable that they even exist; or, if they do exist, then it is hard to see how they are really any different from a bundle of properties, in which case the objections given above will be applicable. So, I say that Aristotelian substances can't be the bearer of the property Being Individuate.

So, then, if the bundle theory and the Aristotelian substance theory are unable to account for exemplification in general, and the exemplification of Being Individuate in particular, and if these are the only viable alternatives to bare particularism, then it follows that (C) properties are predicated of bare particulars.

To summarize, then, Loux objects to BP* by arguing that since bare particulars are qualitatively identical, then they will need something to individuate them—we will need an individuator for the individuator. But I said that bare particulars have the property Being Individuate, which is a built-in property that just individuates one bare particular from another. My opponent will say that we ought to predicate this property of either a bundle of properties or an Aristotelian substance. But given the arguments presented above, I conclude that these are not adequate candidates for the predication of properties in general and Being Individuate in specific, and I conclude that properties, instead, are best thought of as being predicated of bare particulars. The upshot is that if there is a property Being Individuate at all, it will necessarily subsist in a bare particular.

ONLY ONE BARE PARTICULAR

Now we will turn our attention to the second objection to the claim that bare particulars have essential properties like Being the Bearer of Properties and Being Individuate. According to

D.W. Mertz, “All bare particulars in having no constituents have exactly the same constituents and so are identical” (2001, p. 52). Put differently, if all bare particulars have the same constituents, i.e., they all have the constituent “*having no constituents*,” and given the principle of constituents (entities having exactly the same constituents are identical, i.e.,

$(x)(y)[(z)(z \text{ is a constituent of } x \equiv z \text{ is a constituent of } y) \supset x = y]$), then there is—absurdly—only one bare particular after all (ibid). And this allegedly suggests that BP* is intrinsically flawed.

Is Mertz right? Does the fact that all bare particulars have the same nature or the fact that they all are bare, along with the principle of constituents, entail that there is only one bare particular? I highly doubt it. I offer four remarks to support my doubts.

First, it should be noted that Plato—the philosopher who formally introduced us to the world of abstract universals—did not affirm just any universal or property; he denied that there were platonic universals, or “forms”, such as Hair, Mud, or Dirt (Plato 1996, p. 130). I agree. I doubt that there is just any sort of universal, abstract object, or actual property—including of course the property Having no Constituents. And if this property doesn’t exist, then Mertz’s objection does not go through.

Second, I highly doubt that negative properties like Having no Constituents exist. Here, I follow Moreland and Pickavance in analyzing negative properties in terms of positive properties. They say:

Clearly, there are innumerable linguistic expressions assertible of bare particular, e.g., “is simple” and “is colored if green”. In our view, each of these linguistic predicates is analysable in such a way as not to require corresponding ontological properties. (2003, p. 9)

The idea is that we do make true statements about objects involving negative properties, but such statements do not entail a corresponding universal or property. So, for example, the proposition “The red ball lacks green” need not entail that there is an abstract universal *Lacks Green*. Instead, by the principle of parsimony

mony, we can accommodate such sentences by analyzing them as derivative of normal, everyday, *positive* properties. Thus, Lacks Green is merely derived from Green, and Lacks Green is not itself an object. In a slogan, if x lacks the property P, then x doesn't have a negative property. Instead, x *lacks* a positive property. Thus, Having no Constituents is analyzable into talk of, say, the property Being Simple. And so negative properties are unavailable to Mertz. But even if negative properties don't exist, we might try running Mertz's objection in terms of positive properties. So I will not hang my theory on the non-existence of negative properties.

Third, Mertz's objection is contradictory and so it can't be used against BP*. Here's why it's contradictory: bare particulars are said to have the constituent Having no Constituents, which is to say that bare particulars are P and \sim P. Analogously, x can not have the color Colorless or the weight Weightless, for in both cases we affirm and deny the same property. As a result, Mertz's objection is rendered impotent.

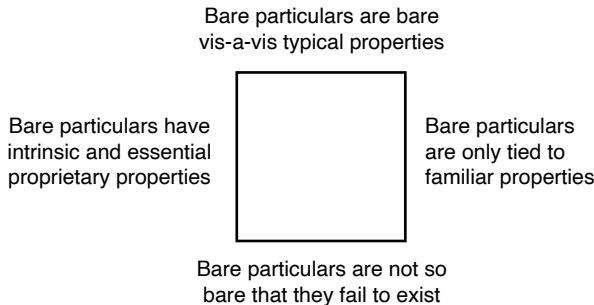
Fourth, and finally, there is still a more powerful reason to be suspicious of Mertz's objection: I say that his objection begs the question by presuming the bundle theory vis a vis bare particulars.¹¹ He presumes that bare particulars are constituted entirely of properties like Being the Bearer of Properties and Being Simple. But the friend of BP* denies this. He stipulates that bare particulars have, besides their repeatable properties like Being Simple, the *non-repeatable* property Being Individuate 1, Being Individuate 2, etc... (See diagrams 5 and 6 above). And even if objects just are bundles of properties, it doesn't follow that bundles having the same properties are identical. The reason they're not identical is because there are different theories vis-a-vis properties, some of which have the resources to account for numeric diversity of qualitatively identical objects. For example, if we assume platonic or immanent universalism, then it is true that objects with the all same properties are identical. But if we think of properties as *instances of transcendent universals*, then if a bare particular has a property, that property is an instance which is numerically distinct from another instance of that property. For

example, if BP_1 has P , P is numerically distinct from the instance of P that BP_2 possesses. Nevertheless, both BP_1 and BP_2 possesses instances of the same transcendent universal. So it simply begs the question to claim that bare particulars are identical given that they have the same properties.

Allow me to summarize the dialectic. In order to keep bare particulars from slipping into non-existence I claim that they do have have properties, e.g. Being the Bearer of Properties, Being Simple, Being that which Individuates, and Being Individuate. But Loux and Mertz argue that if bare particulars all have the same properties, then this entails certain problems; viz., the “Individuating the Individuator” objection and the “Only One Bare Particular” objection. But I argued that both of these objections are not a problem for bare particularism, and so I say that there is no problem with the fact that bare particulars have essential properties and are qualitatively identical.

CONCLUSION

So I think that bare particularism is going in the right direction vis-à-vis the structure of ordinary objects. Underlying the properties of an objects is a bare particular which is the bearer of properties and also that which is responsible for individuation. However, there are many objections to this claim. This paper concentrated on six objections, and my rejoinders entail a version of bare particularism, BP^* , that is composed of the following four propositions:



Notes

1. Philosophers who analyze objects into more basic constituents are sometimes called “constituent ontologists” (following Nicholas Wolterstorff, “Bergmann’s Constituent Ontology,” *Nous*, 1970). However, some deny that such analysis is possible
2. See Loux’s *Metaphysics: a contemporary introduction* for a fantastic overview to the bundle theory-substratum theory dialectic.
3. This response was inspired by a conversation with Professor David Pitt.
4. It should also be noted that BP* clashes directly with physicalism. The physicalist will argue that it is simply “spooky” to claim that objects have some non-physical entity beyond the properties that we are directly aware of. Where exactly is this non-physical entity? How can we, physical beings, know of a non-physical entity? As a result, many cast a suspicious eye towards BP*. Nevertheless, I will not assume any physicalist constraints.
5. Bailey coined the “Classic Objection” label.
6. Bailey provides the following bibliography for the Classic Objection (Bailey, 2012, p 33): Most begin discussion of bare particularism and the Classic Objection by referencing Plato’s receptacles, Aristotle’s ‘prime matter’, and Locke’s substrata ‘I know not what’. The usual citations: *Timaeus* 48c-53c, *Metaphysics* 1029a-33, *Essay II*, xxiii, Sect 2. The Classic Objection shows up in many discussions of bare particularism; some of these include Anscombe (1964, p.38), Armstrong (1989, p94-96), Campbell (1990, p, 9), Davis (2004, p 267-268), Hoffman and Rosenkrantz (1994, p 46-52), Loux (1978, p 149-152), Lowe (2003, p. 86), Mertz (2001, p 48), Sellars (1952, 1963, p 282), Quilter (1985), and Simons (1994, pp, 565-567).
7. Professor Mark Balaguer made this point in a conversation.
8. No matter how we conceive of universals, it’s difficult to see how a property can bear a property. So a transcendent universal can’t bear another transcendent universal, and neither can it bear an immanent universal or trope; an immanent universal can’t bear another immanent universal, and neither can it bear a transcendent universal or trope; and a trope can’t bear another trope, and neither can it bear a transcendent or immanent universal. And this is true of any other combination.
9. I am borrowing this argument from Loux who originally argued that bundles cannot bear impure properties like “Being identical with Oneself”. See Loux’s *Metaphysics: a contemporary introduction*, (chapter 3, p. 115).
10. I am of course assuming compositional nihilism, which is the view that there are no composite objects; there are only material simples. For an introduction to compositional nihilism see Ryan Wasserman’s, “Material Constitution” in *The Stanford Encyclopedia of Philosophy* (Wasserman, Ryan, “Material Constitution”, *The Stanford Encyclopedia of Philosophy* (Summer

2013 Edition), Edward N. Zalta (ed.), URL = <<http://plato.stanford.edu/archives/sum2013/entries/material-constitution/>>.), section 4, Elimitivism.

11. I thank David Pitt for alerting me to the fact that Mertz's argument assumes that objects just are bundles of properties.

Bibliography

- Bailey, Andrew. (2012) "No bare particulars", *Philosophical Studies*, Volume 158, Issue 1, pp. 31-41
- Loux's, Michael. (2002) *Metaphysics: a contemporary introduction* (Routledge, New York, New York)
- _____. (1978) *Substance and Attribute: a Study in Ontology* (D. Reidel Publishing Company, Dordrecht, Holland)
- Mary Louise, Gill and Paul Ryan. (1996) *Plato, Parmenides* (Hackett, Indianapolis, Indiana)
- Mertz, D.W. (2002) "Individuation and Instance Ontology", *Australasian Journal of Philosophy*, Vol. 79, No. 1, pp.45-61
- Moreland, J.P. (1998) "Theories of Individuation: a reconsideration of bare particulars", *Pacific Philosophical Quarterly* 79, pp 251-263
- Moreland, J.P. and Timothy Pickavance. (2003) "Bare Particulars and Individuation: Reply to Mertz", *Australasian Journal of Philosophy*, Vol. 81, No.1, pp. 1-13
- Wasserman, Ryan. (2013) "Material Constitution", *The Stanford Encyclopedia of Philosophy* (Summer Edition), Edward N. Zalta (ed.), URL = <<http://plato.stanford.edu/archives/sum2013/entries/material-constitution/>>

NIHILISM WITH A BERKELEYAN TWIST

Lorena Guandique

INTRODUCTION

Both nihilism and Berkeley's ontology stray from what people normally think of as common sense. I do not fully embrace nihilism, but Rosen and Dorr (2002) make a convincing case for it (or at least for not completely dismissing it) in their paper "Composition as a Fiction". After comparing other possible choices they come to the conclusion that none of the options stand out from the rest. All of the possible choices are equal and thus they see no way of actually making a choice between them all. They claim that things that we usually think of as wholes don't actually exist and when we talk about them, our utterances are false. The claim hardest to believe is that people do not exist; "on this view, it is probable that you do not exist... if the only objects in your vicinity are material objects—then strictly speaking, there is no such thing as you" (Rosen and Dorr 2002, p. 152).

Berkeley maintains (1982) that he is on the side of common sense, yet he makes similarly outrageous claims as Rosen and Dorr do. He claims, "[material objects'] *esse* is *percipi*, nor is it possible they should have any existence, out of the minds or thinking things which perceive them" (Berkeley 1982, p. 24). An object's essence just is someone perceiving it.

Even though when I first heard his position it sounded completely crazy, just as in the case with Rosen and Dorr, it was not so easily dismissed. In an effort to make these metaphysical theories easier to comprehend and accept for the non-philosopher, and thus get more people thinking about these fascinating metaphysical puzzles, I want to try to combine Rosen and Dorr's nihilism with some of Berkeley's notions about spirits and ideas, and I want to see if combining these two views helps to lessen

the strangeness the two philosophies pose without completely rejecting compositional nihilism. The upshot is that compositional nihilism is not so easily dismissed and a very likely candidate for a serious ontology, provided it is more in accordance with what common sense dictates.

BRIEF EXPLANATION OF NIHILISM AND BERKELEY'S ONTOLOGY

Compositional nihilism is the theory that composite objects, things with proper parts, do not exist. Nihilists only believe in the existence of the most basic particles, atoms. These atoms (or simples, as they are called in Rosen and Dorr's paper) are then arranged in the shape of the material objects we are used to dealing with, for example tables, chairs, cats, and plants. So when we talk about what we call a "coffee table" we're actually referring to a bunch of simples arranged coffee table-wise; but there is no actual single object that is the coffee table. This is weird. It's strange to think of a coffee table as just a bunch of atoms arranged in a certain way and not as a single object. What is stranger still is that they claim that people do not exist. We are just a collection of simples arranged person-wise. This is really paradoxical, because of all the things we humans tend to be pretty sure of, existing is definitely one of them. We've grown up thinking that all the things and people around us are real, that they do exist, that they are concrete, solid objects. Nihilists come along and tell us that all the material objects around us do not actually exist in the way we believe them to. They are not as they seem, single material objects, but tiny little particles arranged in such a way as to give us the illusion of the objects we experience. The theory is more than just a little bit hard to swallow. What nihilism suggests goes against common sense.

Another weird ontological theory is George Berkeley's. He states that there are no such things as material objects. According to Berkeley the existence of objects of material substance depend upon some mind or spirit perceiving them. These items do not

exist outside of them being perceived. The chair I am sitting in, the laptop I am typing this paper on, the book I keep referencing, none of these things exist without being perceived by me (or by someone else). This being the case it follows that when I leave the room where the chair, laptop, and book are that they would disappear because I am no longer there to perceive them. This is completely crazy though! We do not believe that just because we do not perceive some object that it ceases to exist. We actually believe the opposite. We believe that things exist independent of their being perceived. People don't wonder whether all their possessions go out of existence just because they leave their homes. We are fairly confident of the existence of material objects, whether they are being perceived or not. This is the strangeness in Berkeley's theory that an object's existence depends on our perceiving them.

COMBINING NIHILISM WITH BERKELEY'S IDEAS

In an effort to make some of the ideas from both philosophies a little more appealing, in terms of common sense, I want to put together some concepts from each philosophy to see if I can find a third, more plausible way to think of these things. First of all I want to get rid of the nihilistic idea that people do not exist. That is the idea that, I think, clashes the most with common sense, and the reason for which people would most likely reject nihilism. There is an endless amount of things people can doubt have existence and have intelligent discussions about; our existence, as people, is not one of them. More likely than not the average person on the street will be absolutely positive about one thing, and that is that they are really here, that they truly exist. They will be pretty sure that they exist as a person, made of flesh, bone, and with blood running through their veins. This is not something most people doubt. The nihilist just believes in simples, and not in composite material objects, but if we incorporate Berkeley's "spirits" into Rosen and Dorr's nihilism it makes the theory less difficult to accept.

For Berkeley, unlike for the nihilist, two types of things exist: spirits and ideas. In trying to stay as close to nihilism as possible I want to admit the existence of this extra thing, the spirits, but not as a composite material object. A spirit would be indivisible and outside of the material realm. So we still have only one simple material object existing, but we are not forced to say people do not exist. People do exist, and they are these spirit/mind entities, not simples arranged person-wise. Even though a spirit/mind is by no means equivalent to the flesh and blood person we are used to thinking about, it is still easier to imagine and accept because we are more familiar with the idea of having a soul; some people have even come across the mind/body problem, and are willing to accept that their mind is not the same thing as their brain.

The idea of thinking of people (or any sentient being) as spirits is consistent with Nihilism because it allows for their existence without committing to extra stuff existing outside of the simples the nihilists believe in. Maybe this isn't the same as the more common sense view but it's easier to accept this because of the concept of souls. The concept of souls is more commonly accepted, so not as weird as saying that what we think of as "people" are actually not people but just atoms arranged person-wise. The spirits I am thinking of would be like what we think of when we talk about a soul, but with a mind.

Some of Berkeley's beliefs about ideas are appealing in trying to make nihilism less outlandish. If we think of nihilistic simples as Berkeleyan ideas it takes a little bit of the strangeness away. Berkeley believes that, "the objects of human knowledge [...] are either ideas actually imprinted on the senses, or else such as are perceived by attending to the passions and operations of the mind, or lastly ideas formed by help of memory and imagination" (Berkeley 1982, p. 23). The first example refers to things we experience through the senses; the second example refers to ideas we get from reflecting. The third example talks about ideas we get when remembering, or ideas we simply make up using our imagination. I toyed with the idea of making the nihilist's simples simply be like Berkeley's notion of ideas, things that we receive in

our minds through the senses, reflections, or memories and imaginings. In this scenario things would be a little like the way we already understand things to be: for example, we have this notion that we can create and destroy (most of) the material objects around us. People build and destroy houses, carve marble statues then break them apart, and put together and take apart computers.

We feel we have some sort of manipulative power over the objects around us. We, to a certain extent, have some control over things of nature too: we can plant a tree, breed dogs with particular characteristics, destroy forests, and wipe out entire species. Our thoughts are similar to material objects in that we can manipulate the ideas in our minds just like we can manipulate concrete objects. I can play around with the ideas of a mare and griffin and come up with a hippogriff. Just as I can throw my laptop on the floor and wreck it, I can take apart the griffin and mare and destroy the hippogriff. Either way, I have some power over these thoughts of mine and can manipulate them in different ways, i.e., I can change the color of the hippogriff from red to grey, I can change its size from 11 inches to 25 feet tall, I can attribute special powers to it such as the ability to breathe fire, or manipulate lightning, etc. By likening the nihilist's simples with Berkeley's ideas we get a way of thinking about these things that seems less foreign.

There is a problem with likening simples to Berkeley's ideas. There's the fact that we tend to think that material objects can exist without our constant perception of them. This is Berkeley's odd insinuation: that things require our perceiving them to exist. He actually manages to wriggle himself out of this weird notion by saying that things do not cease to exist when we leave the room because God constantly perceives everything, so even though we stop perceiving something, He does not, and so it continues to exist. Attributing objects' constant existence to God seems to not be a strong enough argument. Saying things don't disappear when we leave the room because God perceives them at all times is fine if you believe in God, but if you don't, Berkeley doesn't have another logical explanation for the continued existence of things once they're not being perceived. Likewise saying that our ideas

(thus the simples) do not disappear once we stop perceiving them because God perceives them is too crazy of an idea to accept. In order to stay with common sense we want our simples to be more like regular concrete objects, i.e. we want our simples to exist independent of our (or anybody else's, omnipotent or otherwise) perception. We want our simples to be like our idea of objects and not Berkeley's idea of objects. Because we think we have some, but not total control over the objects around us, I do not want to equate nihilistic simples with Berkeleyan ideas. We feel we have more control over our ideas than we do over the material objects around us. Instead I want to think of simples as primary qualities.

Thinking of material objects as ideas was a little problematic because we have this notion that we can manipulate the things around us (just like we manipulate our ideas), but the objects existence, unlike our ideas, does not entirely depend on us. We can build houses, carve statues, and construct computers but they do not stop existing once we leave them alone. They continue to exist and do whatever it is they do independently of our attention. On the other hand, with ideas we can be thinking about them one second, but a split second later they are gone as soon as we turn our attention to something else. This is why whatever it is that simples are, they need to be independent of the people experiencing them, but somehow, in some way or another, be capable of being manipulated by us.

Thinking about simples as ideas was close to what I wanted but thinking of simples as primary qualities is more accurate. In nihilism simples exist independently of anything else. The simples are not created or destroyed by anything else. The simples are just there and other simples are in some sort of relationship with them. Let's take a simple sentence like 'someone is looking at the blue chair'. The nihilist would say that what is happening is 'there are simples arranged chair-wise and some simples arranged person-wise are having some experience of it'. When combining nihilism with Berkeley's idea of spirit we can have spirits perceiving some primary qualities (simples), and manipulating them like ideas by adding secondary qualities to them e.g. by attributing to some

group of simples color, size, texture we get the idea of a blue chair. The simples exist independent of us, just like primary qualities exist independent of us, and the spirit dresses up these simples (primary qualities) with secondary qualities that come from their own mind. We can apply bundle theory to this idea and think of objects as bundles of simples that are arranged in such a way that they have the potentiality to suggest a certain object to us.

We have all had an experience where we do not agree with some other person as to the specific shade of some object. Thinking of an object as a bundle of simples would explain why. The simples are really out there in the world, outside our mind, but the color is not in the thing itself; it comes from the person (spirit). The perceiver dresses up the simples, thus each perceiver might experience something a little bit different from others. Having a spirit entity experiencing some simples and then adding the details to it sounds closer to how we normally think of things, than just a bunch of atoms being close to each other and thereby getting the complex and varied ideas of chairs, dogs, people and so on.

A spirit experiencing the simples and then attributing the secondary qualities to them would be similar to someone playing an on-line video game where one can share one's character with other on-line players and interact with them. When playing on-line games where one has a character that one manipulates, one usually starts off with a basic figure of a human (or elf, or alien, or dwarf or whatever character choices the game offers) then adds different physical traits to make it look however one wants. The basic character would be like the simples, and adding blue hair, and face paint, and giving them destructive magical powers and elven armor would be like the secondary qualities. The players interact on-line through their characters. The players have choices to make for their characters as to what to do while on-line. There are consequences that happen depending on what choices are made; characters get killed off, or join rebellion groups, or sneak into other characters' palaces and then get sent to jail. This is similar to the way I'm thinking about this issue. There are the spirits (which would be the gamers), and they attribute certain secondary quali-

ties to primary qualities or, simples. The simples or primary qualities exist independent of the spirit. The secondary qualities come from the spirit and dress up the primary qualities. The spirits then interact through the simples clad in secondary qualities.

PROBLEMS

In thinking about things in this nihilistic/Berkeleyan way many problems arose. I want to put forward a couple problems now, and try to find some possible solutions to them. The first problem I came across is about secondary qualities; we obtain our experiences of secondary qualities through the senses, but how can we obtain experiences of colors, smells, tastes, textures, etc. through the senses if we (as spirits) are not material substance? If we are merely spirits, I assume, we do not have ears, eyes, and all the other proper equipment to be open to obtaining information through the senses. There must be some vehicle through which we receive the information if not through the senses. I figure we, as spirits have some way—maybe not the traditional way, but some way—of receiving information that allows us to have ideas of color, smells, etc., and thus to be able to apply them to the simples. The only possible response I can think of is that we do not receive any such information about secondary qualities. That information comes from us, we attribute it to things outside of us, we do not receive it, and since we do not receive it we do not need to be equipped with any receptor for this information.

The second problem I really have no conceivable answer to. If as spirits, we attribute secondary qualities to the simples, how is it that we come to ascribe such similar qualities to the simples arranged object-wise? So, for example, it seems to me that if we ourselves are assigning a color to some particular object, how is it that we all seem to agree on what the color is? How do we all see the Statue of Liberty as some shade of blue-green, and how is it that nobody sees it as mustard, or burnt sienna? It seems strange that having the power to assign whatever secondary qualities we wish upon the simples, we would all assign very similar, some-

times the exact same qualities to them. But, of course, there is no sure way to know that we are agreeing on colors and other secondary qualities in the first place. It could be the case that what I experience as turquoise and have learned to call turquoise someone else also calls turquoise but actually has an experience of orange. That being the case it's a little disconcerting to think that we are not all in agreement about what we experience.

CONCLUSION

Thinking about everyday objects in this nihilistic/Berkeleyan manner keeps nihilism's main idea, that there is only one type of material thing that exists (the primary qualities), but we also get a different way of thinking about ourselves without things being too different from what we are used to. We do not have to give up the idea of being a single sentient creature that acts, and interacts in the world.

Bibliography

- Berkeley, George. (1982) *A Treatise Concerning the Principles of Human Knowledge*. Kenneth Winkler (Ed). (Indiana: Hackett Publishing Company)
- Berkeley, George. (1979) *Three Dialogues between Hylas and Philonous*. Robert M. Adams (Ed). (Indiana: Hackett Publishing Company)
- Rosen, Gideon and Cian Dorr. (2002) "Composition as a Fiction," in: Richard M. Gale (Ed), *The Blackwell Guide to Metaphysics*, pp. 151-174 (Oxford: Blackwell)

THE WORLD IS NOT ENOUGH: HOW HYLEMORPHIC DUALISM OVERCOMES PROBLEMS WITH CARTESIAN SUBSTANCE DUALISM

Mark H. Morrow

That which is a whole and has a certain shape and form is one in a still higher degree; and especially if a thing is of this sort by nature, and not by force like the things which are unified by glue or nails or by being tied together, i.e., if it has in itself the cause of its continuity.

— Aristotle, *Metaphysics*

For the form is that through which a thing is the very thing that it is.

— Thomas Aquinas, *Questions on the Soul*

I. INTRODUCTION AND GENERAL CONSIDERATIONS

Substance dualism is a view in philosophy of mind that is often cast aside as an object of ridicule and misunderstanding. Most of its adherents are considered to be mavericks that run “against the grain”, mainly because of certain problems that arise for the Cartesian form of dualism. Thus, when almost any discussion of substance dualism is on the table, it is nearly always the Cartesian form that takes center stage, despite the growing number of property and event dualists.¹ Nevertheless, Cartesian dualism retains the unassailable pride of place as the whipping post on which most dualists are ritualistically flogged (Oderberg 2005, p.71). The typical “Queen Elizabeth” problems, well rehearsed in the literature, crop up continually against Cartesian interactionism.

How can an immaterial non-spatial substance enter into causal relations and exert causal powers on a spatiotemporal physical substance? Don't causes need to be spatially paired with their effects? What are the identity conditions for such a substance? It seems that we intuitively distinguish between mind and body in ordinary contexts as often as we do in philosophical ones, and in a way that implies that the difference between them goes deeper than a mere difference between part and whole. We do not, after all, make the same intuitive distinction between "hand and body" or even "body and brain" (Feser 2006, p.19). The philosopher that takes the inquiry I speak of here seriously will discover that while Descartes does leave behind some problems with his dualist solution to the mind-body problem (e.g. the causal pairing problem and causal interactionism), there is much that Descartes has in common with another kind of dualism for which I shall presently argue.² Property and event dualism, while occupying a respectable place, have also encountered great difficulty in providing an adequate solution to the mind-body problem. Property dualists face two unattractive alternatives, epiphenomenalism or overdetermination, neither of which seems very satisfying. The difficulty in securing an adequate solution to these problems is a reflection of the materialist's steadfast determination to either eliminate or reduce the mind. What Adam Pautz calls the "whole hard problem of naturalizing the mind" is not likely to be going away any time soon (Pautz 2013, p.195). One reason that adequate functionalist or property dualist solutions to this problem have not been forthcoming is that materialist approaches to solving the mind-body problem presuppose a compatibility with our best-known conceptions of the natural sciences. If materialist approaches to solving this problem of mind fail to ground themselves in our best-known scientific theories, then the explanatory power of naturalist epistemology is threatened, thus inhibiting the naturalist's ability to find a robust solution to the hard problem.³

Recently, however, there has been a "turning of the tide" which heralds a resurgence of interest and serious intellectual attention to a unique form of dualism (Oderberg 2005, p.71). I

shall argue that substance dualism of the Cartesian sort is not the only game in town and instead I shall present a form of substance dualism, found in Aristotle and Thomas Aquinas, that I shall call “hylemorphic dualism” (hereafter HD).⁴ This form of dualism circumvents many of the problems encountered by Cartesian dualism, and thus provides a more plausible and robust dualism that carries greater explanatory power. I shall argue for substance dualism *simpliciter*, and then for HD specifically, by presenting three arguments. First, I shall present a modal argument from conceivability that shows that disembodied existence is epistemically possible. Second, I shall argue that the dualism of Aristotle and the Aristotelians overcomes many of the worries that Cartesian dualism faces. In particular I shall address the problem of causal interactionism and causal pairing. Third, I shall argue that hylemorphic dualism is the best explanation for appeals to notions of proper functioning and in particular a conception of irreducible structure that is currently enjoying support from the biological sciences. Moreover, I shall argue that HD provides the best explanation for thinking that living organisms *qua* irreducible wholes are autonomous centers of biological development. Finally, I shall conclude that HD proves itself to be a reasonable view, able to overcome Cartesian worries and explain irreducible structure of living organisms, therefore demonstrating itself to be a satisfying metaphysical theory of mind.

II. THE MODAL ARGUMENT FOR DUALISM

What case can be made for any kind of substance dualism? Is the mind identical to the body or are they distinct substances? If distinct, what is the relationship of the mind or soul to the body? In order to make a case for HD, I shall begin by making a case for what is common to all views of substance dualism. Here I present and defend a modal argument for substance dualism from the conceivability of disembodied existence. While this argument is not original, as versions of it date back at least to Descartes himself, it nevertheless aims to show that disembodied existence

is possible. Moreover, it shows that disembodied existence is not only a feature of Cartesian dualism, but is a feature of the HD view as well. It does not follow that just because disembodied existence can be conceived, and is thus possible, that one is locked into a commitment to a Cartesian view that posits a soul, related by an external casual relation to a body. The distinct differences between Cartesian substance dualism and HD will be explicated in the sections to follow. The modal argument is grounded in conceivability. If something is conceivable then this is ample justification for the claim that it is a least possible. There are several objections that have been raised against this. Many have said that the very notion of conceiving is vague and can be used in too many different ways (Tidman 1994, pp. 297-298). However, two things can be said against this. First, conceiving is not imagining and second, conceiving is not the same as understanding. We can conceive of something without having an image of it in our minds and we can understand something, like the idea of a square circle, without conceiving of it. Here I shall use conceive as “what seems to be coherently supposed” (Moreland 2000, p. 172).⁵

This argument can be constructed in the following straightforward deduction, after the fashion of Plantinga (1978, pp. 65-69).

1. What is imaginable is logically possible.
2. I can imagine myself existing while my body does not.
3. Therefore, it is logically possible that I could exist while my body does not. (1, 2)
4. It is not logically possible that my body could exist while my body does not.
5. Therefore, I have at least one property that my body does not have. (3, 4)
6. For any entities A and B, if A and B are identical, then A and B have all of their properties in common. (Leibniz’s law of the indiscernibility of identicals)

7. Therefore, I am not the same thing as my body (i.e., they are not identical). (5, 6)

One objection to this argument is that if it can be conceived that I am possibly not my body, then it can equally be conceived that I am my body. However this response confuses conceiving of oneself to be identical with one's body with conceiving of oneself to be integrally embodied (i.e. to be interrelated to one's body). Another objection takes the following form. Your natural kind is essential to you. Given that being an animal with a body is part of your natural kind, then having such a body is essential to you, and thus you cannot exist without a body. It can be said that our natural kind is "human person" and it is essential that we belong to this kind. It is not essential, however, to being a human person that we are embodied. On the HD view, it is essential to being a human person that we have the capacities for developing a body. When a biologist studies the body, he studies the actualization of capacities for embodiment within the soul. It is necessary, then, in order to be a human person that these capacities are present, but it is not necessary for them to be actualized. So while I would happily concede that the modal argument is not infallible, conceivability does provide a good test for possibility.

III. CAUSAL INTERACTION, CAUSAL PAIRING, AND THE NATURE OF HD

Can an immaterial soul causally interact with a physical body? How would it even be possible for a soul to enter into some kind of pairing relation with a body? These questions, which appear to present significant objections against dualism, are not as problematic as one might think. In fact, the possibility of producing a theory that is plausible and answers these objections depends on one's prior assumptions and metaphysical commitments. Here I shall contend that the materialist objections to Cartesian dualism turn out to be minimal. I will explain why these "problems" are only apparent, and why causal interactionism amounts to not much more than materialist question begging against the Cartesian or

advocate of HD. The mystery of how something immaterial can exert causal forces on a material body is not a problem with the dualists' conception of mind, but rather the materialists' conception of the physical world. As I will explain in the following section, HD is not compatible with causal interactionism only in that it is not consistent with efficient or material causality. However, it is consistent with final and formal causality.

There is no mystery about how soul and body get into causal contact with one another, for the soul-body relationship is just one instance of a more general relationship existing everywhere in the natural world, namely, the relation between forms... and the matter they organize. If this general relationship is not particularly mysterious, neither is the specific case of the relationship between soul and body. The mistake of Cartesian dualists and materialists alike... is to think of all causation as efficient causation, [so] when it is allowed that there are other irreducible modes of explanation—in particular; explanation in terms of formal causation—the interaction problem disappears. (Feser 2006, p. 223)

In this section, I shall explicate what I mean by HD and what I think are two primary distinctions between the Cartesian and HD views. First, the interactionist objection presupposes that all causation is physical-to-physical causation. Now barring the materialist's commitment to the causal closure of the physical domain, the materialist is simply asserting what either form of dualism denies. Is it patently obvious what the correct theory of causation is? What justification can be provided to show that all causation involves physical contact between cause and effect and the transfer of some physical magnitude from cause to effect? Given that in physics there are a host of competing theories of causation (e.g. transfer, counterfactual, and nomological theories), it seems that the nature of causation is up for grabs. Since the interactionist objection presupposes a transfer theory of causation, it is unclear that any kinds of interactionist objections are decisive.

If there is any truth whatsoever to Cartesian dualism, then some kind of immaterial causation is possible (Vallicella 2009, p. 1). Simply asserting that all causation is physical begs the question from the outset. Moreover, even if it were the case that all causation is physical-to-physical causation, it would not follow that all physical-to-physical causation would be intermediate and indirect. When one asks how one thing can exert casual forces on another, this presupposes a causal intermediary (like asking how turning a key can start an engine). Not all causation, however, is indirect. It could turn out that some types of causation, even physical-to-physical causation, can be immediate and direct. One might ask what the intervening mechanism of mental-to-physical causation is. The answer is that this too begs the question against the dualist in that there may be no intervening mechanism at all.

One of the leading strengths of HD is that it overcomes typical objections to Cartesian dualism. What I shall show here is how the HD view presents a dualism that is free from most of the worries that come with Gilbert Ryle's "ghost in the machine". This worry is valid in the context of the Cartesian conception, which views the body as merely a machine. A difficulty then arises in proffering an explanation of what it is that makes the body human. For the Cartesian, the sole relationship between mind and body is an external causal relation, so while the soul or mind is a substance, the body-mind composite is a "property thing". This means that the body is composed of ordered aggregates and the mind is related externally to a body that is an ordered aggregate. Contrary to Descartes, the body on the HD view is not an ordered aggregate. Rather, it is human in virtue of its soul being diffused and fully present in every body part, as the essence of it. This presents a much richer ontology than that of a Cartesian mind, causally connected to a body that is solely physical. Moreover, there is a modal distinction to be made between soul and body. The soul on the HD view can exist without the body, but the body cannot exist without the soul. Thus we do not wind up with two distinct separable substances as depicted in the Cartesian conception. Rather there is only one substance, although it should not

be identified with the body-soul composite. On the HD view the soul is one substance, and the body is an ensouled biological and physical structure that depends on the soul for it to exist (Moreland 2000, p. 201).

One of the strongest objections to the Cartesian conception of substance dualism has been raised by John Foster, and more recently by Jaegwon Kim. Kim calls this the causal pairing problem. He asks how non-spatial, immaterial substances could possibly enter into causal relations with material substances and other non-spatial mental substances. What does it mean for a non-spatiotemporal immaterial mind to be united to a spatiotemporal body? According to Kim, in order to explain mental causation there would necessarily need to be a causal chain. Kim presupposes that since Cartesian minds are wholly outside of space, and causal chains necessarily require space in order to produce pairing relations, then causation by immaterial souls would be functionally impossible. If body and soul are united to form one person, then according to Kim “there must be a relationship R such that a mind stands in relation R to a body, if and only if that mind and that body constitute a unitary person” (Kim 2005, p. 78). According to Kim, we must understand what R means. If we cannot know what R is then a body-soul union is unintelligible.

Kim provides an illustration of this by describing two guns, A and B, being simultaneously fired and subsequently killing two persons. Gun A kills Adam and gun B kills Bob. Why is it that gun A kills Adam and not Bob and why is it that gun B kills Bob and not Adam? This same line of reasoning is employed against the idea of Cartesian souls as causal agents. Kim postulates two immaterial souls, A and B, which have the ability to exercise causal powers on a material thing M. If A and B exercise their causal powers but only A has an effect on M, supposing this is possible, then given the fact that A and B do not exist in space, what could possibly account for any causal pairing of A on M? Why does A have causal influence on M and not B? Kim’s argument against Cartesian dualism and immaterial souls hinges on two assumptions. The first assumption is that a causal chain must

necessarily be spatial in order to pair a given cause with its effect. Kim wants to know what relation could possibly pair a cause with its effect, *not why one is necessary*. The second assumption is that immaterial souls are necessarily non spatial. My response to these assumptions is threefold. First, Kim wrongly presupposes that anything non-spatial lacks causal powers. It seems to be possible to conceive of a world where no space exists, such as our situation prior to the big bang. According to big bang cosmology, the singularity was caused but not within space. It seems then that Kim would have to reject any understanding of the origins of the universe that required the singularity to have a cause. Second, the possibility of spatially extended souls is a viable dualist option.⁶ Kim only considers souls as extensionless geometric points. Why can't souls occupy regions of space and be extended, but be neither physical nor material? It seems plausible that a particular soul could be tied to a particular body in virtue of occupying the exact same region of space as that body. Since souls are spatially situated on this view, they can enter into spatial or pairing relations, and therefore have the possibility of entering into causal relations. This explanation provides a solution to Kim's pairing problem. Kim asks, what it is about soul A that could evoke causation on M and not B? The answer is that A can cause this change in M because A exactly occupies the same region of space as M. B occupies another region of space, and therefore cannot cause the change in M.

One might object by suggesting that if souls were to exist spatially, then they would not be capable of causing anything, in virtue of their being co-located. Then it could be asserted that co-located causes are unable to be differentiated spatially. However, this does not necessarily prevent two different things that are coincidental from each possessing their own distinct causal powers. A sugar cube could in theory possess distinct causal powers from its size and shape, even though the sugar cube is not spatially differentiated from its size and shape. Since this seems plausible, this would negate the claim that a cause necessarily needs to be spatially paired with its effect. Finally, it could

be objected that A cannot be determined as the source of causation, as opposed to B, if it's the case that A and B share all their non-haecceitous properties (properties that are neither unique nor essential to a thing). The problem here is that there are a number of conditions like this that could be required to have causation and not result in a causal pairing problem for any immaterial entity. Given that conditions like this may not be true, it seems to be possible that A could exert a causal influence on M instead of B even if there were no non-haecceitous property of A or of B, such as being spatially oriented in the right way, whose instantiation distinguishes A from B (Tooley 1997, pp. 89-92). If this is true, then a pairing relation doesn't seem to be required for causation. It seems at this point that there needs to be ample justification as to why there can't be singular causation between an immaterial soul outside space and a spatial material object that is governed by non-haecceitous laws. Kim asserts that a pairing relation is required, but doesn't explain why one is necessary. In any case it seems that this objection can be successfully refuted.

In order to begin to get a grasp on what HD is, it will first be important to understand one of the distinctions between modern thought and ancient or medieval thought. This is found in the notion of "substantial form". I contend that it is this notion of persons as substances as opposed to "property things" that is essential to this view.⁷ Moreover, a rediscovery of holistic cause, which is similar to the notion of substantial form, is already taking place in the biological sciences. Ironically, the very kind of science that was responsible for the rejection of substantial form in the early period of modern philosophy is now responsible in part for its recovery (Nichols 1996, p.1). This notion will be necessary to the understanding of the central claims of HD.⁸ First, all substances or self-subsisting entities that serve as bearers of properties but are not themselves properties of anything are compounds of matter (hyle) and form (morphe). Second, the form is substantial since it actualizes matter and gives the substance its very essence and identity. The human person, being a substance, is a compound of matter and substantial form. According to HD, the person is defined as

an individual substance having a rational nature. The substantial form is the rational nature. The operation of rationality is essentially an immaterial one. Therefore, human nature itself is essentially immaterial, and so the rational nature's existence does not depend on its being united to matter. It is a subsistent form. It will follow from this that a person is capable of existing by means of his rational nature independently of the existence of the body, and so we as human beings are immortal. I shall argue presently that although we are immortal, our identity and individuality require we be united to a body at some time in our existence. In this sense there can be duality of matter and structure as substance along with the substantial structure or rational nature, even though the rational nature has the capacity to exist independently. This is very differently from the Cartesian conception of the person, in that the rational nature is more intimately related to the body (Oderberg 2005, p.71). It will be important then to understand the distinctions between HD and Cartesian dualism.

While Descartes gets many things right in his ontology, he does make a mistake with respect to *the soul's relationship to the body*. It is on this point that I believe the contentious issues about substance dualism turn. Consequently, two primary distinctions can be made between Cartesian dualism and HD. The first is that, for Descartes, there is a sharp distinction between the body and the mind. Descartes thought of the immaterial ego as a mind more than a soul, and as a result we have a mind-body problem rather than a soul-body problem, which is captured on the Aristotelian-Thomistic conception. This reduction of the soul to the mind seems to identify the person with purely conscious substance, or an ultimate capacity for consciousness. On the HD conception, the mind is a faculty of the soul and possibly requires certain physical states in the brain and central nervous system to obtain before it can function. However, the soul itself does not require certain states of the brain to obtain before the soul itself is able to function. The soul, on both a Thomistic and Aristotelian account of HD, is responsible for the development of the brain and nervous system and, more generally, the body (Moreland 2000, p.200). On

the HD view, the soul is more than just a capacity for consciousness. It serves as an organizing principle of matter into structure. Thus, it is responsible for organic functioning and the living activities of the body. Contra Descartes, who argues that the soul is in the body, the HD view says that while the soul is in the body, the body is also in the soul.⁹ According to Aristotle all material things are composed of a potential principle, prime matter, and an actualizing principle known as substantial form or structure. Prime matter, as the potential for existence, has no actuality of its own and consequently no existence apart from structure. The HD view I advocate, however, allows for a subsistent form that, contra Aristotle, can have existence apart from structure. All characteristics of the person as matter and structure flow from the structure as the actualizing principle. “Immaterial substance” on the HD view is found in both matter and structure organized by life’s form, as opposed to Descartes who locates immaterial substance only in the soul as mind (Crem 1979, p. 160). Causal commerce between body and soul as efficient causality, which is seen in the Cartesian conception of the person, becomes a moot point on the HD view. Causation is formal causality related according to the structure imposed on matter or the human body by life’s subsistent form.

IV. INFERENCE TO THE BEST EXPLANATION

At the heart of HD is the notion of structure as it is imposed on matter by the rational nature as an organizing principle. Here I argue that structure is a real and irreducible ontological and explanatory principle. This is reasonable to believe based on the evidence of appeals to structure in biology and most notably neuroscience. William Jaworski describes HD in the following fashion:

Organisms are not just chunks of matter and energy; they are chunks of matter and energy that are structured or organized in various ways. That structure or organization is what is responsible for organisms having the distinctive capacities they have such as the capacities for growth and development, reproduction, perception, movement, and

cognition. It is because organisms have these capacities that they qualify as living things as opposed to nonliving ones, and it is because they possess distinctive types of developmental, reproductive, and other capacities that they qualify as living things of one or another kind: mammal, fish, bird, primate. The capacities that categorize us humans as the kind of living things we are include capacities for engaging in activities described and explained in a vocabulary of psychological predicates and terms: thought, feeling, intentional action, personality, character, and so on. (Jaworski 2011, pp. 296-298)

In this way biology, and even neuroscience, seems to provide evidence for the HD view rather than against it. HD, I contend, is also the best explanation for a view in the biological sciences currently on the rise, namely the “organocentric” view. This view is increasingly supported in the scientific literature as a worthy opponent to the “genocentric” view. According to Jaworski, HD explains why we can successfully explain the behavior of organisms by making appeal to their structure. All living things exhibit some sort of behavior that seems to be grounded in their structure. Many in the natural science community make use of free concepts such as organization, order, arrangement, and structure. The best explanation for this is that there is something other than the body as an ordered aggregate that makes a unique contribution to what a particular living thing does, and to the very nature of what that particular thing is. This explanation should be accepted unless superior claims can be offered to the contrary (Jaworski 2011, p. 296-297).

HD also seems to be the best explanation for the organocentric view of biology. There are two primary views regarding DNA and the process of morphogenesis (i.e. the biological process of organism development). These are the genocentric view and the organocentric view. According to the genocentric view, DNA is primary. That is to say that DNA is the fundamental unit of life and provides everything needed to produce an organism. On

this view, DNA is what is responsible for producing everything needed to construct the organism and control morphogenesis, as it contains all the instructions needed to make this possible. Brian Goodwin, in his work *How the Leopard Changed its Spots*, argues instead for the organocentric view. On this view living organisms are irreducible wholes and are the basic loci of morphogenesis. It is not DNA, which needs the organism as a unity before it can function, but the organism itself that is the primary source of information. Goodwin argues that it is not DNA that is responsible for your traits or genetic information, but something else. Modern biology can currently attest that DNA has the ability to coat the proteins, but that is all that science has reason to think that it does. What many biologists are now saying is that DNA seems to need a driver (Goodwin 1994, p.75). The two main functions of DNA require the coordinated activity of many complex molecules, and it can occur only within the context of the entire cell. This is species specific, so it depends on the nature of the specific organism for its unique activity. It is now known that DNA is not the only thing that is passed on in reproduction. Experiments have shown that if the nucleus of, say, a human egg is transplanted into an enucleated egg of a different species, then the egg of that particular species with the human nucleus will continue to develop in accordance with the pattern characteristics of its own species. The development will eventually cease, leading to premature death, but the embryo during that period it is able to grow and behaves as if there were no DNA present in it whatsoever (Wells 1993, p.15). What is responsible for this kind of growth, given that the DNA thought necessary to explain this growth is not present? The HD view says that an organism is an irreducible whole with its own internal structure. Interestingly Goodwin argues the very same thing, claiming that if the organocentric view is correct, then organisms have a species-specific principle of development whereby the various parts are genuine functional entities that exist for, and by means of, each other and the whole of which they are parts. If Goodwin is right, then organisms are substances and not merely ordered aggregates.

The Cartesian view of the body as an ordered aggregate would not be able to explain these facts as HD can.

VI. CONCLUSION

Substance dualism has had a bad reputation in modern philosophy for far too long. If HD turns out to be a more viable form of substance dualism, then attacking Cartesian dualism in order to refute substance dualism *simpliciter* amounts to a straw man argument. What I have attempted to show is that a better form of dualism can be borrowed from the ancients. This is HD. This view provides answers, or makes obsolete certain questions that arise regarding causal interactionism. Efficient causality is replaced with formal causality, which is grounded in the form of matter. Additionally HD seems compatible with the findings of neuroscience. Cartesian dualism seems open to the objection that, if the mind were as independent of the brain as the theory implies, then we shouldn't expect that brain damage could so severely impair mental functioning. But on the HD view, the soul is (on the Thomistic conception) almost as close to the body as the form of a chair is to the matter of the chair. Just as the form of a chair cannot function apart from the chair's matter, neither can the soul, for the most part, function apart from the matter of the brain and body. So we should expect, on the Thomistic version of HD, that damage to the body and brain would impair mental functioning. This is especially so given that, on the Thomistic form of HD, sensation and perception are, unlike the higher intellectual mental operations, purely material processes which cannot exist or function independently of the body (Feser 2006, p. 226). In this way HD provides a much more robust solution to the mind-body problem.

Notes

1. See Jaegwon Kim's first edition of his Philosophy of Mind text, (Kim 2006), where there is no discussion of dualism whatsoever. After receiving some helpful critique, he devotes an entire chapter to substance dualism in his second and third editions of this text. However, he only addresses Cartesian dualism and does not include any discussion of the Aristotelian or Thomistic

forms. Eric Olson goes so far as to attempt a treatment of hylemorphism in his text on human ontology, *What Are We?* (Olson 2007), and clearly admits that he simply does not understand it, which obviously makes for limited discussion of the view.

2. I realize that there are philosophers that would deny the title of “dualist” to Aristotle and Aquinas. However, I maintain that this is the result of disenchantment with Descartes view, which has created a widespread revisionist tendency among philosophers. See Howard Robinsons “Aristotelian Dualism,” in *Oxford Studies in Ancient Philosophy* (1983, pp. 123-144).
3. By naturalist epistemology I mean a presupposed epistemic backdrop that includes, but is not limited to, a commitment to the reduction of first-person conscious states to objective physical processes, to valuing theoretical simplicity sometimes at the expense descriptive accuracy, and to combinatorial modes of explanation. The hylemorphist, as well as certain kinds of dualists, will view consciousness not as a theory, but rather as a phenomenon that is directly experienced by those who have it. Although the materialist’s desire is to situate her solution to mental causation within this sort of epistemic framework, the question that arises is whether property dualism or functionalism is compatible with this framework. For example, John Searle claims that consciousness should be located within the purview of the evolutionary theory of biology and the atomic theory of matter, but does his dual attribute theory comport with his epistemic framework? If not, then it seems he has given up his claim to a superior form of explanation. To be sure, not all naturalists would hold to an epistemology exactly as the one described here, as arguably there are different denominations of naturalist. Some, like David Papineau, would fall into a strong naturalist camp, while others such as John Searle might be seen as advocating a weaker form of naturalism. If the naturalist epistemology gains its justification from its alleged superior explanatory power, then a more probing question to ask in light of this would be what is the advantage of “weak naturalism”? Does it make sense to claim that you have a weaker form of superior explanatory power?
4. Current HD advocates include John Haldane, Peter Kreft, David Oderberg, Ron Ticelli, and Edward Feser.
5. A.D. Smith admits that if we were to sever conceivability from our belief in possibility, we would end up with extreme Megareanism and the possible and the necessary would collapse into the actual (Smith 1997, p. 243).
6. While I entertain this option here for the purpose of refuting the causal pairing problem, I am in no way claiming that this is the HD view of souls.
7. By ‘substance’ I am referring to Aquinas’ and Aristotle’s view of substance, particularly as taught by Aristotle in the *Categories*. According to Aristotle, a substance is a unity of parts, properties, and dispositions. It is that which exists in itself, by itself and not in another, as in a subject (Connell 1988, p.15). In the context of the discussion here, a human person or any kind of living organism would be an ideal example of substance, in that a person

is a unity of properties, parts, and capacities. Substances “own” properties, and remain the same through change. In sum, a substance is an individuated nature, which is a whole that unifies its parts, properties and capacities, possesses accidental properties, remains the same through change, and matures in law-like ways towards the realization of its appropriate end or “maturity”. It could be referred to as a “this-such”, where ‘this’ refers to the particularity, and ‘such’ refers to the nature, e.g. humanness. This is in contrast with something non-living, like a measure of salt. This would be an example of an ordered aggregate or “property-thing”. It is described by the stuff it’s made of and what unifies that stuff. It is treated as the loose association of parts that it is. What gives it unity is a temporary spatiotemporal configuration. For ordered aggregates there is no absolute sameness or strict identity through change, as it gains new parts and loses old ones.

8. As I use the labels ‘Cartesian dualism’ and ‘HD’, I do so in a widely accepted way. I do not claim to accurately represent Descartes, Aristotle, or Aquinas in every detail. Nevertheless the way that I represent them captures the spirit, and in many cases the letter, of each thinker’s view.
9. The soul is “in” the body as its individuated essence or substance that “stands under”, informs, animates, develops, and unifies all the body’s parts and functions, thus making it human. The body is “in” the soul in the sense that the body is a spatially extended set of heterogeneous parts that stand to each other via internal relations that serve as an external expression of the soul’s need for a body (i.e. the structural set of capacities for forming a body to realize particular functions that are latent within the soul itself) (Moreland 2000, p. 205).

Bibliography

- Churchland, P. M. (1988) *Matter and Consciousness: A Contemporary Introduction to the Philosophy of Mind* (Cambridge: MIT)
- Connell, R. J. (1988) *Substance and Modern Science Vol. 3* (Houston: Center for Thomistic Studies, University of St. Thomas)
- Crem, Theresa M. (1979) “A Moderate Dualist Alternative to Cartesian Dualism,” *Laval Théologique Et Philosophique* 35(2), pp. 153–175
- Feser, E. (2007) *Philosophy of Mind: A Beginner’s Guide* (Oxford: Oneworld)
- Goodwin, B. C. (1994) *How the Leopard Changed Its Spots: The Evolution of Complexity* (New York: Phoenix Giant)
- Jaworski, William (2011) *Philosophy of Mind: A Comprehensive Introduction* (Oxford: Wiley-Blackwell)
- Kelly, E. F.; Crabtree, A.; Kelly, E. W.; Gauld, A. (2009) *Irreducible Mind: Toward a Psychology for the 21st Century* (Lanham: Rowman & Littlefield)

- Kim, J. (2005) *Physicalism, Or Something Near Enough* (Princeton: Princeton University Press)
- Kim, J. (2011) *Philosophy of Mind* (Boulder: Westview Press)
- Kim, Jaegwon (1996) *Philosophy of Mind* (Boulder: Westview Press)
- Kriegel, Uriah (2013) *Phenomenal Intentionality* (Oxford: Oxford University Press)
- Moreland, J. P. and Rae, S. B. (2000) *Body & Soul: Human Nature & the Crisis in Ethics* (Downers Grove: InterVarsity Press)
- Mourant, John A. (1957) "Cartesian Man and Thomistic Man," *Journal of Philosophy* 54(12), pp. 373–382
- Nichols, Terence L. (1996) "Aquinas's Concept of Substantial Form and Modern Science," *International Philosophical Quarterly* 36(3), pp. 303–318
- Oderberg, David S. (2005) "Hylemorphic dualism," *Social Philosophy and Policy* 22(2), pp. 70–99
- Olson, Eric T. (2007) *What Are We?: A Study in Personal Ontology* (Oxford: Oxford University Press).
- Pautz, Adam (2013) "Does Phenomenology Ground Mental Content?" In Uriah Kriegel (Ed.): *Phenomenal Intentionality* (Oxford: Oxford University Press), pp. 194–234
- Plantinga, A. (1978) *The Nature of Necessity* (Oxford: Clarendon Press)
- Robb, J. H. (1984) *Questions on the soul* (Milwaukee: Marquette University Press)
- Robinson, Howard M. (1993) *Objections to Physicalism* (Oxford: Oxford University Press)
- Smith, A. D. (1993) "Non-reductive physicalism?" In Howard M. Robinson (Ed.): *Objections to Physicalism* (Oxford: Oxford University Press)
- Tidman, Paul (1994) "Conceivability as a Test for Possibility," *American Philosophical Quarterly* 31(4), pp. 297–309
- Tooley, M. (1997) *Time, Tense, and Causation* (Oxford: Clarendon Press)
- Van Cleve, James (1983) "Conceivability and the Cartesian argument for dualism," *Pacific Philosophical Quarterly* 64(1), pp. 35–45

CONTRIBUTORS

Samuel Batzdorff. B.A. Philosophy, California State University, Northridge. Metaphysics, Philosophy of Mathematics, Philosophy of Language, Philosophy of Science. Samuel intends to apply to Ph.D. programs.

Lena Becerra. B.A. Psychology, University of California, San Diego. Metaphysics, Epistemology, Logic, Social and Political Philosophy. Lena plans to apply to Ph.D. programs.

Mark Gaynor. B.A. Philosophy, California State University, Chico. Moral Philosophy, Metaphysics, Logic. Mark intends to apply to Ph.D. programs in 2014.

Jared Gee. B.A. Psychology, Occidental College. Continental Philosophy, Post-structuralism, Phenomenology, Philosophy and Literature, Film Studies. Jared plans to continue to research and write on Derrida and his relationship to the history of philosophy.

Nathaniel Greely. B.A. Communications, University of Southern California. Philosophy of Mind, Philosophy of Language, Metaphysics. Nathaniel plans to apply to Ph.D. programs in the Fall of 2014.

Lorena Guandique. B.A. English, California State University, Los Angeles. Metaphysics, Plato, Berkeley, Philosophy of Mind, Existentialism. Lorena plans to apply to Ph.D. programs in the Fall of 2015.

Alysha Kassam. B.A. Cognitive Science, University of California, Irvine. Philosophy of Science, Epistemology. Alysha plans to apply to Ph.D. programs.

Nicole Lavoie. B.A. Philosophy, B.A. Linguistics, California State University, San Bernardino. Philosophy of Language, Metaphysics, Meta-Ontology, Semantics and Pragmatics, Meaning-theory, and the history of the English language. Nicole is applying to Ph.D. programs in the Fall of 2014.

Mark H. Morrow. D.C. Southern California University of Health Sciences. Philosophy of Mind, Epistemology, Metaphysics. Mark plans to apply to Ph.D. programs in the Fall of 2014.

Letizia Ragusa. B.F.A. University of California, Los Angeles. Letizia was kind enough to donate the cover art for this issue.

Ric Saenz. B.A. Philosophy, Biola University. Metaphysics, Epistemology, Philosophy of Religion. Ric intends to apply to Ph.D. programs.

Neil Sanchez. B.A. Philosophy, California State University, Northridge. Metaphysics, Epistemology, Philosophy of Language, History of Modern Philosophy. Neil intends to apply to Ph.D. programs.

Pedro Viramontes, Jr. B.A. Philosophy, California State University, Northridge. Ethics, Social and Political Philosophy, Legal Philosophy. Pedro intends to apply to law schools in the Fall of 2014.

MASTER OF ARTS IN PHILOSOPHY CALIFORNIA STATE UNIVERSITY, LOS ANGELES

The Department of Philosophy at California State University, Los Angeles offers a program of study leading to the Master of Arts degree in Philosophy. The program aims at the acquisition of a broad background in philosophy. It is designed for those preparing for further graduate study or community college teaching, and for self-enrichment. Although the department is analytically oriented, it encourages work in other areas, for example Asian philosophy, feminist philosophy, and the interaction between European and Anglo-American thought. The Department includes faculty members with diverse backgrounds and interests actively working in a wide range of philosophical specialties. Classes and seminars are small with a friendly, informal atmosphere that facilitates student-faculty interaction.

The academic programs in philosophy at California State University, Los Angeles are intended to engage students in philosophical inquiry. They aim to acquaint students with noteworthy contributions by philosophers to the tradition; to explore various philosophical issues, problems, and questions; to provide students with principles of inquiry and evaluation relevant to the many areas of human activity, such as science, law, religion, education, government, art, and the humanities; to develop in them skills of analysis, criticism, and synthesis needed for advanced work in various scholarly fields; to encourage the development of skills and attitudes leading to self-reflection and life-long learning.

PHILOSOPHY IN PRACTICE SUBMISSION INFORMATION

Each of the student contributors was specially selected to submit a paper for this issue of *Philosophy in Practice* by one or more faculty members in the Department of Philosophy at California State University, Los Angeles. All writers are currently either students in the master's program of philosophy or undergraduate majors in philosophy. All philosophy students at California State University, Los Angeles are eligible for nomination, and those who were chosen to contribute have demonstrated a superior ability to develop and compose works of advanced philosophical writing.

For more information on *Philosophy in Practice*, please contact:
philosophyinpractice@gmail.com