



α

Aletheia

Texas A&M Journal of Undergraduate Philosophy

Fall 2020

TABLE OF CONTENTS

NOTE FROM THE EDITOR-IN-CHIEF	4
NOTE FROM THE EXECUTIVE EDITOR	5
CONTRIBUTORS	6
On Plato’s Cognitive Powers and Theory of the Soul Julian Perilla, New York University	12
The Hylomorphic Soul View of Our Identity and Essence: Overview and Appraisal Philippe-Alexis Young Melody, George Washington University.....	24
Does the Partly-Social Narrative Account of Identity Resuscitate Person Essentialism in the Face of Criticism? Mia Ehudin, George Washington University	34
Buddhism, Stoicism, and Naturalistic Ethics Trevor Woodward, University of Idaho	44
Being-There as Mood: Fear and Anxiety in Heidegger’s Being and Time Arundhati Singh, University of Pennsylvania.....	56
EDITOR’S CORNER	
The Problem of Mental Causation and the Autonomy Solution Daniel Lightsey, Texas A&M University.....	66

NOTE FROM THE EDITOR-IN-CHIEF

It has been an honor and privilege to serve as *Aletheia's* Editor-in-Chief this academic year. It brings me great joy to see our goals come to fruition in this Fall 2020 edition of the journal. With the recent reformation of our organization, I hope this edition proves to be an exceptional beginning to a new age of *Aletheia*.

To the executive editor, Jake: Your contribution to *Aletheia* has reached far beyond this edition to the very core of our organization. You have put forward innovative ideas and systems which I believe will be in practice for years to come. For the effort you put forward these past several months, you have my deepest thanks.

To the editorial board: Between coronavirus and our organizational shift, this semester was wholly unprecedented. Nevertheless, you persevered and brought about an academic journal worthy of our admiration. It is due to your scholarship and diligence that we were able to push *Aletheia* to new heights and publish this edition. I am fortunate to have been a part of this board. Thank you for making this journal a reality.

To the scholars: This was the first year *Aletheia* welcomed submissions from outside of Texas A&M, and after observing the quality of your research, it was clearly the right decision. You are more than authors, you are scholars. Thank you for furthering the academic field of philosophy with your contributions.

To the reader: In these pages, you will discover arguments and ideas worth contemplating. I hope you become inspired and invigorated by the wisdom these scholars have shared, and continue to find joy in philosophical discourse.

Now, please enjoy this first-ever fall edition of *Aletheia*.

Eric W. Nash
B.A. in Philosophy
Class of 2022

NOTE FROM THE EXECUTIVE EDITOR

I would like to begin my thanks by acknowledging all the editors and contributing authors. It is only because of y'all that we have a journal to publish. Specifically I would like to thank Eric Nash for keeping spirits high, having enough faith in my abilities to elevate me to a position of experimental leadership, enduring my chaotic proclivities, and overall being an excellent editor-in-chief. Thanks as well are deserved by Professor Raymond for helping take Aletheia to the next level on the institutional end and Bec Morris for the lovely cover. Lastly, I would to acknowledge all those future readers attracted here by the edition year. May you find here some insight on this strange, strange moment of our present crisis only evident in past reflection.

Jake Donohue
B.A. in Philosophy
Class of 2021

CONTRIBUTORS

Editors

Eric Nash

Eric Nash is a junior philosophy major with minors in English and classics from New Braunfels, Texas. He is curious about all fields of philosophy but has a particular interest in philosophy of jurisprudence, environmental ethics, philosophy of religion, axiology, political philosophy, and ethics. In his undergraduate career, Eric has had the privilege of writing a thesis as Glasscock Summer Scholar, being a Fish Camp counselor, serving as a director for the sexual assault awareness committee of SLIDE, interning at the Hernandez Law Firm, and serving as the Editor-in-Chief for *Aletheia*. After graduating from Texas A&M, but before attending law school and pursuing a career in international law, Eric hopes to either work in the public sector or pursue a higher degree in philosophy.

Jake Donohue

Jake Donohue is a Texas A&M senior majoring in philosophy and minoring in gender studies. They serve as the executive editor of *Aletheia* and the president of private philosophy club Antithesis. Poststructuralism, gender theory, philosophy of art, and Nietzsche compose their areas of philosophical interest. Their current in-development project seeks to challenge Judith Butler's conception of gender through the ideas of Elizabeth Grosz, Mario Perniola, and late Roland Barthes.

Deborah Tate

Deborah Tate is a Junior Society, Ethics, and Law major and minor of philosophy, business and classics from New Braunfels, Texas. Her areas of philosophical interest include the philosophy of law, the philosophy of the mind, and ethical theory with particular interest in their impact on societies throughout history. Deborah is a member of several university clubs and organizations including the TAMU Women's Rugby Team. In her free time, she enjoys being with friends, playing the guitar, and blacksmithing. After graduation she plans to attend law school and pursue a career in corporate law.

CONTRIBUTORS

Editors

Daisy Cisneros

Daisy Cisneros is pursuing a bachelors degree in University Studies- Society, Ethics, & Law along with minors in philosophy and human resource development. Her philosophical fields of interest include ethics, epistemology, and science. Daisy describes making the Dean's List to be her favorite academic accomplishment throughout her time at A&M. She is currently President of another organization on campus, Gamma Phi Omega International Sorority, Inc. For the near future, Daisy plans to continue on to graduate school and eventually start a career in the field of human resources.

Joshua Legg

Joshua Legg is majoring in Society, Ethics, and Law and minoring in philosophy and military studies. Joshua's area of philosophical interest includes epistemology, axiology, ethics, and human rights. In addition to working as an editor for *Aletheia*, Joshua has previously worked as an intern for U.S. Congressman John Carter of the Texas 31st district. Joshua has also spent multiple years interning for Cofty Law PC. He plans to attend law school upon completion of his undergrad degree and hopes to concentrate in constitutional and libel law.

Daniel Lightsey

Daniel Lightsey is a junior philosophy and physics double major from Leander, Texas. His philosophical interests include philosophy of mind, philosophy of religion, philosophy of science, free will, and philosophical questions on the meaning and purpose in and of life. He is a recipient of the Phillip and Doris Moses Ranch Fund and the Crawford and Hattie Jackson Foundation Scholarship in Science for his academic success in college. In his free time, he likes to play board games with his friends and fiancée, cook, and read. Once he has graduated from Texas A&M, Daniel plans on going to graduate school and getting a PhD in philosophy and hopes to teach and influence the lives of his students.

CONTRIBUTORS

Editors

Savannah Benson-China

Savannah Benson-China is a Society, Ethics, and Law major with a double minor in philosophy and military studies from Mannheim, Germany. She is interested in political philosophy, social philosophy, and medical ethics. Savannah is a sophomore in the Corps of Cadets and a member of the Corps Women's Basketball Team. She is also a member of Pre-Law Society and First Generation Aggies. Upon graduation, Savannah plans to commission as a Second Lieutenant in the United States Air Force.

Haley Muff

Haley is a senior Society, Ethics, & Law major from Bastrop, Texas. She is primarily interested in environmental philosophy, ethical theory, and philosophy of religion. Haley is also Awareness Coordinator of Muster as well as a member of the women's organization Ryllies. After graduation, she plans to attend The Bush School of Government & Public Service for the ultimate goal of public service.

Patrick Henry-Featherston

Patrick "Henry" Featherston from San Antonio, Texas recently graduated summa cum laude with a double major in philosophy and history. While pursuing his education at Texas A&M, Henry served as a judicial advocate for the Student Government Association's judicial court, participated in the Public Policy Internship Program in Washington, D.C., was a member of Young American for Freedom, and was appointed to Board of Directors for Keep Brazos Beautiful. Henry's philosophical interests focus primarily on western political philosophy and philosophy of law. Henry has happily accepted the opportunity to join Representative Steve Toth's office as a Policy Analyst for the 87th Legislative Session. He is excited to serve the people of House District 15 while working first-hand with public policy as he goes on to law school in the Fall.

CONTRIBUTORS

Editors

Simon Holmes

Simon Holmes is a senior communication and philosophy major from Houston, TX. He enjoys studying political philosophy and philosophy of mind, with a focus in consciousness and seeking to better understand how the immaterial mind meets the material world. He is also passionate about education and learning more about our education system and its history from a cultural and societal perspective, as a way to better improve it for generations to come. He has served in the Memorial Student Center Freshmen in Service and Hosting committee (MSC FISH) as an Assistant Director and Executive Director as well as a two-year counselor and later Co-Chair for Texas A&M's extended orientation program, Fish Camp, which helped transition over 3,000 new students to college each year. After completing his undergraduate degree, Simon will be continuing his education at Duke University where he will obtain his Master of Science in Management. One day he hopes to work in the Consulting or Marketing field, helping companies' visions come to life.

CONTRIBUTORS

Authors

Philippe-Alexis Young Melody

Philippe-Alexis Young Melody is a fourth-year philosophy major and political science minor at George Washington University. He is from northern Virginia, and his fields of philosophical interest include political theory, Confucianism, philosophy of religion, medieval philosophy, and applied ethics. He would like to pursue a graduate degree in East Asian studies and enter academia.

Mia Ehudin

Mia is a fourth-year student at the George Washington University studying Philosophy with a Public Affairs focus and minoring in Sustainability and German. She is also a part of Phi Sigma Tau, the international Philosophy honor society. She is most interested in applied ethics as well as the philosophy of race and gender. Mia currently serves as Head Peer Educator at GW Students Against Sexual Assault, where she trains a team of peer educators to give workshops on sexual violence prevention and response to other student organizations on campus. Upon graduating in the Spring of 2021, she hopes to continue her work in this field and expand her horizons to gain experience educating others about reproductive health.

Trevor Woodward

Trevor Woodward is a third-year philosophy major at the University of Idaho. His philosophical interests are often in flux, but currently include the philosophy of language, metaphysics, and metaethics. He attended the 2020 Philosophy, Politics & Economics Society annual meeting in New Orleans as a moderator and received the University of Idaho's "Logic Award" for his work in symbolic logic in 2020. Trevor is a jazz drummer and enjoys listening to and performing music as a hobby. He hopes to attend graduate school in philosophy.

CONTRIBUTORS

Authors

Julian Perilla

Julian Perilla is an international student from Bogotá, Colombia, and a junior at New York University, majoring in philosophy with a minor in Latin and Ancient Greek. His main interests lie in moral and political philosophy, particularly in normative ethics. He is a Presidential Honors Scholar at NYU's College of Arts and Science and has been the recipient of several scholarships, such as NYU's Global Pathways Scholarship and the Herman Berkman Undergraduate Research Scholarship. His current research focuses on Nietzsche's ethics and metaethics, particularly on Nietzschean constitutivism, while his previous research has ranged from Platonic philosophy to Dante's *Commedia*. After graduation, he hopes to attend graduate school in order to continue his path in academic philosophy.

Arundhati Singh

Arundhati Singh is a senior at the University of Pennsylvania studying Comparative Literature and Philosophy.

On Plato's Cognitive Powers and Theory of the Soul

Julian Perilla, *New York University*

To this day, much has been discussed and written on Plato's notions of knowledge and opinion, yet, unfortunately, just as much has been neglected about his notion of ignorance. As difficult as it may be to attempt to understand the latter, I believe it to be a task worthy of our efforts. In this paper, therefore, I revisit Plato's argument in Republic V about the nature of these three cognitive powers, with a specific focus on the power of ignorance. In doing so, I offer a particular reading of the argument by presenting the correspondence between Plato's three cognitive powers and his three parts of the soul. First, I evaluate such correspondence at a general level through Socrates' characterization of each individual part of the soul and their corresponding part of the city. Subsequently, I reflect on how this correspondence should allow us to further understand Plato's notions of knowledge, opinion, and ignorance. I shall also show how this reading resolves much of the problems faced when attempting to understand his overall notion of a cognitive power. Finally, I conclude with some brief considerations on how this proposed reading of Plato may also shed light on the relationship between his cognitive powers and his theory of virtue developed throughout the Republic.

Introduction

So often when we speak about *knowledge*, we find ourselves in need of mentioning *ignorance* as well. It is not surprising, then, that Plato falls victim to this same need when writing about the difference between knowledge and opinion in *Republic V*. Starting at 475d, Socrates and Glaucon begin a discussion on the distinguishing factor between “true philosophers” and mere “lovers of sights and sounds.” This discussion ultimately leads to Plato's argument about the nature of three, not two, cognitive powers: knowledge (*epistēmē*), opinion (*doxa*), and ignorance (*agnōsia*). Granted, ignorance is clearly not the main concern at the moment, nor is it ever throughout the rest of the *Republic*. However, the fundamental role it plays in the argument and, most importantly, the problems it poses suggest that to simply overlook the matter would be a mistake. Consequently, I have attempted a closer consideration of Plato's power of ignorance and its relationship to the rest of cognitive powers in hope of it shedding light on much more central topics of the Platonic corpus. In this paper, I will offer a possible way of further understanding Plato's three cognitive powers—with a specific focus on ignorance—through the Platonic tripartite division of the soul by evaluating the possible correspondence between each of the abilities set forward in Book V and each part of the human soul established throughout the preceding sections of the *Republic*.

I.

Let us begin with the fact that *knowledge* and *opinion* are powers—which is to say abilities—and *ignorance* is at least suggested to be one as well (477d-e). Despite it being true that ignorance, as opposed to knowledge and opinion, is not explicitly said to be a power by either Socrates or Glaucon, it is safe to say that Plato most clearly treats it as such. We are told that powers ought to be distinguished in virtue of two attributes: (i) what they are related to and (ii) what they accomplish (477d). Each of these two conditions are invoked in order to distinguish ignorance from both knowledge and opinion as if Plato were, in fact, solely distinguishing among powers. Ignorance is related to ‘what is not’, while knowledge is related to ‘what is’, and opinion to what lies in-between both (I will have more to say on this later). And, opinion, Socrates says, is both darker than knowledge and clearer than ignorance (478a-e). This should be understood as a characterization of the corresponding cognitive state that each cognitive power accomplishes—knowing, ignoring, and opining, respectively. We are, thus, moved to take ignorance as a power in order to accept the soundness of these comparisons among powers, the distinctions that follow from them, and the definition of opinion as an intermediate between knowledge and ignorance.

Ignorance is a power—a cognitive power to put it in more contemporary terms—and yet, because of the obscurity of this assertion, little is accomplished by it as far as clarifying the nature of ignorance goes. To define ignorance as an ability is clearly an idea that lacks complete intuitive sympathy, to say the least, from current readers of Plato. We rarely think of ignorance this way since we are much more prompt to understand it in terms of knowledge itself; we think of ignorance, in some way or another, as an absence of knowledge. Thus, the question of how we are to understand Plato’s power of ignorance remains. All the more so, if one considers that, if ignorance is a power, then ignorance must have a set of intrinsic relata as a distinguishing factor from the rest of cognitive powers.

Powers are said by Socrates to be ‘set over’ their relata in such an exclusive manner that it constitutes one of the two distinguishing attributes among powers. In other words, no two distinct powers can have the same relata. The importance of this differentia is evident, but the meaning of the expression “set over” used to describe the nature of the relationship between each power and its corresponding set of relata has been a matter of much scholarly debate. It is often thought that each power is *about* or *of* its relata, as if a power’s relata were its subject-matter. The fact that knowledge, for instance, is set over ‘what is’ would mean, under this particular interpretation of Plato’s writing, that knowledge is about or of ‘what is’. Nevertheless, as common as this interpretation may be, when applied to Plato’s power of ignorance it seems rather limited, clarifying little of how we are to

understand ignorance as a power while raising further, and equally challenging issues. If we were to follow said interpretation, we would ultimately have to conclude that ignorance is about or of ‘what is not’, or—considering Socrates’ later clarification at 478b—about ‘nothing’ at all. But, the meaning of this is unclear at best.

The matter further complicates itself as one considers the logical consequences of asserting that each set of relata serves as a distinguishing factor of its corresponding power amidst the rest of powers. Namely, it follows that ignorance, opinion, and knowledge are about different things since they are, in fact, different powers. Not only is this counter-intuitive¹, but it leads Plato’s audience to an even deeper difficulty: the nature of these relata themselves. If we were to accept that knowledge, opinion, and ignorance are about different things, what exactly are these things? I have already mentioned that Socrates refers to these as ‘what is’, ‘what is not’, and what lies in-between both, respectively. Nevertheless, these characterizations are just as problematic as the problem itself. Therefore, we as Plato’s readers, at the end of *Republic V*, are left with the task of construing an interpretation of this argument comprehensive enough as to allow us to fully make sense of the nature of each power’s relata and the nature of the relationship between a power and its set of relata. This would certainly allow for a better understanding of the Platonic notions of knowledge, opinion, and ignorance.

II.

The human soul, Plato writes, is constituted by three different, coexisting parts: the rational, the spirited, and the appetitive parts of the soul. This particular division of the soul seems to be a product of Plato’s own advancement towards a much more complex theory of human psychology. It is the result of Plato’s rejection of the Socratic theory of human motivation, developed in his own earlier dialogues, and the recognition of a multiplicity of motivating factors—beyond intellect itself—for human action².

Socrates in Book IV of the *Republic* holds:

- (i) We humans, more often than not, experience conflicts among opposing desires in regards to the same thing at the same time (439c-440d).

¹ It seems natural to me to think that for any “x” that can be known, one could also claim that it can be ignored, and, most likely, subject of mere opinion. Thus, I have taken it to be that we commonly think as these three cognitive powers as being *about* the same things.

² For more on Plato’s theory of human motivation and the differences between it and the Socratic theory of human motivation of his earlier dialogues, see John M. Cooper’s “Plato’s Theory of Human Motivation.”

(ii) No same thing can be, do, or undergo opposites, at the same time, in the same respect, and in relation to the same thing (436b-437a)³.

From accepting these two claims it then follows that the human soul must have different parts which can, in fact, hold opposing desires among each other, allowing for the oppositional conflicts we have all experienced (439c-441c)⁴. Plato further believes that when one carefully examines the different possible motivational conflicts, his proposed tripartite division of the soul becomes evident, since we, he claims, have at least three different sources of desires and motives. Moreover, these three parts of the soul, Socrates explains, correspond to the three main parts of the city he and his interlocutors have identified beforehand—the rulers, the auxiliaries, and the producers. During the remainder of this section, I will be primarily concerned with this correspondence in an attempt to show that each part of the soul also corresponds to one of Plato’s cognitive powers. In doing so, I will primarily rely on Socrates’ characterization of both the parts of the soul and the parts of the city, as well as his remarks about their corresponding virtues—insofar as they are part of said characterizations. However, I will return to consider in full the relationship between virtues and cognitive powers in Section IV.

Perhaps the most intuitive correspondence, and thus the one I shall begin with, is that between knowledge and the rational part of the soul. This part, we are told by Socrates, is in charge of “rational calculation” (439d), and because of this reason-based deliberative ability it can ultimately come to possess “the knowledge of what is advantageous for each part and for the whole soul” (442c)—that is, when the virtue of wisdom is achieved. The same is true, *mutatis mutandis*, for the rulers of the city. Accordingly, if the corresponding virtue of the rational part of the soul is, in fact, a type of knowledge—and its virtue is that through which the function of the rational part is performed well (353b-e)—it seems reasonable for knowledge to be, at least, a fundamental part of its function. This would, indeed, suggest a strong correspondence between knowledge as a cognitive power and this particular part of the human soul.

Let us now consider the spirited part and, what I hold to be its corresponding power, opinion. Both honor and obedience become central in Plato’s characterization of this part of the soul. It is said to be honor-loving, suggesting a strong dependence on external influences and a motivation primarily determined by

³ In much more general terms, this claim can be restated in the following terms: It is impossible to have A and ~A at the same time, in relation to the same thing, and in the same part. This has been commonly known as Plato’s Principle of Opposites.

⁴ A useful example of how the argument works is found through 439b-439e: there are often thirsty people who do not wish to drink, thus there must be a part of their soul that wishes to drink and another different part that forbids it.

expected praise or blame rather than reasoned deliberation (475a)⁵. Furthermore, its corresponding virtue, courage, and its adequate functioning seem primarily based on its ability to obey the rational part of the soul and to preserve its opinions (429b-c). It, therefore, seems that this particular part of the soul is not only nurtured by opinion—that is to say, it is brought up and moved by external opinion—but its entire function is strongly dependent on it as well. Plato’s designated education for the auxiliaries, the part of the city that corresponds to the spirited part of the soul, further emphasizes the role of opinion in its well-functioning: while the auxiliaries share in the musical education of the rulers, providing them with a sense of the good, they have no access to the ruler’s later philosophical education that imparts them with true knowledge (402a seq)⁶. In other words, the auxiliaries—and by extension, the spirited part of the soul—ought only to have an opinion of the good, rather than pure knowledge of it. It seems then that Plato holds opinion to be central to the upbringing, functioning, and even to the particular virtue of the spirited part of the soul, which would suggest opinion as its corresponding cognitive power.

Lastly, the correspondence between ignorance and the appetitive part of the soul seems rather natural and intuitive. According to Socrates, this part of the soul is moved by its lusts, hungers, thirsts, and other many desires (439d) rather than by any reasoned calculation or educated belief. It is for Plato the naturally worst part of the soul, which, ideally, shall be ruled by its better counterparts (431 seq). Thus, it seems that this inherent lesser condition of the appetitive part of the soul could be, at least partially, explained by its corresponding cognitive ability. Perhaps, given that the rational part of the soul is naturally capable of knowledge and the spirited part of the soul of opinion—which, as I briefly mentioned above, can be trained and educated to hold correct beliefs—then the appetitive part of the soul is left to hold this inferior status and, in those who are moderate, must be restrained. This particular characterization of the appetitive part of the soul makes it increasingly difficult to hold that it could be capable of any other cognitive power different from ignorance.

If we were to be wrong about this matter and, in fact, the appetite part could, in turn, have opinion or even knowledge, then it would be possible for it to become better than the other parts of the soul. Why would Plato, then, so emphatically hold it to be the naturally worst part of the human soul? There ought to be an inherent

⁵ Socrates’ description of how the timocratic youth—he who is governed by the spirited part of the soul—comes into being is also continually marked by the influence of external and, mostly, uneducated opinion.

⁶ For more on this distinction see *Meno* 97d-100b on the statues of Daedalus. Through the analogy of the statues of Daedalus one can easily understand the difference between the sense of good that the auxiliaries have in the city and the rulers’ knowledge: it is only through philosophical education that those of gold blood tie down their opinion of the good provided by their previous musical education.

condition for this to be true in any and every case, namely the impossibility for opinion and knowledge⁷. This, I take it, does not mean that this part is necessarily bad nor that it cannot be controlled in order for it to judge in accordance to the rational or the spirited part of the soul—that is, in fact, part of the Platonic definition of moderation⁸. I will return to this in the following section. In any case, this shall be sufficient to show that both the characterization of the appetitive part of the soul and its innate inferior status, given by Plato, surely suggest a correspondence between ignorance and this part of the soul.

III.

In this section, I will present, what I take to be, the consequences of accepting these existing relationships between each part of the soul and their corresponding powers. I shall begin with some preliminary clarifications.

The human soul, Socrates explains, is ultimately able to grasp the Forms and acquire true knowledge (517b seq). Since we have no reason to suspect that knowledge could be an ability of the soul while the other two powers abilities of the body, it seems reasonable for all three cognitive powers to be abilities of the soul. Accordingly, as I explained above, each part might have its corresponding power. Socrates' characterization seems to favor, at least to some extent, the agent-like interpretation of the parts of the soul, insofar as each can create its own judgments⁹. The rational part of the soul and the rulers of the city are often referred to as “lovers of learning” (376b), the spirited part and the auxiliaries as “honor-lovers” (475a), and the appetitive part and producers as those who have “love for money” (436a). This would entail that each part is able to create judgments, at least of the sort of “learning is desirable”, “honor is desirable”, and “money is desirable”, respectively. If so, Plato's cognitive powers may be the abilities through which their corresponding parts of the soul produce their judgments—which seems not too distant from how we nowadays think of cognitive powers.

⁷ It is worth mentioning that Plato, in his *Timaeus*, explicitly denies that the appetitive part of the soul could have either opinion or knowledge: “This type [the appetitive part of the soul] is totally devoid of opinion, reasoning or understanding, though it does share in sensation, pleasant and painful, and desires” (77b).

⁸ Allow me to clarify that I am here, and for the remainder of this paper, understanding opinion (*doxa*) as distinct from mere judgment (which has no corresponding greek word). It then follows that when Plato uses the word *homodoxein* to refer to that which appetite has when it agrees with the other parts of the soul he is not saying, I take it, that it literally has *doxa*.

⁹ The agent-like interpretation of the parts of the soul claims that Plato treats each part of the soul as the subject of psychological affections, activities, and capacities that are normally attributed to a person as a whole. At this point, I am not concerned with determining if Plato truly gives complete agent-like nature to each part of the soul, which would mean that each part has its own desires, practical goals, beliefs, reasoning, and the ability to communicate; but rather I will endorse this view partially, insofar as I believe that each part has at least the ability to create its own judgments. For more on this matter, see Bobonich's “Parts of the Soul and the Psychology of Virtue”.

The onus of differentiating the three powers by following Socrates' two distinguishing conditions should not be a disincentive to follow this interpretation. From the fact that all three cognitive powers produce judgments, it does not follow that they ultimately accomplish the same thing, but rather that each power when producing its particular judgments, based on different levels of reasoning, ultimately produces completely different cognitive states. Consider the judgment "Socrates is a just man." I would believe that, even intuitively, one could easily assert that this judgment could be the product of any of the three cognitive powers—that is to say, of knowledge, opinion, or ignorance. Thus, the distinguishing factor among powers must lie beyond the fact that all three are abilities to produce judgments and, *a fortiori*, beyond the content of the judgment produced by any of them. Rather, following Smith¹⁰, I hold that what truly distinguishes Plato's cognitive powers from one another is how each judgment has been produced.

In order to clarify this, I will briefly consider the matter of the relata of knowledge and the relata of opinion. Near the end of Book V, Socrates holds that the Forms are the only things that truly *are*—as well as the only objects of knowledge—while perceptible things are always in between being and not being (479d-480a). This seems to match quite nicely with Socrates' previous characterization of the relata of knowledge as "what is" and the relata of opinion as what lies in between what is and what is not. Therefore, one may conclude, reasonably enough, that the Forms are, in fact, the relata of knowledge and that perceptible objects are the relata of opinion. Let us now return to the judgment "Socrates is a just man" to clarify this point and the previous conclusions on the matter of each power's differentiae.

He who has knowledge judges that Socrates is, indeed, a just man after he has grasped the Form of Justice and, thus, has reached the highest level of understanding, producing an infallible judgment. Most differently, he who has opinion may well have judged that Socrates is a just man, but on the basis of some perceptible thing, such as the decision of a jury or a ruling of a judge—or even the mere opinion of some other man—demonstrating a considerable lower level of understanding and producing a fallible judgment. In this case, the two distinguishing factors among powers become easily identifiable: that which produces the judgments and the level of understanding each one accomplishes. The

¹⁰ I follow Smith primarily in his interpretation of the expression "set over" in regards to the relationship between each power and its corresponding relata (a matter I will consider next), as well as his interpretation of the accomplishment condition that serves to distinguish among all three cognitive powers. It is each power's relata *by* which each corresponding cognitive state is produced, thus knowledge, opinion, and ignorance could, indeed, have the same subject matter and even produce the same judgment while still being distinguished by Socrates' two conditions, their relata and what they accomplish. However, I differ from him insofar as he holds the accomplishment of each power to be a "conception" rather than a cognitive state—or its underlying level of reasoning—and, all the most, as his resulting notion of Plato's power of ignorance claims it to be produced by a misconception of any given F-ness. For more, see Smith's "Plato on the Power of Ignorance".

chain of reasoning of he who judges with knowledge would be something of the sort of “Justice is x thus Socrates partakes in Justice,” while the chain of reasoning of he who opines would be something like “the Assembly has decided that Socrates is a just man; thus, he is a just man.”¹¹ It seems then that this particular account of Plato’s cognitive powers of knowledge and opinion coherently maintains the two distinguishing conditions between powers Socrates establishes in Book V while allowing the relation of each power to be that *by* which each corresponding judgment is produced.

Accordingly, consider the case of the ignoramus. Based on the correspondence explained above between the appetitive part of the soul and the power of ignorance, it seems that any of its judgments would then be entirely based on lusts and bodily desires—or any other irrational motive. For instance, the mad lover of the *Phaedrus* (250d-256e) or the erotic tyrant of the *Republic* (572e-576a), if they were to be in love with Socrates, would probably claim that Socrates is indeed a just man. However, their chain of reasoning in doing so would, probably, be of the sort “I love Socrates; thus, Socrates is a just man,” which demonstrates no underlying reasoning and no true level of understanding about the subject matter in question. Perhaps, this is how the expression “what is not” should be understood: ignorance relates to considerations which have nothing of rationale and which produce nothing of understanding. Therefore, it seems that Plato’s power of ignorance is the mere ability to produce judgments completely based on irrational desires and emotions. As a consequence, these judgments produced by ignorance are fallible and, most importantly, much more prone to error than those produced by opinion and knowledge. Plato’s notion of ignorance, thus, seems very much coherent with our own intuitions. Even some expressions we use today to refer to situations in which we judge based on passions, such as “love clouded my judgment” or “I was blinded by fear,” are not too distant from Plato’s own examples of those who judge through the appetitive part of the soul, not too distant from the “mad and deranged” tyrant of the *Republic* or the horse with “bloodshot white eyes” of the *Phaedrus*¹².

¹¹ This view is somewhat similar to what Julia Annas argues in the chapter “Belief, Knowledge, and Understanding” of her *Introduction to Plato’s Republic*. In this regard, I do believe that an interpretation of Plato’s cognitive powers in favor of some different levels of understanding would entail different levels of reasoning and vice versa. Nevertheless, Anna’s account seems to fall short in other respects such as the solution to the relation problem and the way these powers differ in regards to what they relate to.

¹² In the *Phaedrus*, Socrates describes the human soul’s structure through the renowned chariot analogy: the soul is like “a team of winged horses and their charioteer”, where one horse is “beautiful and good” and the other one is the opposite (246a-b). As the discussion progresses, Socrates offers a more detailed characterization of each horse and refers to the unruly one—purportedly believed to represent the appetitive part of the soul—as with “bloodshot white eyes” (253e). I am not interested here in raising any theoretical point about the analogy or the general argument developed in the *Phaedrus*. Rather, I am merely pointing out the extraordinary similarity between some of our common expressions and Plato’s own descriptions of those who judge out of passion.

IV.

Finally, allow me to conclude with a few remarks on the relationship between Plato's virtues and his cognitive powers. As I mentioned above, each virtue considered by Socrates through *Republic IV* becomes a fundamental part of his characterization of each part of the soul. They ultimately allow us to take a glance at how each part ought to work and what each one lacks when functioning in a deficient manner. They additionally help us understand the role each cognitive power plays in the correct or incorrect functioning of its corresponding part of the soul, as I have attempted to show. However, this shall not mean that the present interpretation of Plato's cognitive powers depends on each part of the soul having already achieved such virtues nor does it rely on the soul ever achieving a virtuous state at all. The claim altogether is that even those who are non-virtuous are capable of, and indeed possess, all three cognitive powers.

Consider the oligarchic man with his corresponding oligarchic soul described in *Republic VIII*. He is a lover of money who has established his appetitive part as the ruler of his soul, above both the rational and the spirited parts. Nonetheless, Socrates does not speak of him as completely dumb, indeed he is capable of some kind of reasoning and he, in fact, is most capable of admiring some things. It is in these abilities where we find the power of knowledge and the power of opinion in him. The rational part, in this case, thinks of ways of gaining money and the spirited part values and admires wealth and wealthy people (553d). In other words, one has knowledge of money or profit and the other opines about wealth and wealthy people. Is it, therefore, possible to have knowledge and opinion without being virtuous? Plato, in the *Republic*, seems to allow for this possibility. The problem, thus, with the oligarchic soul, as well as with the rest of non-virtuous souls, is their constitutions instead of their lack of knowledge as a whole.

It would indeed be a mistake to claim that virtues and cognitive powers, for Plato, are completely independent—after all, wisdom is a type of knowledge. But, it would be equally mistaken to claim that possessing the power of knowledge, this mere ability Socrates has defined in such a technical and narrow sense, would necessarily entail virtue. The key to virtue, Socrates seems to suggest throughout the *Republic*, is the soul's constitution. Each constitution provides a principle or set of principles based on which one will ultimately live one's life. These principles are the product of a scheme of values, and a notion of value itself, that determines one's ends and pursuits. It is because of the constitution of his soul that the timocratic man values honor and victory above all other things, much like the oligarchic man values money and wealth. And, in their own particular organization of the three parts of the soul, each part continues to possess and perform its corresponding cognitive power, but to the ends determined by the ruling part. If one is virtuous,

reason will rule the soul and one's ends will be established by reason through knowledge; otherwise, each part's cognitive ability will be no more of a slave to the vicious ruler than the enslaved part itself.

Knowledge in general, as Socrates treats it in the *Republic*, does not make one virtuous nor does it require virtue at all. Nonetheless, there are some *types* of knowledge or knowledge of some things that constitute virtue or necessarily move towards a virtuous constitution—or are the product of one. For instance, I have already mentioned in Section II that wisdom is knowledge of what is advantageous for the soul. Furthermore, virtue itself requires knowledge of the Good and the individual virtues. However, there is nothing in Socrates' account of each of the non-virtuous souls, nor in our intuitive thinking, that suggests that non-virtuous people are indeed incapable of any sort of reasoning or knowledge. Quite the opposite, Socrates repeatedly refers to the condition of the rational part of the soul and its labors determined by the ruling part. It does not seem unreasonable for the timocratic soul to have knowledge of honor nor for the oligarchic soul to have knowledge of money. After all, the rational part is made a slave but remains active. The true problem, then, lies in the constitution of the soul and its resulting ends. If the rational part of the soul is not allowed to rule, one's judgments and value-judgments are to be produced by either opinion or ignorance—and will thus be more prone to error, to some or other degree.

Works Cited

- Annas, J. (1981). Belief, Knowledge, and Understanding. In *An Introduction to Plato's Republic*.
Oxford: Clarendon Press; New York: Oxford University Press.
- Bobonich, C. (2002-07-11). Parts of the Soul and Non-Rational Motivations. In *Plato's Utopia Recast: His Later Ethics and Politics*: Oxford University Press. Retrieved 11 Oct. 2019, from <https://www.oxfordscholarship.com/view/10.1093/0199251436.001.0001/acprof-9780199251438-chapter-4>.
- Cooper, J. M. (1984). Plato's Theory of Human Motivation. *History of Philosophy Quarterly*, 1(1), 3-21. Retrieved from <http://www.jstor.org/stable/27743663>
- Cooper, J. M. (2013). *Pursuits of Wisdom: Six Ways of Life in Ancient Philosophy from Socrates to Plotinus*. Princeton, NJ: Princeton Univ. Press.
- Plato. (1997). *Plato: Complete Works*. J. M. Cooper (Ed.). Indianapolis: Hackett.
- Smith, N. (2012-08-02). Plato on the Power of Ignorance. In *Virtue and Happiness: Essays in Honour of Julia Annas*: Oxford University Press. Retrieved 11 Oct. 2019, from <https://www.oxfordscholarship.com/view/10.1093/acprof:oso/9780199646043.001.0001/acprof-9780199646043-chapter-4>.

The Hylomorphic Soul View of Our Identity and Essence: Overview and Appraisal

Philippe-Alexis Young Melody, *George Washington University*

The problem of numerical identity is typically framed in terms of a contrast between soul-based theories of identity that posit a non-physical essence of human persons and materialist theories according to which human persons are essentially physical structures of some sort such as organisms or embodied minds. Both theories suffer from conceptual problems that conflict strongly with intuitions about the nature of identity. Aristotle's hylomorphic approach to metaphysics offers an alternative that has not been analyzed as thoroughly in contemporary debate and that avoids the problems associated with other theories. Although not free of all conceptual difficulties, hylomorphism is a viable theory.

Introduction

In the popular chestnut known as the ship of Theseus, a ship's planks are removed one by one and replaced with fresh ones over a prolonged period until eventually none of the original planks constitutes the ship. Intuitively, the ship is understood as being the same one as it was when the last plank was removed, but its composition is now totally changed. This thought experiment raises questions about what the criterion should be for something to persist over time after undergoing alterations. In the case of the ship of Theseus, the ship undergoes a complete change of its parts. However, it still appears to be the same ship despite these modifications. In other words, the ship S_2 at time t_2 is numerically identical to the ship S_1 at time t_1 .

This issue of numerical identity--that is, establishing the persistence conditions for objects after qualitative change--also reveals itself when we attempt to understand human persons. An adult is radically different, qualitatively, from what s/he was as an infant but appears to be the same individual despite that fact. According to the hylomorphic thesis human persons are essentially composites of form and matter. The form of human beings is the soul which is something all organisms have from trees to conscious creatures. When this form is realized in matter, an individual organism exists.

In what follows, the hylomorphic theory will be explained and analyzed beginning with a presentation of the case in its favor. Afterwards, a series of challenges to this theory will be presented. Finally, responses to these challenges will be advanced, and the theory will be assessed to see what its prospects are in addressing the issue of identity.

The Cartesian Soul

Hylomorphism is a theory based on the idea of the soul. The word “soul” might carry religious connotations or seem to evoke theological or pre-scientific ideas about psychology, but the hylomorphic soul is distinct from this idea of the soul as will be explained below and avoids the central problems that confront typical understandings of the soul.

In spite of the possible connotations of the term “soul,” a theory of identity based on something beyond mere matter is appealing because it makes sense of various intuitions that link identity with subjectivity. It is easy to imagine someone’s thoughts and psychological qualities inhabiting a different body after a swap occurs in a *Freaky Friday* scenario. In such a case, it makes intuitive sense to suppose that this is because what is essential to each person is not the body but the soul, or in more modern terminology the mind, regardless of whether such a case could transpire in this possible world. Therefore, a theory of identity that can make sense with this point in mind might resolve the issue, at least more adequately than those that do not.

The popular conception of the soul finds a strong articulation in the work of Descartes. According to Descartes, human persons are in essence *res cogitans*, that is, thinking things. Because it is possible to doubt the existence of the body but not the existence of the mind that does the thinking, Descartes believed that our essence was mental. In other words, human persons are fundamentally subjects of experience as opposed to extended bodies in space. Although a Cartesian might believe the self cannot in reality be separated from the body, it can conceptually be separated without forfeiting identity because it is said to be an independent substance. Thus, the Cartesian theory makes sense of intuitions that suggest psychological or subjective properties are sufficient to retain the numerical identity of human persons.

There emerge two major problems that the Cartesian must address: why each soul seems causally connected in some way to a body and how consciousness can appear to split. The first challenge is known as the interaction problem and can be variously phrased, but the central issue is how an immaterial soul can have any causal connection with a material body without violating the law of conservation of energy. The body in general, and the brain in particular, interacts regularly with our mental life insofar as bodily events cause changes in attitude, feeling, cognition, and other subjective properties. The Cartesian might respond with the parallelist view that changes in the soul happen to accompany changes to the body, but that would entail an awfully lucky number of coincidences. Without appeal to further theological premises, a resolution seems far-fetched. It is more parsimonious to jettison the idea of the soul being an independent substance when explaining its

interactions with the body than to accept this response or to be forced to accept a theological premise that not everyone can countenance.

The second challenge can be found in the recent work of Derek Parfit. Parfit begins with a real-life medical advancement that results in split brain patients. In the case of a split brain, the patient's corpus callosum is severed so that each hemisphere operates independently thereby creating two separate spheres of consciousness linked to one body (Parfit, 245-246). This scenario illustrates the problematic relationship of soul and body. A change in the body causes a change in the soul even though they are said to be independent substances. Moreover, the creation of new centers of consciousness by splitting the brain would entail the splitting of the soul if it is the center of consciousness. A soul, however, is supposed to be a thinking thing, not a spatially extended thing. The absence of spatial extension suggests that a soul is not divisible, which implies that a split in the brain should not lead to multiple souls.

Parfit then offers a second case of dividing minds but this time in a case that has not yet been carried out. He imagines that after having had his body paralyzed, his cerebrum, the part of the brain believed to give rise to conscious experience, is divided into two and transplanted into the bodies of his two twin brothers whose brains have been removed (Parfit, 254-255). The resulting individuals are both psychologically continuous with Parfit but cannot be the same numerically since they have now branched off into two separate bodies and centers of consciousness. The Cartesian account is inadequate in addressing such a problem because it would again fail to explain the relationship between an immaterial soul and a spatially extended cerebrum. Furthermore, it would be unable to account for the emergence of a second center of consciousness from a split brain.

Minds and Organisms

In light of the apparent issues with positing the presence of a soul and its constituting the essential property of human persons, one may be inclined to favor a materialist theory. When identifying individuals of other species, it is common to do so on a biological basis. Defining human persons in terms other than their biological organism may appear to be an instance of speciesism. This view will be referred to as animalism. According to animalism, human organisms do not merely constitute human persons, rather, we essentially are human organisms or members of a biological category such as the genus *Homo* or perhaps the species *Homo sapiens* (DeGrazia, 48). Instead of understanding ourselves as essentially persons, the category person is simply one phase of human identity.

Jeff McMahan argues against animalism and puts forward a view that puts more emphasis on the psychological dimensions of human persons. What I take to

be the two strongest objections McMahan presents against animalism are the case of dicephalic twinning and the possibility of brain transplants. In the case of dicephalic twinning, there is one organism and there appear to be two persons. McMahan imagines an extreme case where the twins are conjoined everywhere except their brains or everywhere except their cerebra (McMahan, 35). In such a case, a single organism consisting of one set of organs and limbs gives rise to two centers of consciousness, each with its own interests and beliefs, and therefore two persons. If they are essentially their organism, they would be identical with each other even though they are two different subjects of experience.

The theoretical possibility of brain transplants, or more specifically cerebrum transplants, also give reason to be skeptical of animalism according to McMahan (McMahan, 31). If a person, Michael, has his cerebrum transplanted into a new body that becomes psychologically continuous with (pre-operation) him after the operation, it would appear that he would follow his consciousness instead of following the surviving human animal that is no longer conscious. Therefore, people would probably refer to the new body and its mind as Michael even if the old body is kept alive but unconscious. According to animalism however, Michael continues to be the unconscious body rather than the psychologically continuous being in the new organism because that animal is still alive.

To account for our intuitions, McMahan rejects animalism in favor of what he calls the embodied mind view according to which the essential properties of human persons are the capacity for consciousness and the embodiment of what gives rise to that capacity for consciousness (McMahan, 67). The embodied mind theory can deal with the two challenges above by stipulating that the relevant parts of the brain continuing to function without division is necessary for a person to remain the same one over time and that when these parts are embodied in such a way that they give rise to a connected consciousness, the person is the same numerically.

Several problems arise for the embodied mind account of identity. The first is that the qualities associated with people do not always coincide with those of their brains or cerebra. That is why most people do not describe their coloration as grey or their weight as a few pounds for example (DeGrazia, 72). The next is the fetus problem according to which human persons were early (pre-sentient) fetuses at some point, but early fetuses have no capacity for consciousness (Schechtman, 53). Therefore, the embodied mind account must either reject the belief that anyone was an early fetus at some point or be incorrect, and in everyday speech, most people do not seem to have a problem speaking of having been a pre-sentient being at some point. The next problem is the vegetable problem: that the embodied mind account is incompatible with the idea that a person could continue to exist in a permanent vegetative state because that person would have no capacity for consciousness

(Schechtman, 53). The embodied mind theorist would thus have to say that when Christina enters a persistent vegetative state (PVS), it is not she who is in a PVS; rather, she has ceased to exist while the body that coincided with her continues to exist, but Christina's relatives would probably continue to treat her vegetative organism as being Christina in some way.

McMahan can respond to the first objection by saying that the rest of the body is part of the person, and the brain is what is essential to the person without which the person could not exist (McMahan, 93). It is ordinary to think of the most basic level of an entity as being the most real part, such as subatomic particles, and the complete structure as existing in a nominal sense, so the same could be said of persons and their brains. He could respond to the fetus and vegetable problems by suggesting that individuals are essentially persons and that since their organisms no longer have the capacity for consciousness, the person is gone and only the organism remains even if in everyday speech or in emotional contexts people act as if there is no such distinction.

These responses seem feasible even if they are not conclusive, but there is one objection that provides grounds to ultimately reject the embodied mind account. The thinking animal problem responds to McMahan's reply that the person and organism are distinct entities that coincide under normal conditions. If that is the case, the human animal must be thinking when the person is doing so. Thus, there must be two thinkers when someone is thinking and it cannot be known from the first-person perspective whether it is the animal or the person that/who is thinking (Schechtman, 54). The distinction made between animal or organism on the one hand and person on the other cannot be sustained in light of this problem.

The Hylomorphic Solution

The apparent flaws of the Cartesian thesis may seem to preclude the possibility of a soul, but the proposed alternatives of animalism and the embodied mind account also have their difficulties. A theory that can make sense of the intuitions that make these other theories attractive while also avoiding their problems is thus the most likely to succeed. Hylomorphism can retain the idea of a soul and some of its strengths as a theory while avoiding its central problems. It can combine the soul with animalism so as to constitute a theory that accounts for both the spiritual and material dimensions of human persons while steering clear of the pitfalls associated with the embodied mind account of identity.

Hylomorphism holds that physical objects are material instantiations of forms. The form is the nature of the object or what makes it the type of thing it is, while the matter is the underlying base (Ainsworth, 1). Matter according to Aristotle has potential, whereas the form is a matter of actuality. As an illustration, clay can

potentially take the form of a brick when it has in actuality the form of a pot. The distinction between form and matter is similar to that between universals and particulars in that a particular piece of matter can take the form of a universal idea. Despite the word “form” being used, it is not simply the shape that determines an object’s form. For example, a toy turtle might have the same shape, including size, as a real turtle, but it does not have the same form in an Aristotelian sense. Rather, it has the form of a toy and is composed of plastic, so both its form and matter differ from those of a living turtle.

Aristotle’s analysis of change is based on his understanding of matter and form. His account of change supposes that for a change to transpire something must persist through the change. When this persisting entity is a substance, an accidental change has occurred (Ainsworth, 1). For example, if a car gets painted or has a part removed, it still exists but has undergone a change to its properties. When matter survives a change of substance, a substantial change has occurred (Ibid, 1). For example, when a brick is melted so that it can be molded, and it is turned into a clay bowl, the matter continues to exist throughout the process. However, the substance went from being a brick into being a bowl.

Under hylomorphism the form of an organism is its soul. Unlike the Cartesian soul, the hylomorphic soul is not a substance separate from the body. Rather, it is the substantial form that animates an organic body and makes that living body what it is (Aristotle, 22). A living body cannot exist without its soul, and a soul cannot exist without its living body. Hylomorphism therefore avoids the interaction problem associated with Cartesian dualism.

Three kinds of souls exist according to this theory. Plants and other non-sentient life have vegetative souls because they perform basic life functions such as growth and reproduction (Aristotle, 25). Non-person animals have sensory souls meaning they have the powers of the vegetative soul as well as sentience and commonly locomotion (Ibid, 25). Finally, persons, whether human or not on my interpretation, have intellectual souls. Intellectual souls have the powers of sensory souls as well as the power of the intellect which receives forms like the sensory soul receives external stimuli. Therefore, the intellect is purely potential until it receives a form (Marenbon, 81). The thinking animal problem does not apply to this theory because it does not treat the soul as a separate substance from the organism, so the organism is thinking because its form is that of a rational animal.

Objections

The hylomorphic theory, like the others, has problems of its own. Firstly, McMahan argues that the hylomorphic conception of the soul implies we do not begin to exist until well after birth. Next, since hylomorphism posits the presence of

a soul, some have argued it still runs into certain difficulties associated with Cartesian dualism even if it does not have the interaction problem. Finally, it also must deal with the objections to animalism since it can be interpreted as a kind of animalism.

McMahan argues that by conceiving of human persons as composites of matter and an intellectual soul, hylomorphism leads to the self-defeating conclusion that they do not come into being until their rational capacities develop (McMahan, 12-14). Because the rational capacities of fetuses and infants are below those of members of other species, rationality cannot be what distinguishes individual humans from them. It must then be conceded that the early human organism possesses a vegetative and later a sensory soul before it develops the capacity for rationality. This implies that new souls are being generated and old ones shed because the rational soul possesses the powers of the other two which then implies that the organism is acquiring a new form as its capacities increase.

The proponent might reply that the fetus and infant do have a rational soul, but it cannot yet be exercised. The important fact is that they belong to the kind of being that is internally oriented to eventually do so (McMahan, 13). McMahan dismisses this response by noting that the potential for rationality is not the same as the capacity for it and that if the soul is the organizing principle of an organism the kind of soul it has is based on its current capacities, not on possible future ones (Ibid, 13).

McMahan is correct that the early human organism does not yet have the capacity to exercise rationality, only a potential, but it does not follow that it does not have a rational soul. As mentioned above, the intellectual soul has the capacities of the vegetative and sensory souls meaning that a rational animal does not have three souls but rather one with the capacities of three (Hess, 84). This suggests that because the natural potential to eventually exercise rational capacities is ingrained in the typical early human organism, it has a rational form as its organizing principle. If it did not, that would necessarily imply that a substance change is undergone when the sensory and later intellectual capacities arise because the matter would remain continuous as the original form is replaced (Hess, 90). Given that a new substance does not emerge when a child first begins to conceive of universal forms (which Aristotle characterizes as the unique work of the intellect and thus of the intellectual soul), it should be understood as having had the same form to begin with whose capacities have developed as accidental changes have enabled their realization.

Although the hylomorphic account does not have the interaction problem associated with other soul-based theories, some might argue it still fails to make sense of divided consciousness. If two centers of consciousness can animate a

single body, or a brain can be split in two and be transplanted into two new bodies, there seems to be a challenge for any idea of the soul. The first problem can be addressed by thinking of the two centers of consciousness as being the product of one soul. The two centers of consciousness are the work of one organism, and each organism has one unifying principle or soul. Even if two theaters of experience exist, there is still a single living being undergoing those experiences. This contrasts with the Cartesian concept of the soul which is not defined as a unifying principle but rather as a source of subjective experience.

Before addressing the problem of dividing and transferring two halves of a cerebrum into new bodies, it is first necessary to answer one of the objections to animalism, namely, the brain transplant issue since Parfit's problem builds off of it. If a brain or perhaps just its cerebrum can be transplanted into another body, the organism is not the same, the matter lacks continuity with that of the original organism, and yet the person seems to have survived. To deal with this, the proponent of hylomorphism can concede that the new person is psychologically connected and even psychologically identical with the original while disputing that that person is numerically identical because that person is not the same composite of matter and form. This response might seem *ad hoc* since the new version of the individual will likely be treated as the original in terms of relationships, entitlements, and so on. However, viewing the response this way is a result of sneaking the premise that numerical identity is what matters in interpersonal relations, so disagreeing with this premise saves the response.

As for Parfit's problem involving a cerebrum being divided and transplanted into two different bodies, hylomorphism can make sense of their different numerical identities based on the fact that they are distinct organisms composed of different matter. Because the newly animated bodies are made of different matter, they are not the same individuals even if they are psychologically identical with the original and with each other.

Finally, there is the issue of dicephalic twinning. In this case, two persons are part of one organism. If they share the same matter and both have the form of an intellectual soul, hylomorphism would seem to entail that they are the same person. A similar problem arises when imagining one individual going through enough accidental changes so as to be composed of the same matter as another individual at some earlier point (Ainsworth, 3). If they both have the form of a rational animal, they would then be numerically identical against the judgements of common sense.

The most viable reply to the dicephalic twinning problem is to say that the two persons can be individuated on the basis of particular forms rather than understanding the form as only the nature of a kind (Ainsworth, 3). Universal forms suffice for making sense of how an individual remains the same over time, but, as

the dicephalic twinning case shows, it is necessary to invoke particular forms to individuate members of the same category in instances where matter is insufficient (Ibid, 3). As for the thought experiment about an individual being composed of the same matter as another at an earlier point, position in time and the continuity of the underlying substance can be used to make sense of the differentiation.

Conclusion

This analysis of hylomorphism has shown that it is the most viable soul-based theory of personal identity and has certain advantages over strict materialist theories. It can adequately address the interaction and divided mind problems associated with Cartesian dualism as well as the brain transplant and thinking animal problems linked to animalism and the embodied mind account respectively. However, the challenge of dicephalic twinning poses a serious challenge to the theory.

Making sense of the dicephalic twinning problem by positing the existence of particular forms opens the door to charges of circularity (Ainsworth, 3). If the grounds for distinguishing persons is their particular forms and the grounds for distinguishing particular forms is the persons who have them, the question is being begged. It is also open to debate whether positing particular forms is a faithful interpretation of Aristotle and his medieval commentators.

A possible escape is to conclude that there is indeed only one being like us despite there existing two centers of consciousness with their own mental states. In other words, there is only one composite of organic matter and intellective soul even though two subjects are present. If what matters is not numerical identity, then it is not unreasonable to think that there could be two sources of value for a single being like us. This response is not definitive since there will likely remain a strong intuition among some people that the two are different individuals, so the question is still open as to whether hylomorphism can survive the dicephalic twinning argument, and a future investigation into hylomorphism and standards of death could show whether this is so by looking at whether it is possible for one of the twins to be dead while the other lives thereby offering a means to determine whether there are two individuals.

Because it can answer some of the more challenging objections to other theories of numerical identity, the hylomorphic theory should be examined alongside the others in investigations into the subject. Additionally, its ethical implications would make for a valuable study. If there is merit to the hylomorphic theory as an account of numerical identity and numerical identity has implications for questions about ethics, it is surely worthwhile for there to be an exploration of the consequences of the theory for ethics.

Works Cited

- Aristotle, Book 2 in *De Anima* (Indianapolis, Indiana: Hackett Publishing Company, 2017). Translated by C.D.C Reeve
- David DeGrazia, *Human Identity and Bioethics* (New York, NY: Cambridge University Press, 2005)
- Derek Parfit, *Reasons and Persons* (Oxford: Oxford University Press, 1984)
- Henry Veatch, Chapter 3: “The Nature of Things Animate and Things Human” in *Aristotle: A Modern Appreciation* (Bloomington, Indiana: Indiana University Press, 1974)
- Jeff McMahan, *The Ethics of Killing* (New York, NY: Oxford University Press, 2002)
- John Marenbon, Chapter 7: “Mind, Body, and Mortality” in *Medieval Philosophy: A Very Short Introduction* (Oxford: Oxford University Press, 2016)
- Keith Allan Hess, Chapter 2: “Hylomorphic Animalism” in “Matter and Form: Towards an Animalist Conception of Personal Identity,” in *UC Santa Barbara Electronic Theses and Dissertations*, (<https://escholarship.org/uc/item/2hk26956>)
- Marya Schechtman, *Staying Alive* (Oxford: Oxford University Press, 2014)
- Thomas Ainsworth, “Form vs Matter,” in *Stanford Encyclopedia of Philosophy* (<https://plato.stanford.edu/entries/form-matter/>)

Does the Partly-Social Narrative Account of Identity Resuscitate Person Essentialism in the Face of Criticism?

Mia Ehudin, *George Washington University*

The theory of person essentialism posits that being a person is essential to our existence. However, under a traditional, Lockean sense of personhood, this theory seems to exclude many human beings from having moral status. Among them are infants, those with severe mental disabilities, and those in the late stages of dementia. In this paper, I aim to prove that the theory of person essentialism may be resuscitated through a partly-social narrative account of identity.

Introduction

When philosophers ask the question “What are we?”, they are in search of what we *essentially* are, and possibly that which sets us apart from other living beings. The way people answer this question proves to be of paramount importance when it comes to how we treat others and what matters in our survival. The debate of what we essentially are is ages old, and those in pursuit of the answer often fall into one of a few categories. Some believe that we are essentially organisms, so the continuation of our (biological) lives is what matters to our numerical identity. Others believe that we are essentially souls that might be able to exist beyond matter or experience. Another view is that we are essentially our brains— we go where our brains go. Lastly, some believe that we are essentially persons. Person essentialism raises several questions, such as how to define ‘persons’ and who qualifies for personhood. Many of these questions arise as a result of disputes surrounding our conceptions of numerical identity. Answering these questions is of utmost importance when we assume that only persons have full moral status.

I plan to demonstrate that despite its perceived flaws, person essentialism can be resuscitated as a tenable theory of our essence through a partly-social narrative account of identity. I will first introduce the theory of person essentialism and the problem of non-paradigm humans. Next, I will reconstruct objections to person essentialism as outlined by David DeGrazia and Jeff McMahan. Later, I will present the partly-social narrative account of identity and the case in its favor as advanced by philosophers including Mary Anne Warren, Hilde Lindemann Nelson, and Marya Schechtman. Finally, I will critically evaluate possible objections to these arguments.

Person Essentialism and its Drawbacks

The thesis of person essentialism is that “any being that is ever a person cannot exist at any time without being a person” (DeGrazia, 30). In other words, being a person is essential to our existence. The human organism merely houses or constitutes the person, so for beings like us, the human organism may exist without possessing rational thought or self-awareness in the form of a ‘pre-person’ or a ‘post-person’. In this view, ‘pre-persons’ are not beings like us (they are instead a precursor to us), and ‘post-persons’ are no longer considered to be alive. There is certainly contention surrounding the definition of ‘persons’, but it does take on a fairly standard meaning in the context of the Psychological Criterion for Personal Identity, which posits that psychological continuity is essential to our continued existence. One of the earliest accepted definitions of personhood was introduced by John Locke. He describes a person as a “thinking intelligent Being, that has reason and reflection, and can consider it self as itself, the same thinking thing in different times and places” (Locke). Lynne Rudder Baker believes that personhood requires a first-person perspective, “a perspective from which one thinks of oneself as an individual facing a world, a subject distinct from everything else” (Francescotti, 44). Mary Anne Warren describes five relevant criteria for personhood: consciousness, reasoning, self-motivated activity, the capacity for complex communication, and self-awareness (Francescotti, 44). Harry Frankfurt requires self-motivated activity and freedom of the will for personhood (Francescotti, 44). When considering all of these definitions and criteria, one might notice that several groups of humans do not fit within these standards, yet denying them personhood (and thus, full moral status and rights), would be extremely controversial. Among them are infants and small children, individuals with severe mental disabilities, and those in the late stages of dementia. For the purposes of this essay, I will not discuss humans such as anencephalic infants and those in persistent vegetative states, as they lack any mental activity altogether. Some believe that this ‘problem of non-paradigm humans’ is best solved by putting them in a category of ‘borderline persons’. However, it becomes difficult to conceptualize a way to decide who qualifies in a non-arbitrary way when considering that phenomena such as self-awareness and free will are nearly impossible to measure.

Challenges to Person Essentialism

I have already established that a problem with person essentialism under a Lockean definition of personhood is that it excludes many sentient beings, thus putting their moral status into question. In this section, I will explore some of the objections to the implications of this view. David DeGrazia outlines an objection to

person essentialism entitled the Fetus Problem. To summarize, if persons are defined as beings with psychological capacities and fetuses do not fall within this category, then how may we describe the transition from fetus to person (DeGrazia, 31)? There are three possibilities. The first is that the fetal predecessor to the person has died, but this is implausible, as there would be a death associated with every person that comes into existence. The second is that the fetus simply disappears without dying. This is also implausible, as there would be no explanation of how the succeeding person would come into being. Matter does not simply spring from nothing. The third possibility is that the existence of the fetal predecessor and the person overlap, but this would require two distinct beings to be associated with one body at once. It is true that an individual can have two differently described beings associated with them, like how I am both a daughter and a sister; however, the fetus and the person are two different substances (DeGrazia, 32). If we determine that the transition from 'pre-person' to 'person' occurs in early childhood, these three possibilities begin to seem even more absurd.

Jeff McMahan also discusses the permissibility of infanticide as a possible implication of person essentialism if we assume that only persons have a right to life. Nowadays, abortion is becoming an increasingly accepted practice, but a majority of people find infanticide to be morally objectionable. Under person essentialism, the moral status of the fetus is almost identical to an infant, as both lack rational self-agency and a first-person perspective. The low moral status of infants has rendered infanticide a historically common act of convenience (Newman). Even in some modern societies, selective infanticide occurs in extreme cases. Sometimes, infants who are born with Down's Syndrome require throat surgery in order to swallow properly. In the past, some parents made the decision to let their children die by opting out of this surgery and allowing them to starve, as their lives were deemed to not be worth living and their personhood would never be actualized (McMahan, 341). This is not quite the same as infanticide, but still quite controversial, as the infant is denied a chance at life. This is why the practice was banned by the Reagan Administration. McMahan objects to this practice, stating that from birth, infants engage with those around them and form close bonds and relationships. While much is dependent on how others perceive the value of their lives, it seems morally objectionable for a third party to make such a decision so early on.

Another objection to person essentialism has to do with what happens at the end of life, when the 'self' begins to wither away. Under a view of person essentialism, a patient at the onset of dementia and the same patient in a later stage of the disease would not be categorized as the same individual (McMahan, 493). Instead, the patient would be categorized as a 'sentient nonperson', and thus no

longer have moral status. According to DeGrazia, this is “implausible and morally outrageous” (DeGrazia, 127). A person who has just been diagnosed with Alzheimer’s would probably not feel relief that the future experiences and sensations of their body would be experienced by an individual of a different kind (McMahan, 47). Instead, we instinctively fear becoming deeply demented and what happens to us as a result. It is true that we have no way of knowing how much we will be able to experience during the ends of our lives, but many people intuitively believe that it will be their own consciousness that is experiencing it.

Case in Favor of the Partly-Social Narrative Account

The answer to the problem of non-paradigm humans in person essentialism is a partly-social narrative account of personhood. Narrative identity may be described as one’s life experiences woven together through self-narration into an integrated and unified self-conception that evolves over time. While non-paradigm humans may not be able to contribute to their own narrative identities, I shall argue that the people around them can. Many philosophers have attempted to elaborate on this topic using different terminology, but their views share the fact that the basis for personhood involves one’s relationships and interactions with others rather than only the traditional criteria.

Mary Anne Warren has outlined what she calls a ‘Principle of Full Moral Status’. She uses the term ‘moral status’ rather than ‘personhood’, but for the purposes of this paper we are assuming that only persons have full moral status.

Principle of Full Moral Status: “The degree to which an entity E possesses moral status is proportional to the degree to which E possesses morally relevant properties until a threshold degree of morally relevant properties possession is reached, whereupon the degree to which E possesses morally relevant properties continues to increase, but the degree to which E possesses moral status remains the same” (Lovering, 509).

In simpler terms, there is a plethora of potential morally relevant properties that entities may possess. However, once they reach a certain level, adding on more morally relevant properties does not affect one’s moral status. At first, it may not be clear how the Principle of Full Moral Status relates to the problem of non-paradigm humans; however, Warren later writes that to have moral status “is to be an entity toward which moral agents have, or can have, moral obligations” (Lovering, 510). Included under this umbrella are beings for whom “normal adult humans” care (Lovering, 518). It is nearly impossible to argue that beings who fall under the

traditional definition of personhood do not have obligations towards infants, mentally disabled people, or the deeply demented. In fact, many individuals' entire livelihoods are based on being caregivers for those who do not possess the ability to care for themselves. Infants at birth can rarely ever survive without their mothers. Without proper acknowledgment of these moral entitlements, human lives take on a form that is not quite worth living (Lovering, 515).

The idea of the caregiver serving as the narrative author for non-paradigm humans is further explored by Hilde Lindemann Nelson. Lindemann Nelson ponders whether it is possible for personhood to be granted to someone who is seriously ill or disabled and unable to construct their own narrative identity. She touches on her personal experience with her little sister Carla, who was born with hydrocephalus, a condition where an excess of spinal fluid in the brain causes a pressure buildup that interferes with cerebral function. Carla was unable to perform basic functions like lifting her head, sitting up, speaking, swallowing, or grabbing objects. According to doctors, her care would become increasingly difficult as she grew, but sadly, she passed away at eighteen months (Lindemann Nelson, 30). Through reflecting on Carla's short life, Lindemann Nelson considers the possibility that individuals can be brought into personhood through how they are treated. Even in paradigm cases, humans use socially shared narratives to create a sense of group identity, and one of the most common socially shared narratives is that of the family (Lindemann Nelson, 31). Before birth, identity-formation can begin through choosing a name for the child, performing proper prenatal care, and even talking to or playing music for the fetus. Paradigm children are able to accommodate themselves to their families by acting or speaking, but any type of child may be accommodated by their family. Although Carla was mostly unable to express herself, her family was able to perceive her in a certain way and treat her accordingly, further molding her into the person that they saw (Lindemann Nelson, 32). However, Lindemann Nelson questions whether the personhood granted to Carla was arbitrary or honorific, given that she did not fit most of the traditional criteria. She turns to Carl Elliott, who writes that treating someone as a person merely involves having a certain attitude towards them (Lindemann Nelson, 33). Human interaction involves responding to the perceived psychological states of others; something one would not be able to do with a plant or a piece of furniture. Even though Carla lacked language, she was able to convey her desires and intentions through other modes such as expressions, gestures, and sounds (Lindemann Nelson, 34). It is also certain that Carla had the ability to be comforted, experience sensations, and focus her attention (Lindemann Nelson, 35). Even though Carla lived a very short life, she is remembered as a person through the experiences that her family had with her and the stories that they tell about them. In

this view, it follows naturally that Carla had moral status. Her rescue from a burning building would take priority over that of a cat or dog. She was fed and clothed and given a name. Her family had an obligation to care for her as a person, and they treated her as such.

This phenomenon of caregivers providing the narrative that is necessary for personhood may also be seen at the end of life in those in the late stages of dementia. Cases of humans with dementia are characteristically different from infants or those who are severely mentally disabled, as these individuals were at once considered to be persons. In fact, dementia is often viewed as the ‘loss of the self’ or ‘loss of the person’ (Higgs & Gilleard, 773). This is perhaps because paradigm humans often assume that their inner worlds are “so bizarre and disordered that [they] could not be discussed in the categories of rational discourse” (Kitwood, 13). This assumption is not baseless, as these individuals often have a difficult time recognizing others, speaking, and performing basic daily functions. However, observers have no clear way of seeing into their mental lives. Like Carla, humans with dementia have basic needs that are attended to by others, such as comfort, attachment, inclusion, occupation, and identity (Kitwood, 19). It may even be argued that some of these needs are of a higher faculty than those of an infant. Humans with dementia are sent to facilities that are specialized for their treatment and are often visited by family members and others who care about them. When my grandmother was hospitalized for a stroke shortly before her death, I would paint her nails the way she liked them and play her favorite music. Although there was little chance of her mentally processing my actions, it can be said that I was promoting her personhood through continuing and honoring her narrative identity.

Perhaps the most comprehensive description of the partly-social narrative account of identity was established by Marya Schechtman. She begins by advocating for a narrative-self construction view of identity, which involves conceptualizing our lives in a holistic way through a first-person perspective rather than seeking relations between time-slices (Schechtman, 100). Rather than requiring psychological continuity in order to be the same person, an attribute that people may lose over time, this view requires a diachronic unity. Diachronic unity implies that a person’s experiences must be intertwined, but not necessarily sequentially linked. Schechtman also addresses the problem of those with no first-person perspective by taking on a view similar to Lindemann Nelson’s; that they can be given an identity through narratives created by others (Schechtman, 104). After establishing the basis for the identity of non-paradigm humans, she then turns to the problem of personhood by advocating for a ‘person life view’. Schechtman asserts that “To be a person is to live a ‘person life’; persons are individuated by individuating person lives; and the duration of a single person is determined by the

duration of a single person life” (Schechtman, 110). The ‘person life’ that Schechtman describes typically includes characteristics like a period of social dependence, eventual maturity through substantial growth in cognitive capacities, possible loss of capacities through aging, and community systems like friends and family (Schechtman, 112). Those who live ‘person lives’ also often fill a certain societal role and adhere to cultural norms (Schechtman, 113). But what about non-paradigm humans and those who are on atypical development trajectories? Schechtman argues that despite a lack of ‘standard’ capacities, these individuals are still clearly seen as persons, and thus afforded a place in ‘person-space’ (Schechtman, 120). Many elements of a ‘person life’, such as being named, being registered with a Social Security number, and wearing clothes do not require mature capacities or even the prognosis of them. Under the person-life view, individuals are not granted personhood because of their capacities, but because they fit in with and are accepted by the humans around them. When we treat individuals under the pretense that they are the same kinds of beings as us, this dramatically shapes the way in which we interact with them.

Critical Evaluation and Conclusion

Now that I have established the case in favor of the partly-social narrative account of identity, I will outline and evaluate some possible objections to it. The first objection is that affording individuals ‘honorary’ personhood renders the distinction between persons and non-persons completely arbitrary. If infants like Carla are considered persons, what’s to stop us from deciding that the family cat or dog is a person too (Lindemann Nelson, 35)? Pets experience ‘person lives’ in that we give them names, feed them, and form bonds with them. Some owners even dress them in little clothes, let them sleep in their beds, and buy them insurance. Lindemann Nelson responds by stating that animals who we keep as pets simply do not share our form of life; they are never realistically deemed to be of the same kind as humans. Similarly, Schechtman argues that the basic expectations for humans and animals are completely different (Schechtman, 121). If my cat does not end up attending GWU as my legacy, I will not be all that disappointed. Instead, I praise her for doing cat things, like using her litter box or chasing her tail. The way that I view and treat her is fundamentally different from how I would treat a human, regardless of their mental capacity.

Another possible objection to a partly-social narrative account of personhood is that some non-paradigm humans may not be treated like persons or have people that care about them. If we operate under a view like Lindemann Nelson’s this argument may hold true, however, it can be rebutted using Schechtman’s Person Life View. Consider a person who lives alone, does

not go to work, and orders in all their food. They do not have any friends, and do not interact with their neighbors or others in their community. Would they still be considered a person? Technically, yes, as there are aspects of a person-life that do not necessarily involve being cared for by others. Some examples are having a registered birth, having a Social Security number, and being named. If this person did go out into the world, people would treat them as one of their kind. As mentioned before by Schechtman, many of the aspects of personhood do not require the possession of mature capacities. If a non-paradigm human truly lacked all aspects of a person life, it would be an extreme case and regarded as an exception.

Through critically evaluating these objections to the partly-social narrative account of identity, it remains clear that this view is our best hope in resuscitating person essentialism, and thus granting moral status and rights to non-paradigm humans. Person essentialism is an important view to legitimize, as it answers the problem of what makes us different from other living beings while still being based in biology. The partly-social narrative account of identity grants personhood to non-paradigm humans by allowing those who care about them to assist in the construction of their narrative identities. It also allows those who do not have supportive social units to participate in personhood simply by sharing aspects of a 'person life'. In a society where those who lack certain mature capacities are heavily discriminated against, reframing the way we view 'persons' and adopting more lenient and universal standards has paramount implications in the broader scope of human rights, but most importantly the way we treat others every day.

Works Cited

- DeGrazia, David. *Human Identity and Bioethics*, Cambridge University Press, 2009.
- Francescotti, Roberto. "Person Essentialism and the Problem of Vagueness." *Proceedings Metaphysics 2006: 3rd World Conference*, 2006, pp. 54–60., [http://romemetaphysics.org/actas/2006_Proceedings Metaphysics.pdf](http://romemetaphysics.org/actas/2006_Proceedings_Metaphysics.pdf).
- Higgs, Paul, and Chris Gilleard. "Interrogating Personhood and Dementia." *Aging and Mental Health*, vol. 20, no. 8, 2015, pp. 773–780., doi:10.1080/13607863.2015.1118012.
- Kitwood, T. "The Experience of Dementia." *Aging and Mental Health*, vol. 1, no. 1, 1997, pp. 13–22., doi: 10.1080/13607869757344.
- Lindemann Nelson, Hilde. "What Child Is This?" *Hastings Center Report*, vol. 32, no. 6, 6 Mar. 2012. *Wiley Online Library*, doi:<https://doi.org/10.2307/3528131>.
- Locke, John. "Of Identity and Diversity." *An Essay Concerning Human Understanding*, I, 2004, <http://www.gutenberg.org/cache/epub/10615/pg10615.html>.
- Lovering, Robert P. (2004). Mary Anne Warren on "Full" Moral Status. *Southern Journal of Philosophy* 42 (4):509-30. PhilArchive copy v1: <https://philarchive.org/archive/LOVMAWv1>
- McMahan, Jeff. *The Ethics of Killing: Problems at the Margins of Life*. Oxford University Press, 2002.
- Newman, Sandra. *The Roots of Infanticide Run Deep, and Begin with Poverty*. Edited by Marina Benjamin, 27 Nov. 2017, aeon.co/essays/the-roots-of-infanticide-run-deep-and-begin-with-poverty.
- Schechtman, Marya. *Staying Alive: Personal Identity, Practical Concerns, and the Unity of a Life*. Oxford University Press, 2017.

Buddhism, Stoicism, and Naturalistic Ethics

Trevor Woodward, *University of Idaho*

There has been renewed interest in Stoicism and Buddhism alike in both academic and popular culture. In this paper, I investigate the substance of a seeming similarity between Buddhism and Stoicism: the necessary connection between suffering and desire. I argue that this similarity is superficial. However, the investigation reveals another, deeper similarity that I explicate and defend; namely, that both Stoicism and Buddhism are ‘strong’ forms of naturalistic ethics, whereas Aristotle’s account of Virtue Ethics is a ‘weak’ form of naturalistic ethics.

Introduction

There has been renewed interest in Stoicism and Buddhism alike in both academic and popular culture. Moreover, as is well documented, the two philosophies appear remarkably similar. For example, as outlined by Pigliucci (2020), Stoicism and Buddhism (in virtually all construals) agree with respect to their doctrine on anger; namely, anger and all its associated emotions are categorically bad. That said, the respective philosophies rest on nearly incommensurate metaphysics and are motivated by different theoretical desiderata. In a word, it might appear that most apparent similarities between the schools of thought are superficial as opposed to being of substance. In this paper, I will investigate the substance of a seeming similarity between Buddhism and Stoicism. Pigliucci contends that, in some sense, the two philosophies are aimed at the elimination of suffering. I will take this line, but strengthen the claim to what follows: Both Buddhism and Stoicism claim a necessary connection between suffering and desire.¹³ I then compare the respective justifications for this claim, and argue that the similarity is, indeed, superficial. However, the investigation reveals another similarity between Buddhism and Stoicism that I will argue for in this paper: both philosophies are a ‘strong’ form of naturalistic ethics. More precisely: both ideologies call upon the agent to qualitatively identify their moral psychology with the descriptive nature of the absolute.

I proceed in four steps. First, I introduce a superficial similarity between Roman Stoicism and Zen Buddhism that provokes further exploration; namely, the connection between suffering and desire. Then, I argue that the respective theoretical justifications for this connection are incomparable, and thus that the similarity is mere coincidence. Next, I argue that Hellenistic Stoicism and Zen

¹³ I do not commit to what sort of necessary connection this might be. It seems plausible that it might be one of supervenience, i.e., changes in one’s suffering are possible only given changes in one’s desires; but this is immaterial to the discussion.

Buddhism are instances of naturalistic ethics and demonstrate their similarity with respect to moral psychology. Finally, I juxtapose the two schools of thought with Aristotelian Virtue Ethics—a *weak* form of naturalistic ethics—in order to motivate the significance of the similarity.

Roman Stoicism

Handbook by Epictetus is, perhaps, the most widely acknowledged text on Roman Stoicism. In it, Epictetus begins by drawing a central distinction: that which is in one's power by nature (one's actions), and that which is beyond one's power.

In our power are opinion ([Greek: *hupolaepsis*]), movement towards a thing ([Greek: *hormae*]), desire, aversion ([Greek: *echchlis*]), turning from a thing; and in a word, whatever are our acts. Not in our power are the body, property, reputation, offices (magisterial power), and in a word, whatever are not our own acts (Enchiridion I).

For Epictetus, the Stoic lifestyle is one in which you exercise careful epistemic consideration of these categories and regulate your psychology by their dictates. For the Stoic, to pursue or avoid something beyond one's power—bodily pleasure, reputation, disease—is a simple category error. It is to mistake one's domain of personal concern to contain considerations that, by the simple fact of nature, do not belong there. Therefore, it follows to direct that which is in one's control—opinion, desire, aversion—toward only the things within the same category. Epictetus provides the example of a Roman bathhouse:

If you are going to bathe, place before yourself what happens in the bath; some splashing the water, others pushing against one another, others abusing one another, and some stealing; and thus with more safety you will undertake the matter, if you say to yourself, I now intend to bathe, and to maintain my will in a manner conformable to nature (Enchiridion IV).

There is a category error being implicated here. For many a Roman, to desire to bathe is to set as their ends an implausible goal: a perfectly quiet and uninterrupted experience. For the enlightened Stoic, the pushing and splashing are not deviations from the goal; they are simply entailments of bathing that ought to bear on the agent's intentions from the start. If one commits to creating this rationally informed species of intentions, then they will never be unsatisfied by coming to the bathhouse

or engaging in any other plausibly unsatisfying venture. Going to the bathhouse and wishing for peace and quiet is, in a word, irrational.

Recall that the body, for Epictetus, is also beyond one's control. As Epictetus elaborates:

Disease is an impediment to the body, but not to the will, unless the will itself chooses. Lameness is an impediment to the leg, but not to the will. And add this reflection on the occasion of everything that happens; for you will find it an impediment to something else, but not to yourself (Enchiridion IX).

Simply put, the suffering brought on by disease and injury is elective. It is possible, through Stoic training, to condition a psychology that produces no desires or aversions on the basis of bodily harm. To grip one's broken leg in agony is the product of a less-than-virtuous moral psychology. In other words, if one *suffers*, it is because one has chosen, or perhaps *desires*, to suffer. Further, if one adheres fully to the Stoic doctrine, they thus have no desires with respect to the body or outside world, and therefore, *cannot* suffer in most any of the ways thought typically, *e.g.*, physical pain, loss of a loved one, longing for a loved one, etc. Hence, there is a necessary connection between suffering and desire.¹⁴ Similarly, Epictetus views a fear of death (an aversion toward death) to be elective: "death is nothing terrible, for if it were it would have seemed so to Socrates; for the opinion about death that it is terrible, is the terrible thing" (Enchiridion V). The opinion described here is flawed in that it stretches across categories. Opinions may only concern themselves with actions: those things that we have power over. Hence, the proper opinion of death is the opinion of nature: that it is beyond one's control and thus beyond their scope of concern.

Zen Buddhism

According to D.T. Suzuki, a prominent scholar of Zen Buddhism, the central doctrine of Buddhist thought is *anatman*: the doctrine of non-ego. The argument has two key parts:

(1) That all things are transient as they are composites and go on disintegrating all the time, that there is nothing permanent; and (2) there is

¹⁴ Note that suffering that does not essentially involve one's body or the external world may still be possible for the adherent Stoic. I hold that most all paradigmatic examples of suffering do involve such things, and hence, that the virtuous Stoic does not suffer in any robust sense.

therefore nothing worth clinging to in this world where every one of us is made to undergo all kinds of sorrow and suffering (Suzuki 33).

The second part of the argument, that of suffering, is consonant with Roman Stoicism. It is this superficial similarity between these schools of thought—the necessary connection between suffering and desire (in the quote: the connection between “clinging” to things and “sorrow and suffering”)—that first inspired this investigation. The first point, however, is built upon far different theoretical grounds. For Epictetus, the theoretical constitutes the inescapable categories of nature. For the Buddhist, the theoretical permits of no such categories whatsoever.

Zen Buddhism places a heavy emphasis on the cognition, as opposed to the metaphysical constitution, of its central truths. To become enlightened requires a special mode of apprehension through the *prajne*-eye: “a special kind of intuition enabling us to penetrate right into the bedrock of Reality itself” (Suzuki 34). Upon utilizing this mode of cognition, it will become apparent to the agent that the ego, or *atman*, is imperceptible: “As long as our intellectually analytic eye is hotly pursuing the shadow of reality by dichotomising it, there will be no silent stillness of absolute identity.” In other words, the *prajne*-eye’s epistemic output goes beyond the realm of analysis and categorization entirely. To create a boundary between oneself and the external world—the key prescription of Stoicism—is to dichotomize it, and thus to fail epistemically with respect to Buddhist thought. D.T. Suzuki puts it in western terms: “Personal experience is thus seen to be the foundation of Buddhist philosophy. In this sense Buddhism is radical empiricism or experientialism, whatever dialectic later developed to probe the meaning of enlightenment-experience” (Suzuki 42).

It may appear that to compare Buddhism to Stoicism is to square circles. If theoretical Buddhism transcends dichotomies, and thus the realm of language, then how could one ever compare it fruitfully with Stoicism—a school of thought based upon rational discourse? The similarity with respect to suffering and desire, then, may best be conceived as mere coincidence. However, I would offer that while the descriptive elements of Buddhism and Stoicism are incommensurable, the prescriptive elements are surprisingly similar, and worth further investigation. More specifically, both Buddhism and Stoicism call upon the agent to qualitatively identify their moral psychology with the descriptive nature of the absolute.

That said, it is important to emphasize here that I intend a distinct line of inquiry from the “mysticism” question; that is, whether Buddhism (or Stoicism) calls upon the adherent to absolve themselves *into* the absolute. The fact that one’s psychology is qualitatively identical with the absolute does not entail that it simply *is* the absolute, nor does it entail the opposite: that the adherent and the absolute are

metaphysically distinct. The fact that a and b are red triangles (qualitatively identical) does not entail that $a=b$, nor does it entail $a\neq b$. As such, the soundness of my argument does not necessarily bear on the mysticism debate.

Hellenistic Stoicism: The Logos

In order to continue this investigation, it is necessary to turn toward a more robust account of the Stoic cosmos, or *logos*, in order to understand its connection to Stoic ethics. The starting assumption of Stoic physics is that the cosmos—the whole of existence—is (1) perfectly rational and (2) perfectly benevolent (Cooper 267-268). The conjunction of these assumptions is key. Stoic rationality is teleological—all parts of the universe (people, trees, rocks, etc.) are rational insofar as they serve their constitutive purpose. And given that the universe is benevolent, this purpose is unequivocally good. That said, to know what these ends are requires further understanding of Stoic physics. The building blocks of the Stoic cosmos are two forms of rationally imbued matter: active and passive (Inwood 133-134). Active matter is the cause of change, or mind. Passive matter is statically rational—for example, bones and sinews are constituted such that they may be controlled by the mind. The interplay of static and passive matter forms the basis of the Stoic *animal*: an entity constituted by a leading part (mind, active matter) and its entailing passive matter (bones, sinews, muscle, etc.) (Inwood 134). Quite fascinatingly, the Stoic “animal” does not have the negative valence that “animal” would appear to have in other Greek philosophies. In fact, God is an animal according to the Stoics. For Zeno and Chyrippus—and possibly other Stoics—the substance of God is the entirety of the cosmos (Inwood 135). Its leading part, which is alive and performs sense-perception, is the heavens; I understand this to mean the stars, sky, etc. which lie beyond the earth.

To be rational, for the aspiring Stoic, is to understand one’s role in a teleological chain of instrumental actions done toward the ultimate end: that of God, or the cosmos. Given that humans, by definition, comprise God’s passive matter (the earth), it follows that their ends are analogous to that of bones and sinews. That is, humans are to passively accept and cooperate with the dictates of the leading part of the animal in which they are constituted. Simply put, the ends of human behavior are the cosmos, or rationality in itself. Cicero provides an interesting analogy. A navigator—the partially rational agent—may act upon reasons in bringing a ship to port, but their goal is, of course, to bring the ship to port. Given that, for reasons beyond the agent’s control, the ship may fail to reach port (perhaps a mighty storm knocks it off course), it follows that to set one’s ends as “to bring the ship to port” is plausibly to act in conflict with fate, or the will of the cosmos. In contrast, a dancer acts upon reasons—the commands of the choreographer—but does so for the sake

of the action itself; their goal is not the art, but simply to act artistically. This fully rational behavior is to act in agreement with the will of the cosmos (Cooper 270).

What Epictetus provides in *Handbook* is a prescriptive account of moral psychology—it tells one not explicitly how to act, but rather provides the cognitive processes that the agent ought to employ when engaging in ethical deliberation. And these cognitive processes are, in a word, *rationality*: the same principle that governs the Stoic cosmos and interpenetrates its matter. To prevent category errors, and thus to prevent one's ends from being subject to exogenous forces, is simply to ensure that one acts in agreement with the will of nature. Since the ideal of Stoic moral psychology is perfectly rational behavior, and the *absolute* (or cosmos) is rationality in itself, it follows that Stoic ethics calls upon the agent to qualitatively identify their moral psychology with the descriptive nature of the absolute. In the next section, I will argue that Buddhism is a form of naturalistic ethics in precisely the same way.

Buddhism as Naturalistic Ethics

The Stoic absolute is relational—the ultimate nature is that which unites all things: the rational, omnibenevolent principle of *logos*. Zen Buddhism is best understood as a-relational; where the Stoic absolute is a relational principle, the Buddhist absolute is the principle of no relations—that all relations are mental constructions that stand in the way of ultimate truth. Zen Buddhist scholar Kitaro Nishida provides a comprehensive account of Buddhist cognition in *An Inquiry Into the Good*. This account will be necessary to understand the Buddhist *absolute*.

For Nishida, to have pure experience is to view facts just as they are and relinquish all mental fabrications (Nishida 3). The contrast class is Wilhelm Wundt's notion of mediate experience: “knowledge that is reasoned out discursively on the basis of [pure] experience” (Nishida 4). Phenomena are pure experience if and only if they arise from “the strict unity of concrete consciousness” (Nishida 6). Importantly, this unity arises from the systematic nature of consciousness:

For example, when we think we have perceived at a glance the entirety of a thing, careful investigation will reveal that attention shifted automatically through eye movement, enabling us to know the whole. Such systematic development is the original form of consciousness, and as long as the unity maintains itself and consciousness develops of its own accord, we do not lose our foothold in pure experience (Nishida 7).

Simply put, pure experience is anything endemic to the consciousness-system at present. Importantly, this may include representational experience—the unifying of

phenomena by the subject itself—if and only if it occurs by its own systematic perpetuation (Nishida 7). The place where experience loses its purity, then, is where it becomes a meta-system of consciousness: where the current system of consciousness gets related to some other consciousness. When such a relation obtains, for Nishida, experience becomes *meaning* (Nishida 7).

When coupled with the Buddhist commitment to radical empiricism (as stated earlier by Suzuki), Nishida's account of cognition gives rise to a descriptive account of the absolute. Simply put, radical empiricism introduces a normative framework through which to view the findings of Nishida's theory of cognition; namely, that pure experience is to be prized as absolute truth. In contrast, *mediated* experience can at best count as *discriminatory* knowledge, which for Zen philosopher Shigenori Nagatomo is inherently “delusory/illusory in nature” (Nagatomo). Once this notion of the absolute as pure experience is cognized, its implications for moral psychology are precisely the same as that of Stoicism—the practicing Buddhist ought to condition a moral psychology on the basis of pure experience and only perform actions in line with its dictates. The move here is simple. Recall Suzuki's outline of *anatman*, the central doctrine of Buddhist thought:

(1) That all things are transient as they are composites and go on disintegrating all the time, that there is nothing permanent; and (2) there is therefore nothing worth clinging to in this world where every one of us is made to undergo all kinds of sorrow and suffering (Suzuki 33).

The first point is descriptive; the second point prescriptive. The “impermanence” principle is that of pure experience: an empirical account of the absolute in which all “composites” count as illusory. The cessation of suffering—the central prescription of Buddhist thought—is nothing more than an *entailment* of that descriptive account. The statement that “X is painful” or “X is desirable” is necessarily an instance of mediate experience; it can only be constituted by uniting a present system of consciousness with some other consciousness (presumably a normative one). Hence, the best status a statement such as “X is desirable” may achieve under the impermanence principle is that of discriminatory knowledge. That is, a judgment that something is desirable is illusory by its very nature as a “composite”. As such, pure experience (the absolute) does not permit desire—to accept the cessation of suffering as a moral imperative is nothing more than to accept pure experience as one's moral psychology. Just as the Stoic must act in agreement with the truth of the cosmos, the Buddhist must act in agreement with the truth of pure experience.

Strong vs. Weak Naturalistic Ethics

I have shown that a point of similarity between Zen Buddhist and Stoic thought is that both call upon the agent to qualitatively identify their moral psychology with the descriptive nature of the absolute. However, it remains to be seen why such a finding is interesting. In *Eudaimonism*, John Cooper defends the philosophical respectability of the Stoic conception of virtue. For reasons that are incidental to my purposes, Cooper defends the Stoics' "outside appeal" to nature in their account of ethics. Cooper does so by arguing by analogy from Aristotle's account of Virtue Ethics, which engages in precisely the sort of "outside appeal" in question at *Nicomachean Ethics I 7*. Here's Cooper's gloss of the argument:

All animals (the same goes for plants) have a natural good, consisting in the particular kind of flourishing life that their specific natures fit them for, and this life is the life of a thing that is an excellent specimen of its kind—our common and obvious practice of speaking of some things as being good and others bad for specific kinds of living things rests upon this fact, and brings it to light. The application of this general scheme to human beings produces the result that our good consists in a life directed by the use of our rational powers, when they are structured by their excellences, that is, by the virtues; that is what, in our case, an excellent specimen of our kind leading a flourishing life amounts to (Cooper 266).

Cooper goes on to stress that Aristotle's argument above connects *non-ethical* patterns of thought, i.e., Aristotle's account of natural teleology—the goal-orientedness of nature—to explicitly *ethical* patterns of thought, i.e., to one's justification of human virtue. If one is committed to his non-ethical conception of nature, they have good reason to extend the type of thinking therein to the realm of ethics. As such, on Cooper's account, Aristotle's appeal to nature does "real and valuable work" for his central argument for virtue ethics (ibid). Cooper's next move is simply to argue that the Stoic account of virtue engages in precisely the same rhetorical move. As I argued earlier, the Stoic's prescriptive account of moral psychology—as outlined in Epictetus' *Handbook*—is justified by an appeal to nature; an appeal to the rational, omnibenevolent principle of *logos* that interpenetrates all things. However, the Stoics seems to take the rhetorical project one step further. In considering Aristotle's argument above, Cooper mentions the following (my italics):

It is a further question, on which this argument is silent, whether the *specific* content, the *specific* forms of action and *specific* concerns, of the virtues that

we already accept when we enter philosophical ethics can be given some confirmation similarly from the outside. I believe it does not occur to Aristotle to think so (ibid).

Aristotle's appeal to nature is roughly meta-ethical. In other words, it gives support to the most general aspect of his system for practical reasoning and ethical deliberation; namely, to its teleological (goal-oriented) character. In sharp contrast, the Stoics go beyond the general aspects of their ethics to ground their account of moral psychology in a non-ethical account of nature. To put the point more boldly: there is an *isomorphism* between Stoic ethics and physics (their account of nature) such that the latter determines the moral status of *particulars* (i.e., specific actions, specific beliefs, etc.). This fact can be demonstrated as follows. As a premise, take (*): *the moral status of particulars are within the output of a function of a virtuous moral psychology*.¹⁵ This feature of moral psychologies is trivial. As a minimal criterion of adequacy, any account of moral psychology (*qua* an account of moral psychology) must be able to take certain information about agents and the world as input and then return as the output a set of particulars (specific actions, beliefs) that are deemed moral by a normal agent employing that moral psychology.¹⁶ For example, in a *descriptive* account of moral psychology, one must describe a psychology that can explain how actual agents decide what they think is moral. This is just to say that given a concrete case, one can input certain information into the account of moral psychology that the agents in that case had access to and correctly predict the set of particulars that they in fact deemed moral. In a *prescriptive* account of moral psychology (e.g., that of Stoicism and Buddhism), one must describe a *virtuous* psychology that, if possessed and employed by a normal agent, would take in the information accessible to that agent and return a set of particulars that the agent deems moral and are moral *as such*. Given (*), it follows, by the substitutivity of identicals from my prior argument (Section III), that the moral status of particulars (e.g., whether one ought to go to the Roman Bathhouse) is within the output of a function of Stoic physics (the account of the descriptive nature of the absolute). That is, since the Stoics' account of moral psychology and account of nature (physics) are qualitatively identical (i.e., both are *rationality* or *logos*), both have—as a function they can perform—the capacity to fix the moral status of particulars. In sharp contrast, for Cooper, Aristotle does *not* identify his account of virtuous moral psychology with his account of nature. Aristotelian physics could potentially rule out certain particulars as immoral, but it does not

¹⁵ 'Function' is to be understood loosely in the mathematical sense; each input is mapped to exactly one output.

¹⁶ Here it is assumed that the agent is not self-deceived or otherwise rationally deficient.

purport to fix the moral status of specific actions, beliefs, etc. On Aristotle's view, there is more to the moral status of particulars than is to be found in an account of nature. For the Stoics, the opposite is the case—an account of nature can, and in fact should, be able to fix the moral status of particulars.

It is in this sense that Aristotelian virtue ethics is a *weak* form of naturalistic ethics, while Stoicism, in contrast, is a *strong* form of naturalistic ethics. Given my prior argument in section IV, it can be shown that Buddhism is also a *strong* form of naturalistic ethics. Again, take (*) as a premise. Recall that the Zen Buddhist account of nature is pure experience and that a virtuous Buddhist moral psychology is qualitatively identical to that account. Hence, by again applying the substitutivity of identicals to (*), it follows that the moral status of particulars are within the output of a function of the Buddhist account of nature (pure experience). As an example case, take the moral status of gripping one's broken leg in agony. Since pain and attachment are illusory, such a particular is not moral for the adherent Buddhist. However, the conclusion that suffering is illusory is as much a prescription of *anatman* as it is an entailment of pure experience as such. That is, the moral status of suffering from one's broken leg (a particular) is fixed by the account of nature (*viz.* pure experience) in the same way that it is fixed by the account of moral psychology (*viz.* *anatman*). The Stoic account of nature reaches inside the ethical domain to ground the moral status of particulars—the Buddhist account of nature does just the same.

That Buddhism and Stoicism are both strong forms of naturalistic ethics—despite having developed in near perfect spatiotemporal isolation from one another—is quite remarkable.¹⁷ Such a peculiar similarity at this level of abstraction seems hardly trivial. Moreover, the juxtaposition to Aristotle, whose naturalistic account of ethics is somewhat hegemonic in the literature, suggests that a proper analysis of the inner workings of virtue ethics ought not confine itself to the traditional canon of western philosophy.

¹⁷ See "*Stoicism & Buddhism: Lessons, Similarities and Differences*" for evidence to support the claim that the two ideologies developed in spatiotemporal isolation.

Works Cited

- Cooper, John M. "Eudaimonism, the Appeal to Nature, and 'Moral Duty' in Stoicism." *Aristotle, Kant, and the Stoics: Rethinking Happiness and Duty*, Cambridge University Press, 2000.
- "THE ENCHEIRIDION, OR MANUAL." *The Discourses of Epictetus; with the Enchiridion and Fragments. Translated, with Notes, a Life of Epictetus, and a View of His Philosophy*, by Epictetus and George Long, G. Bell and Sons, 1906.
- Inwood, Brad, and L.P. Gerson, translators. *Hellenistic Philosophy: Introductory Readings*. Hackett Publ., 2006.
- Pigliucci, Massimo. "Stoicism and Buddhism: a Comparison." *Medium*, Medium, 16 Jan. 2020, medium.com/@MassimoPigliucci/Stoicism-and-buddhism-a-comparison-58d3e86587b.¹⁸
- "Pure Experience." *An Inquiry Into the Good*, by Kitaro Nishida and Masao Abe, Yale Univ. Press, 1992.
- Nagatomo, Shigenori, "Japanese Zen Buddhist Philosophy", *The Stanford Encyclopedia of Philosophy* (Spring 2020 Edition), Edward N. Zalta (ed.), URL = [<https://plato.stanford.edu/archives/spr2020/entries/japanese-zen/>](https://plato.stanford.edu/archives/spr2020/entries/japanese-zen/).
- "Stoicism & Buddhism: Lessons, Similarities and Differences." *Daily Stoic*, 26 Mar. 2018, [dailyStoic.com/Stoicism-buddhism/](https://dailystoic.com/stoicism-buddhism/).
- Suzuki, Daisetz Teitaro. *Mysticism: Christian and Buddhist*. Routledge, 2018.

¹⁸ I chose to use an article from a pop-culture publication in order to support my claim that there is renewed interest in *both* academic and popular culture. Massimo Pigliucci is a prominent academic and well-respected authority on Stoicism.

Being-There as Mood: Fear and Anxiety in Heidegger's *Being and Time*

Arundhati Singh, *University of Pennsylvania*

In this paper, I explore “mood” or “Stimmung” in Martin Heidegger’s Being and Time as a fundamental way of “Being-there.” I first clarify the taxonomy of the “there” – i.e., as an existential mode of being in the world rather than a geographical one – and then explain how this way of being concerned with the world relates to the concepts of attunement and mood in Heidegger’s philosophy. Then, using the examples of fear and anxiety (and how they appear in the 2011 film Melancholia), I draw a distinction between fleeting emotions – or what we generally call “moods” today – and moods in the Heideggerian sense, i.e., a way of belonging to, looking at, and experiencing the world.

When we speak of moods, what comes to mind are immediately feelings and emotions. I’m happy, I’m cheerful, I’m annoyed, I’m angry. According to Heidegger, however, a mood is much more important and primordial than a mere feeling; it is a fundamental way of looking at and belonging to the world. However, given the ambiguity of the term, and the way in which Heidegger describes moods, the difference between a primordial mood and a mere feeling or psychological state becomes blurred. In this paper, I will first delineate the Heideggerian notions of attunement and mood, and then, using the examples of fear and anxiety, draw a distinction between moods and emotions.

In order to understand the concepts of attunement and mood in Heidegger’s philosophy, it is first important to understand the context in which they are developed, namely, the existential constitution of the “there.” In §12 of *Being and Time*, Heidegger tells us that Dasein has, as an existentially¹⁹, the property of Being-in (Heidegger, *Being and Time* 79, emphasis added). Dasein exists as Being-in, in the state of Being-in-the-world. Human beings²⁰ are always in the world, absorbed by it, concerned with entities. However, we are not in it spatially²¹, as though sitting in a container; we do not occur in the world. Rather, we are in it *existentially*: in the sense that the world matters to us, we are concerned with it, involved in it. I am not just present-at-hand in the world,²² I

¹⁹ An essential feature of Da-sein (contrasted with “categories,” used for non-Dasein entities) (Blatter, Some Terminology in *Being and Time*, <https://faculty.georgetown.edu/blattnew/heid/terms.htm#disclose>).

²⁰ I’m going to employ the everyday usage of the word ‘Dasein,’ which connotes persons or humans (BT 27).

²¹ This does not mean that Dasein does not have spatiality, but that “Being-in-space” is itself founded upon Being-in-the-world generally (BT 82).

²² Although I may be taken as such. Present-at-hand = the way of being for objects, which do not have Being-in, and therefore cannot matter to each other (BT 81-82).

am in the world by way of being engaged in an activity, and being in a situation (Heidegger, BT 80-82). It is this situation that Heidegger calls the “there;” I am always Being-there (the German term *Da-sein* translates literally to “Being there,” ‘Da’ meaning ‘here’ or ‘there,’ and ‘sein’ meaning Being; BT 27). Hubert L. Dreyfus provides a helpful way of thinking about this term, as in the expression “There I was, in trouble again” (Dreyfus 164). Or “Here I am, writing my final paper at the last minute again.”²³ Once again, as with Being-in, the ‘here’ or ‘there’ is used not in the spatial sense, but in the existential one – I am not here²⁴, sitting in this room, at the geographical coordinates 16.7666° N, 3.0026° W; rather, I am here, engaged in the activity of grumpily trying to finish my paper. And because Dasein exists *as* Being-in – that is to say, it *is* what it does, it *is* the activity in which it is engaged – it *is* its “there” (BT 171). It is the “there” in which Being-in-the-world as a whole is disclosed to Dasein; when we say that Dasein is its disclosedness²⁵, we are saying that Dasein is its “there” (BT 171). In other words, in the “there,” Dasein’s existence in the world is disclosed to it. Ontologically, the “there” is constituted by three structures²⁶ of existence: state-of-mind or attunement, understanding, and falling (BT 400). This paper will focus on the first of these.²⁷

As established, we are always in the world in the sense that the world matters to us. And the way in which it matters to us is determined, in each case, by the mood in which we find ourselves; this mattering-of-the-world-through-a-mood is what is described ontologically as “attunement” (BT 172). The word Heidegger employs is *Befindlichkeit*, which he constructs from the German greeting “*Wie befinden Sie sich?*” which in turn translates literally to “How do you find yourself?” (Greaves 65) Although the translators of *Being and Time*, Macquarrie and Robinson, use the term “state-of-mind” for *Befindlichkeit*, Heidegger specifically states that it is far from anything like a mental state: “... a state-of-mind [*Befindlichkeit*] is very remote from

²³ These two sentences could be equivalent, actually. Hopefully not.

²⁴ Again, that spatiality, where I am in this room, at the geographical coordinates 16.7666° N, 3.0026° W, is founded upon my Being-in-the-world, which I currently am by being involved in the activity of writing this paper (BT 171); see footnote 3.

²⁵ Disclosedness = the character of having been unveiled, unhidden (BT 105). The term is used for Dasein or Being (whereas “uncover” is used for non-human entities) (Georgetown website). Dasein is its disclosedness because Dasein’s Being is made manifest in disclosedness: “Dasein *is* its disclosedness,” means at the same time that the Being which is an issue for this entity in its very Being is to be its ‘there’” (BT 171). Also, a fun fact: ‘disclosure’ or ‘disclosedness’ is also the translation for the Greek word *Aletheia*, the name of the journal in which you’re (presumably) reading this paper.

²⁶ In other places in *Being and Time*, Heidegger also includes discourse or talk as the fourth existential structure that constitutes the “there” (Elpidorou and Freeman, § 2); in § 34, he includes only states-of-mind and understanding (BT 171-172).

²⁷ As Elpidorou and Freeman rightly note in their paper, these structures of existence cannot be separated except in analysis (Elpidorou and Freeman, § 2). So while this paper analyzes state-of-mind or attunement in isolation from the other structures, it acknowledges that the structures are, in fact, “equiprimordial” (BT 172). Explaining understanding and falling, however, will keep me here the entire night.

anything like coming across a psychical condition” (BT 175).²⁸ It is not an inner psychological state; it is not, for example, my current rather grumpy mood. Nor is it an external state, occurring in the world independent of me (BT 176), despite my saying that I am *in* a rather grumpy mood. Instead, it is a certain disposition towards one’s situation, a way of “finding oneself” in the world. Tom Greaves has a useful way of putting this. When asked “How do you find yourself?” (or in English, “How are you doing?”), I could reply with both “I’m in a grumpy mood”³⁰ or “I’m writing my final paper.” But then that would elicit the response, in the first case, “Why are you in a grumpy mood?” or in the second, “How do you feel about having to write a final paper this late when you really should be sleeping?” For Heidegger, the circumstances in which I find myself and my disposition towards them are not separable; they are two sides of the same phenomenon, i.e., my Being-in-the-world in a particular *way* (Greaves 65). Finding myself in the world, therefore, requires my being practically involved in it, such that things within it have practical significance, inseparable from the activity in which I am engaged (for example, while my computer and this paper and copious amounts of coffee are all significant to me in one way or another at the moment, a hammer would appear completely useless). And this practical significance, the way in which things in the world matter to me, is determined by my mood (Ratcliffe 158). Thus, in a mood, the Being of Dasein is brought to its “there” (BT 173) i.e., the situation in which its own existence is disclosed to it (BT 176; see also above).

According to Heidegger, we are always in some mood or another; it is a fundamental existiale of Dasein. Even the “pallid, evenly balanced lack of mood” is, in fact, a mood, albeit a more inconspicuous one (BT 173; Ratcliffe 157). In German, however, the word “mood” (*Stimmung*) has a much broader range of connotations than its English translation. It means, not a particular feeling or emotion, but any way of being affected (Dreyfus 169).³¹ It is not just the current, individual mood in which I find myself, a feeling, a “bad” mood or a “good” mood as we say in English, for these are mere emotional states founded upon *Stimmung*. Instead, it is a way of encountering the world in a comprehensible way (Hahmann, Canvas discussion board), a way of Being-in-the-world, of receiving the world and allowing it to matter

²⁸ This is why, for the rest of the paper, I will use the term “attunement;” other translations, as Matthew Ratcliffe notes, include “affectedness,” “disposedness,” and “sofindingness” (Ratcliffe 158). Dreyfus even uses “where-you’re-at-ness,” but I will skip that one, as does he (Dreyfus 168).

²⁹ It is important to remember that Heidegger rejects the subject-object view of knowledge, which conceives of human beings as isolated subjects engaging with external objects (BT 85-86). Instead, in his view, we construct our subjectivity from the objects with which we are concerned; the view of disinterested knowledge/the subject-object relationship, therefore, is itself grounded in already Knowing the world, i.e., the primary mode of Dasein’s Being-in-the-world (Hahmann seminar, 2/18/20). Since a mental state or “state-of-mind” implies the inner psychological state of an isolated subject (Dreyfus 169), it cannot be what Heidegger has in mind when he says *Befindlichkeit*, for the latter is something more fundamental.

³⁰ Actually, I’m just exhausted, not really glowering or in a dark mood as such. But it is not pleasant!

³¹ Hence Dreyfus’ translation of *Befindlichkeit* as affectedness (Dreyfus 168).

to us. In *The Fundamental Concepts of Metaphysics*, Heidegger writes, "... it is necessary to dispense with the psychology of feelings and experiences and consciousness ... It is apparent that moods are not things that are only occurrent, but rather they themselves are fundamental manners and ways of being" (Heidegger, *The Fundamental Concepts of Metaphysics* § 17, tr. Blattner). However, this distinction between an occurrent feeling or experience and *Stimmung* or mood as a fundamental manner of Being is quite vague, especially if we consider the general opinion that emotions or feelings are brief episodes directed towards objects, whereas moods are more fundamental, long-term states of acting and thinking that either don't have an object or have many objects (Ratcliffe 163). It is a distinction that becomes particularly confusing in the context of Heidegger's discussion of fear and anxiety.

In § 30 of *Being and Time*, Heidegger illustrates what he means by *Stimmung* through the example of fear. Fear, he writes, can be analyzed in three ways: (1) what we are afraid of, (2) fearing itself, and (3) what we are afraid for (BT 179). When I am afraid of something, I comport myself fearfully towards an entity I encounter in the world; my fear is always directed *at* something within-the-world. For example, say I am afraid of the novel coronavirus. In such a case, I have identified (1) something detrimental in the world, the coronavirus, as threatening to me, and (2) by fearing it, have allowed it to matter to me. And in doing so, I have (3) disclosed my own Being-in, my own existence³², as endangered. However, it is unclear why we must look at such a fear not as a psychological experience or feeling but as a way of encountering the world. After all, it has a specific object, i.e., the virus (or more specifically, the disease caused by it). And although it is by no means brief,³³ it is almost certainly temporary. Once a vaccine is developed, or testing becomes widely available, there would be much less reason to be afraid. Thus, my fear certainly looks like a brief emotional state, and not a fundamental way of belonging to the world (Ratcliffe 163).

Matters become even more complicated because, as Heidegger will develop both in § 40 of *Being and Time* and in later works, fear is not even a fundamental mood (*Grundstimmung*). This is because when we experience fear, we experience it in our everyday Being-in-the-world, which is an inauthentic³⁴ mode of existence for Dasein. In our average everydayness, when we are in-the-world, when we are absorbed by and fascinated by (and afraid of) entities within-the-world, we do so when we have fallen

³² This would probably be my health, which you might be tempted to think is not strictly my existence, but Heidegger reminds us that Dasein *is*, for the most part, that with which it is concerned (BT 180-181).

³³ It feels *endless*.

³⁴ Authenticity and inauthenticity as modes of existence in Heidegger are beyond the scope of this paper, mostly because if I start writing about that I will deliver a 30-page final paper, which I have neither the time nor the energy to do. However, in a *very* small nutshell, inauthenticity is the mode of existence of Dasein in which it turns away from its authentic self and falls (see footnote 17) into the world (BT 220).

away³⁵ from our authentic selves and into the world (BT 230). Therefore, the Being of our existence that is disclosed to us in everyday moods is at once un-hidden and distorted, because it is not the authentic Being of existence. In a fundamental mood, or “basic” mood (BT 228), however, we comport ourselves towards nothing in particular, and through this, confront our very existence, “pure Da-sein,” as Heidegger writes (Heidegger, “What Is Metaphysics?” 89). What is disclosed to us in such a mood is Being-in-the-world as a whole, because we are not attuned *to* this or that entity, such and such possibility, but to the entire totality of involvements³⁶ (Greaves 68). In a fundamental mood, therefore, *Befindlichkeit* is divorced from involvement in a particular activity, because I am no longer disposed towards anything *in-the-world* (see above). Rather, I find myself among beings as a whole, am confronted with the “there” of my very existence (“What Is Metaphysics?” 87-89). In his oeuvre, Heidegger sketches out various fundamental moods – profound boredom, love³⁷, anxiety (“What Is Metaphysics?” 87-88), Greek wonder, holy mourning. The fundamental moods are what first expose us to entities (Dahlstrom 135); they are what make our everyday moods possible in the first place.

The fundamental mood Heidegger examines in *Being and Time* is that of anxiety. In anxiety, the public world of significances that Dasein did not produce (Dreyfus 177) recedes, the totality of beings become superfluous (“What Is Metaphysics?” 90). It is anxiety that makes fear possible; they are “kindred phenomena,” except in anxiety, there is no one particular entity that threatens. What threatens is rather Being-in-the-world as a whole, existence as a whole (BT 230). It is to avoid this anxiety that we develop fears – my fear of the coronavirus is only possible because I do not want to feel anxiety in the face of my own existence (Dreyfus 182). When I fear something within-the-world, I flee anxiety and dive into the world instead. We might even conceive of Being-in-the-world, doing this or that thing, taking exams, writing final papers³⁸, as a kind of distraction from the anxiety we feel about our existence. Thus we have another problem: if anxiety is the mood in which fear³⁹ is grounded, it becomes increasingly difficult not to look at the latter as an emotional state founded upon the former, which may qualify as a mood in Heidegger’s sense of the term. And yet

³⁵ Falling = one of the other equiprimordial existential structures of Dasein, by way of Being-in-the-world; Dasein has fallen into the world and away from itself (BT 220). See paragraph 2 and footnotes 8 and 9. This is also beyond the scope of this paper; see footnote 16 for why.

³⁶ “That in the face of which one has anxiety is not an entity within-the-world. Thus it is essentially incapable of involvement” (BT 231).

³⁷ I read the letters he wrote to Arendt for this, because he never quite explains love, or rather, “joy at the presence of another Dasein ... whom we love” (“What Is Metaphysics?” 87) and I think this excerpt might help shed some light: “Why is love rich beyond all other possible human experiences and a sweet burden to those seized in its grasp? Because we become what we love and yet remain ourselves ... The other’s presence suddenly breaks into our life — no soul can come to terms with that. A human fate gives itself over to another human fate, and the duty of pure love is to keep this giving as alive as it was on the first day” (Popova).

³⁸ A welcome distraction on most days, don’t worry.

³⁹ Recall that my fear of the coronavirus is both temporary and object-oriented.

Heidegger tells us quite decisively that fearfulness is not “in an ontical sense ... some factual ‘individualized’ disposition, but an existential possibility of [attunement] in general” (BT 182).

So how do we reconcile that fear is both (a) grounded in a more fundamental mood, and (b) not an individualized disposition or feeling, but a mood itself, a possibility of attunement? A clue to this problem may be found in Lars von Trier’s 2011 film, *Melancholia*. The film centers around two protagonists, sisters Justine and Claire, and their characters and reactions to the world ending when a rogue planet, Melancholia, collides with Earth. For most of the film – for most of her life so far – Justine has suffered from what we would ordinarily dub depression, but what Heidegger would probably characterize as anxiety. She finds herself often unable to do things required of her and of living a life in the world – in the first half of the film, which centers on one night (her wedding night), we see her repeatedly leaving her own reception, which she clearly finds draining and meaningless, to take a nap or a bath; as the evening progressively turns into an unmitigated disaster, she has sex with a stranger, tells her boss she hates him and his firm and quits, and finally, bids goodbye to her new husband, who leaves as soon as he realizes that their marriage is completely doomed. In the second half of the film, we see her absolutely incapable of performing even the most basic of tasks: getting out of bed, eating, washing herself. Claire, on the other hand, is very much the opposite – she clearly has some form of what we might call anxiety, but what Heidegger would refer to as fearfulness. She is happily married, very organized, and in the first part of the film, constantly asks Justine to pull it together and enjoy the party (which she and her husband have thrown). In the second part, she acts as Justine’s responsible older sister; she is the one who cares for Justine, bathing her, feeding her, getting her out of bed and forcing her to get some fresh air and exercise. But as it becomes clearer that Melancholia is approaching Earth, we see a shift in their characters. While Justine becomes more upbeat, Claire grows more and more afraid. And towards the end of the film, when it becomes certain that Melancholia will hit Earth, thereby wiping out all life on it, while Justine is calm⁴⁰ about their fate, Claire begins to panic. Where Justine was the one who had lost her head before the threat of the world ending presented itself, it is now Claire who loses her head entirely. In this shift, and in their reactions, we may see a manifestation of the Heideggerian notions of attunement and mood. For their reactions – what we might dub temporary⁴¹ feelings – are grounded in their fundamental characters, or their *attunements*. Justine finds it impossible to flee anxiety, that “uncanny feeling” (“What Is Metaphysics?” 88), and dive into the world. Claire, on the other hand, is able to do so

⁴⁰ “... a peculiar calm pervades [anxiety]” (“What Is Metaphysics?” 88).

⁴¹ I mean, the world is literally about to end.

competently, turning away from anxiety and towards tasks and responsibilities in-the-world. Although her mode of existence, according to Heidegger, would be inauthentic, because she has turned away from the anxiety that characterizes Justine's character and towards the world, that turning away does not obliterate anxiety, which is still present. Rather, it takes on the form of fear. And it is in this general fearfulness of Claire's (what the psychological community may call a generally anxious nature), her fearful *attunement* towards the world, that her fear of a particularly entity within-the-world, Melancholia⁴² colliding with Earth, is grounded. Fearing *something* is grounded in the mood of *fearfulness*, a possibility of attunement.

Let us go back, then, to my fear of the novel coronavirus. Although my fear of the virus is indeed temporary and object-specific, it is not the same as the *mood* of fearfulness itself. Rather, it was my mood of fearfulness as Being-in-the-world, distracted from the anxiety that pervades my existence, that made it possible for me to encounter the coronavirus as a threat in the first place, and then be afraid of it: "Circumspection sees the fearsome because it has fear as its [attunement]. Fearing, as a slumbering possibility of Being-in-the-world in [an attunement] has already disclosed the world, in that out of it something like the fearsome may come close" (BT 180). In other words, I am already a fearful person, and the world has therefore already been disclosed to me as one where it is possible to be afraid. Only because I am such a person can I be afraid of the virus more specifically.

Therefore, my own current "mood" of grumpiness – which has since turned into relief because I am almost done with this paper – would in fact not be a mood in Heidegger's sense of the word. On the contrary, it would be a feeling, a reaction, an "individualized disposition" (BT 182) grounded in my attunement in general. To find a final assignment annoying itself depends on the possibility of being annoyed, on finding myself in the world in such a way that includes the possibility of being annoyed (Ratcliffe 163). It is something that shapes my entire way of living, acting, thinking, of receiving the world – of Being *there*, and therefore, my existence⁴³. Tomorrow I will be grumpy because my favorite bakery is closed, the day after because I will be in Philadelphia and not be able to see my friends. All of this depends on the possibility of being grumpy in the first place, because of the *way* in which this paper, the bakery, my friends matter to me; in other words, the way in which my world matters to me, which is determined by my attunement, my mood⁴⁴. What Heidegger provides us in his explication of mood is not simply an ambiguous use of a well-

⁴² Technically, you might think, Melancholia might not be in-the-world, given that it is literally outside the world/Earth, but not for Heidegger: the world is not an entity that is encountered, but rather an environment in which Dasein lives (BT 93).

⁴³ Remember that Dasein is its there (see paragraph 2; also BT 171).

⁴⁴ This is of course an example. I am not attuned by grumpiness. I think.

known word – it is instead a way of thinking about how we belong to the world in which we live.

Works Cited

- Elpidorou, Andreas; Freeman, Lauren. "Affectivity in Heidegger I: Moods and Emotions in Being and Time." *The Philosophy Compass*, 2015. URL = <onlinelibrary-wiley-com.proxy.library.upenn.edu/doi/full/10.1111/phc3.12236>.
- Blattner, William. Some Terminology in *Being and Time*, Georgetown faculty website, 2011. URL = <faculty.georgetown.edu/blattnew/topics/index.htm>.
- Dahlstrom, Daniel O. *The Heidegger Dictionary*. Bloomsbury, 2013.
- Dreyfus, Hubert L. *Being-in-the-World: A Commentary on Heidegger's Being and Time, Division I*. The MIT Press, 1991.
- Greaves, Tom. *Starting with Heidegger*. Bloomsbury, 2010.
- Hahmann, Andree. Seminar on Continental Philosophy, University of Pennsylvania, Spring 2020. ---. Seminar on Continental Philosophy, Canvas Discussion Board on §29, University of Pennsylvania, Spring 2020.
- Heidegger, Martin. *Being and Time*, Translated by John Macquarrie and Edward Robinson. Blackwell, 1962.
- Heidegger, Martin. *The Fundamental Concepts of Metaphysics*, § 17. Translated by William Blattner, URL = <faculty.georgetown.edu/blattnew/heid/gbm17.htm>.
- Melancholia*, directed by Lars von Trier, Zentropa Entertainments et. al., 2011
- Popova, Maria. "The Remarkable Love Letters of Hannah Arendt and Martin Heidegger." *Brain Pickings*, 2016. URL = <www.brainpickings.org/2016/04/25/hannah-arendt-martin-heidegger-love-letters/>.
- Ratcliffe, Matthew. "Why Mood Matters," from *The Cambridge Companion to Heidegger's Being and Time* (pp. 157-176), edited by Wrathall, Mark A., Cambridge University Press, 2013.
- Heidegger, Martin. "What Is Metaphysics?" from *Pathmarks*, ed. McNeill, William, Cambridge University Press, 1998.

EDITOR'S CORNER

The Problem of Mental Causation and the Autonomy Solution

Daniel Lightsey, *Texas A&M University*

The purpose of this paper is to analyze a physicalist solution to the “Problem of Mental Causation” (PMC) called the autonomy solution. Many contemporary philosophers endorse some sort of physicalism and therefore think that the mind comes solely from physical substances. I am skeptical of such claims, and in this paper, I aim to show that the autonomy solution to the PMC fails to make the mental causally relevant. I begin the paper with defining the PMC in detail and stressing the importance of finding a solution to the problem. In doing so, I will pull mainly from Jaegwon Kim’s work as he eloquently describes the problem, and presents an argument that shows the mental to be causally inefficacious unless it is fully reducible to the physical. Then, I will move to define the autonomy solution, and raise two objections against it. I conclude that the autonomy solution fails to adequately solve the PMC, and that if we want to keep mental causation, we must look for a solution elsewhere.

I. The Problem

I had a desire to get some coffee this morning due to a lack of sleep, so I went to Starbucks and got my favorite cup of coffee. Naturally, we explain this behavior by saying that my feeling of tiredness *caused* my desire for coffee, and this desire *caused* me to go to Starbucks. These are examples of mental causation. Some mental event (or property) caused some other mental or physical event. We all experience this sort of causation daily, but explaining how something such as an intention or desire—something that is purely mental—can cause some physical event can be difficult because of their contrasting natures. Further, mental to mental causation, as opposed to mental to physical, can be difficult to explain because of the mind’s apparent dependence on the brain.⁴⁵ This is the so-called “Problem of Mental Causation” (PMC), and finding a solution to this problem is tantamount to any other significant philosophical question. For if we give up mental causation, then we give up our status as agents in the world, and thus any significance of being human. Jaegwon Kim has written extensively on this subject, and in his paper “How Can My Mind Move My Limbs? Mental Causation from Descartes to Contemporary Physicalism”, he presents an argument that shows that the only way

⁴⁵ For the purposes of this paper, I assume some sort of physicalism regarding the mind. Mental to mental causation may be easier to explain or to understand if something like physicalism is not true. Since the mind would not have strict ontological dependence on the brain, different mental events could cause others without the need for a physical cause.

for the mental to be causally efficacious given physicalism is for it to be reducible to the physical (Kim 11). I find his argument compelling, and in this paper, I will examine one of the more interesting objections to Kim's argument and solution to the PMC called the autonomy solution (sometimes called the dual explanandum strategy).

There are essentially two views about the mind that entail that the mind is not reducible to the physical, and thus distinct from the brain. One view is to postulate the mind as a separate entity or substance that causally interacts with the body. This view, called substance dualism, was originally proposed by Descartes. It has been largely abandoned today due to its lack of ability to explain *how* the mind and body interact, so I will not consider it here.⁴⁶ The second view, called property dualism, has become popular in contemporary literature to explain mental causation. Property dualism denies that the mind is a separate substance from the body and holds a physical ontology. It also asserts that when physical systems reach a certain complexity, they exhibit properties that are not physical. For example, property dualists claim that the physical complexity of the human brain has given rise to mental properties, like intentions or beliefs, which are understood as being purely mental and distinct from any physical property of the brain. Moreover, these emergent properties cannot be *reduced* down to the physical properties (Kim 9). Rather, the mental properties supervene on the physical properties. *A* is said to supervene on *B*, if and only if, there can be no change in *A* without a corresponding change in *B*. Thus, if a mental property is instantiated, then it has a physical property as its base that instantiates it (Kim 9). As an example, take the desire of mine to get coffee this morning. From supervenience, we can say that the desire I had must have also had a *corresponding* physical state in my brain that gave rise to or generated my desire. Without this physical state in my brain, the desire could not have been generated, *but* the desire cannot be explained and understood fully in terms of the physical state—in other words, the mental is not necessarily *reducible* to the physical even if it supervenes on it.

So, even though there is no separate substance that is the mind, the mind has properties that cannot be explained in terms of the physical such as desires, beliefs, intentions, etc. This seems to leave open the idea for the possibility of mental causation. There are good reasons for considering such a view in order to explain things like consciousness and the fact that mental states are multiply realizable by different brain states. For example, you and I can have the same belief about

⁴⁶ Personally, I think the claims that substance dualism fails to be able to explain interaction between the mind and the brain are vastly overstated. Moreover, I think substance dualism is able to make a more accurate account of reality than any of its competitors—by being able to explain consciousness and other subjective experiences; consequently, I think it, or something like it, is true. However, as I have mentioned above, in this paper I am focusing on physicalism's inability to make sense of mental causation, so I will be ignoring substance dualism for the rest of this paper.

College Station (that it is located in Texas), but different parts of our brains correspond to these beliefs. We have the same mental state but different brain states. However, this is exactly the kind of view that Kim attacks in his paper. If Kim's argument is valid, then any kind of property dualism entails that mental events are also causally inefficacious.

II. Kim's Argument

To understand Kim's argument, we first must consider what assumptions the argument makes. The argument seeks to undermine any physicalism that is non-reductive about the mental, like property dualism. Physicalism is committed to the causal closure of the physical (CCP): Any physical event that has a cause at time t , had a physical cause at t (Kim 10). This is the first assumption that Kim makes, and it is one that is undeniable for any physicalist. The second assumption Kim makes is the principal of causal exclusion (PCE) which states: If an event, e , has sufficient cause, c , at time t , then no other cause brought about e (Kim 10). The last assumption, the principal of generative exclusion (PGE), is more general than the PCE, and it claims: If an event e , or an instantiation of a property P is generated by an event c , then the existence of e or of P , is not generated by anything other than c (Kim 11). Some of these assumptions may be controversial or debatable, but for the sake of the paper I will assume they are valid and reasonable.

Kim's argument seeks to refute any kind of mental to mental or mental to physical causation that is proposed by property dualism. So, let us take the example mentioned at the beginning of the paper where my sensation of tiredness caused my desire or intention to get coffee and apply Kim's argument to it. The argument starts off by claiming that my sensation of tiredness caused my intention to get coffee. Because the mental supervenes on the physical, my intention must have a supervenient physical base. Suppose this base is my b-fibers firing in my brain. But now, my intention has two supposed generative sources, my sensation and my b-fibers firing. By the PGE, my sensation and my b-fibers firing cannot both be generative sources for my intention. The only way for my sensation and my b-fibers firing to be generative sources for my intention is to claim that my sensation caused my intention by causing my b-fibers to fire. Thus, my sensation caused my b-fibers to fire. However, by the CCP, my b-fibers firing must also have a physical cause. Suppose this cause is my a-fibers firing in my brain. Now there are two purported causes for my b-fibers firing: My sensation and my a-fibers; a physical and a mental contributor. But by the PCE, one of these causes must go. By invoking the CCP again, my sensation must go as a possible cause, and therefore the firing of my a-fibers caused my b-fibers to fire, which in turn caused my intention. This

contradicts what we assumed at the beginning of the argument—that my sensation caused my intention—and thus shows that the mental is causally impotent if property dualism is true.

If the mental is causally impotent, then mental events and properties are purely epiphenomenal. This has led some philosophers to adopt epiphenomenalism. This view holds that physical events cause mental events, but mental events have no causal relation or impact on any physical or mental events, and this appears to be the case for property dualism if Kim is right. However, many are reluctant to accept this view. It is no wonder why so many are hesitant to affirm the mental as purely epiphenomenal, since the whole point of property dualism is to make the mental something that is special and non-reducible to the physical. But on physicalism, as Kim's argument shows, the only way to allow for mental causation is to allow for the mental to be reducible to the physical. Kim puts it this way, "Unless you bring the mind fully into the physical domain, there is *no hope* of explaining how the mind can participate in the physical causal processes" (emphasis added, Kim 15).

Furthermore, if the mental is epiphenomenal and can be reduced to the physical, or more specifically the neurobiology of the brain, then haven't we lost the very thing that makes the mental important? The mental is not something over and above the physical; rather it is just a physical causal structure like everything else. Kim eloquently makes this point and writes, "If [mentality is reducible], we can save its causal powers but in the process we may lose just the thing we set out to save, namely minds" (Kim 16). If a property dualist wants to keep anything relevantly special about the mental and hold the view that the mental is causally efficacious, then they need to refute Kim's argument.

III. The Autonomy Solution

Kim's argument is cogent, so in order to refute Kim's argument his opponents need to show that one of his three assumptions are wrong. The CCP is non-negotiable for any physicalist, so many objections to the argument attack the causal principles. These objections take many forms, but I do not have the space to address all of them here. As mentioned earlier, the most intriguing objection is the autonomy solution, so this is where I will focus. According to the autonomy solution to the PMC, the idea that mental properties must do the work of their physical realizers or face exclusion is wrong. Rather, mental properties can have causal relevance in their own way by causing mental effects or, as some suggest, behavioral effects. For example, whenever we use psychological explanations about the mental, we rarely say the physical realizers caused the behavior or some mental

effect. The explanation of the causes can be made independently of the physical and can actually add to our understanding, which is why the autonomy solution is sometimes called the dual explanandum strategy. The mental need not be causally excluded from the effect, since the mental and the physical are causally relevant to different aspects of the effect (Heil and Robb 2018).

Thus, there are two ways to explain some event. We can explain it in terms of the mental, in an intentional/rational way, or we can explain the event physically/mechanically. Those who support the autonomy solution suggest that when we explain some event, both explanations are required since the mental and physical are causally relevant to different effects. The philosopher Donald Davidson has given an example of where both explanations are necessary (Silcox 2020): Consider two climbers, one in front of the other, that are scaling a mountain along a rope. In this unfortunate scenario, one of the climbers accidentally causes the death of the other. This accident happened because the climber in front thought that he would be safer if he loosened the rope, and this thought simply unnerved him, causing him to unintentionally loosen the rope, thus killing his friend. Now, Davidson points out, consider the same situation where the climbers are scaling the mountain, but this time one of the climbers dies because his friend intentionally loosens the rope.

In both cases, the climber who killed his friend wanted to loosen the rope. Moreover, these examples look identical to any observer watching, but there is a clear difference in the latter case where the surviving climber *intentionally* loosened the rope to kill his friend rather than *unintentionally*. If we were to explain things in a purely physical or mechanical way, then these two events would be the same. But if we add the intentional explanation, then we can differentiate between the two events and see them as they really are. This is exactly what is expected from the autonomy solution. According to the autonomy solution, the climber who loosened the rope had a physical cause that caused him to loosen the rope, and this explains his physical behavior and effects. But there are also the mental causes that were relevant to whether or not the action was intentional or unintentional, and in both cases the climber who loosened the rope, wanted to but only in the second did he *intend* to.

Thus, neither the physical nor the mental need to be excluded as causes or explanations because they refer to different parts of the event. Applying this reasoning to the example of my desire to get coffee this morning, we can say that my sensation of tiredness caused my desire or intention to get coffee and also caused me to believe that I desired coffee, and my behavioral actions of actually going to get the coffee were caused by physical causes. Some who support the autonomy solution advocate a stronger version that claims that mental events can

cause certain behavioral effects, but these advocates still hold that the mental does not cause any physical effects as to not violate the CCP (Silcox 2020).

IV. Objections

There are two reasons why the autonomy solution to the PMC is not in fact a solution: It makes an explanatory mistake and it cannot explain mental causation without denying one of the tenants of physicalism, namely supervenience. To consider the former, I must first explain what I mean by an explanatory mistake. The supporter of the autonomy solution argues that the mental and the physical independently cause different aspects of the same event. Thus, the fact that we can explain the event in two different ways, the intentional/rational or physical/mechanical, means that there are two different sets of causes that work independently to produce the distinct aspects. You could explain my sensation of tiredness and my desire to get coffee as reasons or causes as to why I went and got coffee from Starbucks this morning. This would be explaining the event in a rational way. It uses a reference to an agent, me, and the reasons the agent did what he did. The other way to explain the event is to say I had some physiological cause that made me to get tired and had some chemical reaction in my brain due to this cause, and these combined prompted me to go to Starbucks. This is explaining the event in a physical manner. It does not reference any of the mental events or states and makes no reference to agents in order to explain the event.

Of course, using these two explanations adds to our *understanding* of the event. This is how we can differentiate between the climber example given by Davidson. This is to say, they add to our epistemological understanding of the event, but using two different explanations does not necessarily add to our metaphysical understanding. Using two different explanations gives us knowledge we may have not had before by allowing us to use two different concepts to describe the event, i.e. rational concepts and physical concepts. But the fact that we can understand or explain an event in two different ways, does not mean that each of the explanations are metaphysically relevant to the causes that brought about the event. This is a common response that many physicalists make to refute dualist arguments.

For example, in her paper, “In Defense of the Phenomenal Concept Strategy”, Katlin Balog uses this idea to explain why Mary, the character in Frank Jackson’s famous “Knowledge Argument”, learns something new, but this fact does not imply the existence of some new metaphysical thing. To understand the “Knowledge Argument”, consider a woman, named Mary, that learns everything there is to know

about the physical world all while locked in a black and white room. Consequently, Mary knows everything about colors and color vision, but she has not seen any colors. Upon leaving the room, Mary sees a ripe tomato that manifests the color red vividly. Many of us, myself included, have a strong intuition that Mary learned something new when leaving the black and white room, namely “what it is like” to see red. Jackson concludes that because Mary learns something new when leaving the room, even though she knew all the physical facts, some things that are not physical must exist (Jackson 291). However, if we use Balog’s reasoning, then Mary learned something new not because things that are not physical exist, but because she was able to use new concepts, namely phenomenal concepts, to understand red in a different way (Balog 2012).

Nothing about this necessitates that there must be something new metaphysically. Rather, it just necessitates something new epistemologically. Similarly, just because we can explain some event or effect in a rational way as opposed to a physical way, and gain new knowledge about it, does not entail that these rational or intentional ‘causes’ have any metaphysical significance. These ‘causes’ have epistemological significance as we learn why a person did what they did, but we merely have employed new concepts to understand the event in a different way. Thus, the fact that we can explain things two different ways does not entail that both explanations have causal significance. And given physicalism, it is very likely that the only metaphysically significant causes are physical.

Other than the explanatory mistake, the other reason the autonomy solution fails is that it cannot explain the efficacy of the mental without denying supervenience. Suppose that these explanations of the rational or intentional do actually imply the mental has some sort of causal relevance to either itself or some behavioral effects. Does this solve the problem of causal exclusion for the property dualist and explain how mental causation works? Not at all. For if the supporter of the autonomy solution accepts Kim’s assumption of the CCP and accepts the supervenience of the mental on the physical (both seem to be necessary for any physicalist theory of mind), then they must claim that the mental is causally inefficacious. To say it is causally relevant you would have to explain *how* it is causally relevant.

On the weaker version of the autonomy solution, you cannot do this without leading to absurdities. Consider what is being claimed by the weaker version of the autonomy solution, that some mental event causes another mental event or effect without causing any behavioral or physical effects. As some philosophers have pointed out, such as Daniel Dennett, this would mean that someone’s desire to get coffee causes them to believe they are about to get coffee, but is not causally

relevant to their action of getting the coffee (Moore 2020). This seems incredible. How does my desire cause my belief that I am about to go to Starbucks, but have *nothing* to do with my actual action of going to Starbucks? If you think the mental has any causal powers, then you ought to claim that the mental is causally relevant to its related action or behavior if its powers are to do anything meaningful. But if someone were to endorse such a view of the autonomy solution, then they would have to deny this.

Not only are the conclusions from this reasoning implausible, but concluding something such as mental to mental causation without reference to the physical is to deny that the mental supervenes on the physical. As Karen Bennett writes regarding this strategy, “Why does the dual explanandum strategy require denying that the mental supervenes on the physical? Because the supervenience claim extends the exclusion argument to problematize mental to mental causation as well” (Bennett 326). According to supervenience, any mental property or effect *must* have a physical supervenient base that generates it. This physical supervenient base must also have a physical cause that generated it based on the CCP. This excludes the mental from having any causal potency.

There is the stronger version of the autonomy solution that claims that not only can the mental cause mental effects or events but also behavioral effects, so it avoids the absurd conclusions that the weaker version must make. But this does nothing to help the supporters of the autonomy solution’s case with regards to supervenience. For the same argument above can be applied to the behavioral effects, i.e. they must have some physical cause and the mental event that ‘caused’ the behavioral effect must have also had a physical cause. The only way to get around this is denying supervenience and run afoul of the CCP. Again, these are not options for the physicalist. Therefore, if someone supports the autonomy solution, then it seems they have to affirm epiphenomenalism since the mental is causally impotent.

V. Conclusion

Consequently, the autonomy solution does not solve the PMC for the property dualist. The supporter of the autonomy solution’s attempt to claim that mental properties are *distinct* from the physical, and are therefore not reducible to the physical, leads to the conclusion that the mental is causally inefficacious. Moreover, they claim that the mental and the physical enjoy causal relevance independent of each other and explain different aspects of an event. However, just because an event

can be explained in two different ways by referencing the different independent ‘causes’, does not necessitate that two different causes *exist*. Instead, we are able to do this because of our ability to utilize different *concepts*. Further, given physicalism, I noted that the attempt to explain mental to mental or mental to behavioral/physical causation with the autonomy solution fails because in order to keep the causal efficacy it must deny supervenience. Hence, all the causal work is done by the physical, and explaining the rational ‘causes’ of an event simply gives us new knowledge and a different understanding of the event.

I find that many of the other objections to Kim’s argument have similar difficulties that the autonomy solution has, and thus it seems that in order for there to be mental causation, on physicalism, the mental must be reducible to the physical. But as we saw earlier, for a physicalist to affirm reductionism is to give up the very idea of what we mean by the mental. As Kim notes, “If we save minds’ causal efficacy only by giving up minds, that seems like no way to save it” (Kim 13). The problem of mental causation remains unsolved, but at least we can say the autonomy solution, or anything like it, fails as an adequate response.

Works Cited

Balog, Katalin. "In Defense of the Phenomenal Concept Strategy." *Philosophy and Phenomenological Research* (2012): 1-21.

Bennett, Karen. "Mental Causation." *Philosophy Compass* (2007): 316-337.

Heil, John and David Robb. "Mental Causation." *The Stanford Encyclopedia of Philosophy* (2018) <https://plato.stanford.edu/entries/mental-causation/#ProIIIExc>.

Jackson, Frank. "What Mary Didn't Know." *The Journal of Philosophy* (1986): 291-295.

Kim, Jaegwon. "How Can My Mind Move My Limbs? Mental Causation from Descartes to Contemporary Physicalism." *Philosophic Exchange* (2000): 1-13.

Moore, Dwayne. "Mind and the Causal Exclusion Problem." *Internet Encyclopedia of Philosophy* (2020) <https://www.iep.utm.edu/causal-e/>.

Silcox, Mark. "Donald Davidson: Anomalous Monism." *Internet Encyclopedia of Philosophy* (2020) <https://www.iep.utm.edu/anom-mon/>.

ἀ
λ
ή
θ
ε
ι
α