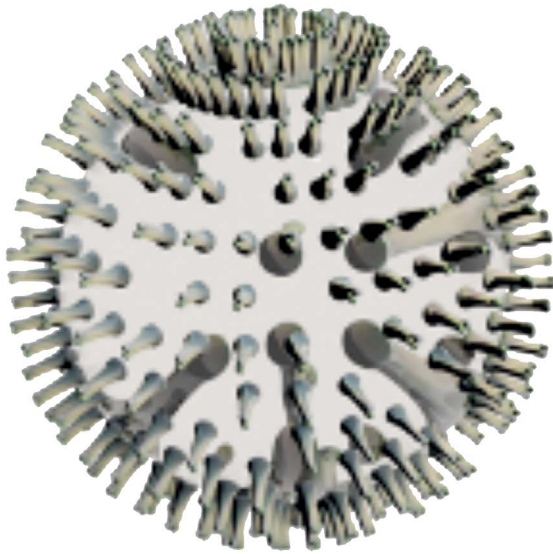# Philosophy in Practice

Volume 14 – Spring 2020

California State University, Los Angeles
Department of Philosophy

Philosophy in Practice

Cover art: 3D-rendered photo of the novel coronavirus. Our hearts go out to those most vulnerable in this difficult time.

CONTENTS

ACKNOWLEDGMENTS

EDITORS:

David Fonth
Sakib Ibne Shahriar

FACULTY ADVISOR:

Dr. Michael K. Shim

CALIFORNIA STATE UNIVERSITY, LOS ANGELES PHILOSOPHY FACULTY

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

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

Bruce Atta (1996–), M.A. California State University, Los Angeles. Metaphysics, Epistemology, Ethics, Social and Political Philosophy

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

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

Jay Conway (2005–), Ph.D. University of California, Riverside. 19th-20th Century Philosophy, Modern Philosophy, Aesthetics, Social and Political Philosophy

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

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

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

Keith Kaiser (2001–), Ph.D. University of California, Los Angeles. Metaphysics, Philosophy of Mind

Steven R. Levy (2005–), Ph.D. University of California, Los Angeles. Epistemology, Metaphysics, Logic, Applied Ethics

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

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

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

EMERITUS PROFESSORS

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

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

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

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

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

Joseph Prabhu (1978–2017), Philosophy of Religion, 19th and 20th Century German Philosophy, Moral and Social Philosophy, Indian and Comparative Philosophy

Kayley Vernallis (1993–2019), Moral Psychology, 19th and 20th Century Continental Philosophy, Feminist Philosophy, Ethics, Aesthetics, Gender and Sexuality

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

Professor Spotlight: David Pitt
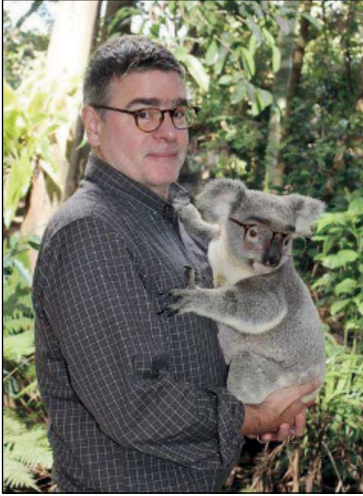
David Pitt began his college education as a music major. Music was the most important thing in his life. He was discovering an ensemble of artists he loved: The Beatles, Duane Allman and the Allman Brothers, and Wes Montgomery. He was into jazz. He moved from jazz to listening to more classical music—he recalls having listened to Alban Berg's opera Wozzeck and being in awe. He wanted to become a classical composer. He started his M.A. program in Music Composition.

It was in the first year of this program that David realized that he didn't have what it takes to become the kind of composer he wanted to be. He started thinking of studying something else. He said to a friend, "I might study philosophy," and his friend advised that he at least finish his M.A. in Composition anyway, so that he'd "have something no other philosopher will have." So he finished the program. Though he diverged towards philosophy, he brought with him compositional skills and a musical image of thought that he incorporates into his philosophical practice.

David is a vivid storyteller. A large chunk of our interview consisted in him sharing with us a great variety of stories about his life, his academics, and his career. He remembers having taken a philosophy course on Hegel's *Phenomenology of Spirit* with Richard Bernstein, a known Hegel scholar. David did not know how to make head or tail of Hegel's writings. He tried his best to understand what Hegel was saying, but to no avail. At one point, sitting in his favorite sleeping chair in a library basement and

brooding with frustration, he threw the *Phenomenology* across the room. He had to tape it back together afterwards. (In class, he got the hang of speaking as if he understood the concepts in the readings, which garnered praise and encouragement from his professor.)

Though he did not find much value other than disdain with Hegel in particular, David sees positive philosophical value in confusion. Confusion means that you are trying to understand something, and there is great value in working your way out of it. It is appropriate to read a philosophical text and have to struggle with it—philosophy is disorienting! Philosophy is also a language unto itself, one with different dialects depending on the subfield and tradition. You have to fully immerse yourself in a language before you start understanding it; you have to throw yourself into the text. David notes that as you continue to grapple with your confusion, you may also have to discern whether the text is confusing because of your own capacities, because of the quality of writing, because of the specific beliefs that the writer holds, or a mixture of all three. David emphasizes the distinction between understanding and believing a viewpoint. Your confusion about a philosophical view need not indicate that you don't understand it. It may well be that you do, but you happen to consider it to be an outlandish view you cannot believe.

David continued to work through his bouts of confusion as he took on his Ph.D. in philosophy. He mentions having felt like an impostor many times, having felt as if he was in way over his head. Yet, he kept at it. He wanted to do something with philosophy, so he stuck with it. He was better at this than he was at music, and though the requisite training was brutal, he felt confident he could do the work.

Despite this divergence in his own career from music to philosophy, David's description of his philosophical work is musically adorned. He mentions a quote about improvisation from jazz musician Charlie Parker: "Master your instrument, master

the music, and then forget all that bullshit and just play." If you don't take the time to do the preparatory work, the quality of your improvising and creative activity will suffer. So it was the case with his own philosophical apprenticeship. After wrestling with the material for 10 years, he finally thought: okay, "*now* I know the techniques and basic literature well enough to start being creative with this stuff; I understand enough not to say something stupid and irresponsible."

Some of that creativity shows when he works in jokes and humorous images into his writings and classes. David talks about a journal article he wrote that got rejected 12 times. It was because the article began with a joke about a giant dog with its head in Brooklyn and its tail in the Bronx. He assures us that the point of the joke was that once the reader got it, they would understand the more important philosophical point he was trying to make in the article. Through humor, David tries to get the reader to see what is important or wrong about a viewpoint. In his lectures, humor has the added benefit of shared laughter, helping students feel more comfortable in the class, more willing to ask questions or raise objections without feeling stupid or put on the spot. In a survey class on the philosophy of mind, he once described a certain aspect of consciousness as the "creamy center" of a Cadbury Creme Egg. He kept repeating that phrase "creamy center," arousing more laughs in the room. Learning should be pleasurable. He does not intentionally craft examples to try to be humorous. His jokes form on the spur of the moment, arising from his own sense of humor, and he is glad he can make good use of them in philosophical contexts.

David falls back on musical ideas when he composes a philosophy paper. Classical composers sometimes hide away small details and secrets into a composition. Bach did this with the letters of his own name, using the notes that corresponded to those letters as a melodic motif in many of his pieces. David himself would sometimes quote a small motif from Bach in his

own earlier compositions. In philosophy, David hides away verbal secrets in his writings. In the book he's working on, he uses many examples using names, many of which point to aspects of his own life—minuscule autobiographical details.

In his writing, David finds himself having many of the same concerns he had when composing music. How can a theme return multiple times in a composition, often coming back changed in a different context? He worries a great deal about structure: of the entire paper, of an argument, and of an idea. He also wants it all to *sound* good, occupying himself with the choice of words, flow, pacing, and rhythm of the piece. "Though an idea may be important, perhaps it belongs earlier rather than later in the piece." It takes a long time for David to write because of this desire to write good prose.

David's movements toward music and philosophy were driven by the same element: the desire he has to *make* something, to create. Part of his joy in doing philosophy and music alike is being able to create a well-structured piece that is beautiful, a piece that stands on its own. David's own musical preferences lean towards, in his own words, "complex as opposed to complicated." Though a piece may be driven by complex rules and relations, the ideas in the piece must be noticeable and easy to grasp to the layperson. A good musical piece should resonate with the listener, even if they have no understanding of the technique that goes into composition. He holds himself and other philosophers to the same aesthetic values. A philosopher should try their best to make their work speak to the reader the way music can speak to a listener. Philosophy is difficult enough on its own. A philosopher's ideal in their work should be "clarity and simplicity with depth." They must show the importance of a philosophical problem, not obscure the problem further in an effort to seem intellectual.

We asked David one last question: "If someone who is not acquainted with philosophy approached you and asked you why anyone should do philosophy or care about it, how would you

answer?" David responded with three points. First, though a great deal of philosophy may focus on trying to change the world, it need not have to do so in order to have value as a profession. No matter what we professionally do, it is important that we as humans fulfill our responsibilities towards one another and don't go out of our way to hurt each other. If philosophers can do their jobs as humans, then they should not be held up to a higher-than-usual standard for changing the world through their profession. Nevertheless, there is still a sizable amount of philosophical work done that addresses social and political issues. Philosophy can help us think with greater clarity and focus about the world; it is capable of helping us see what must be changed and how. And finally, philosophy has its own intrinsic value: it's fascinating to think about conceptual setups and problems, even when those problems are not directly connected to social and political issues. There is joy in thinking. Thought may sometimes be frustrating, but once it starts flowing and concepts start coming together, it brings with it a harmony and beauty akin to art.

Ending the interview, David returned a couple questions to us: "Why do you do philosophy? Why do you like it?"

David Pitt was born in June of 1959 at the Long Island Jewish Hospital in Queens, NY. He grew up in Nassau County. He received his B.A. in Music from Haverford College and his M.A. in Music Composition from Queens College, City University of New York. He finished his Ph.D. in Philosophy at the Graduate Center of the City University of New York. His dissertation was on the semantics of phrases like "plastic flower" and "rubber chicken." He is currently a professor and department chair at Cal State LA's Department of Philosophy. He specializes in the philosophy of mind, the philosophy of language, and metaphysics.

— D.F. & S.I.S.

ON THE DIFFERENT PERMISSIBILITIES OF GENDER DYSPHORIA RELATED AND BODY INTEGRITY IDENTITY DISORDER RELATED MEDICAL PROCEDURES

David Fonth

INTRODUCTION

The recent phenomenon of individuals suffering from Body Integrity Identity Disorder (BIID) has sparked discussion regarding the extent to which medical professionals ought to comply with the request of such patients. One line of reasoning seemingly available to those in favor of these BIID-related medical procedures is to draw a parallel between the acceptance of medical intervention in cases of gender dysphoria (GD), and to argue that, because the medical situation of BIID patients is sufficiently similar to that of GD patients, BIID-related medical procedures, such as elective amputation, ought to be permissible. In this essay, I articulate this argument as I see it used by Tim Bayne and Neil Levy (2005) and by Sabine Müller (2009), and I put forth some possible initial objections to the view. I then critique some of the ableist assumptions behind these objections following Joel Michael Reynolds (2016), introduce a distinction between social and medical models of disability following Richard Dean (2018), and demonstrate that the pro-BIID-medical-intervention argument fails if one adopts a *social* model of disability. Lastly, I discuss some consequences of accepting the argument in favor of BIID-related medical procedures, and I position my analysis within the bioethical discourse as one that is informed by both trans philosophy and philosophy of disability studies. Ultimately, I defend the soundness of the pro-BIID-medical-intervention argument from an explicitly anti-

ableist position.

BIID is a condition in which one desires to "either be paralyzed or to have one or more of their healthy limbs [amputated]" (Blom, Denys, and Hennekam 2012). Given this desire for elective paralysis or amputation, it does not seem immediately clear whether, say, a surgeon ought to comply with such requests and paralyze an otherwise healthy individual or amputate an otherwise healthy limb. At first glance, there are several pertinent considerations: autonomy (Should we not recognize the freedom of a patient to make their own medical decisions?); competency (Is the patient's reasoning for arriving at their selected conclusion akin to that of cognitively sound patients?); and harm-reduction (Would acquiescing to their BIID-related demands reduce the amount of harm or pain that they experience?), to name a few. Given these features of BIID doctor-patient interactions, one could argue in favor of doctors complying with the medical desires of BIID patients by demonstrating that, for instance, one or more of these motivating factors ought to be a sufficient enough reason for a doctor to amputate the limb of a BIID patient. Indeed, depending on how one medically classifies BIID, there are several arguments in the medical and bioethical literature that attempt to do just this (See Müller (2009) and Bayne and Levy (2005), for instance).

One way that one can construct a pro-BIID-medical-intervention argument—i.e., an argument in favor of the acceptance of elective amputation or paralysis in BIID individuals—is by making an analogy to the situation of GD patients. Tim Bayne and Neil Levy (2005) and Sabine Müller (2009), for instance, seem to briefly hint at this argument on the ethical justification of BIID-related medical procedures.[1] Briefly, the idea is the following: if GD patients are granted identity-affirming medical intervention, and if the medical situation of GD patients is similar in kind to that of BIID patients, then should not BIID patients also be granted identity-affirming medical intervention?

In response to this question, I will emphasize and evaluate

the difference between the *nature* of GD and BIID medical proce-
dures: BIID medical procedures intrinsically involve causing the
patient to become disabled, while GD medical procedures do not.
However, as will become apparent further in the essay, there is
more than one way to interpret this difference, and, depending on
the analysis that one offers, it can directly influence the strength of
the pro-BIID-medical-intervention argument.

In Section I, I provide a more robust understanding of BIID
and GD, and I disentangle the medical condition of GD from
the social self-identification of 'transgender' and 'transsexual'.
In Section II, I explicitly detail the argument by analogy for the
acceptance of BIID medical procedures via an appeal to the accep-
tance of GD medical procedures. Section III then sees an initial
critique of this argument, as well as my response to some antici-
pated objections to this position. In Section IV, I bring in consider-
ations from disability studies to help illuminate the sense in which
some of the ideas in the previous section may be seen as unjustly
privileging abled bodies, and I briefly evaluate the social dimen-
sion of disability. Lastly, in Section V, I situate this entire analysis
within the broader bioethical discourse, and I briefly argue for a
way of doing bioethics that incorporates the contributions of both
trans philosophy and philosophy of disability.

Section I: What BIID and GD Are, and What They Are Not

BIID is a condition in which an individual desires to "either be
paralyzed or to have one or more of their healthy limbs … ampu-
tated" (Blom, Denys, and Hennekam 2012, p. 1). This desire
of these individuals is fueled by a mismatch between the actual
bodies with that of their perceived, or requested, bodies (ibid.). In
other words, BIID patients suffer from an identity disorder: they
perceive their body to be other than the way that it *actually* is
(ibid.). So, for instance, a BIID patient may express the intense

desire to amputate their left leg, with the main motivating factor behind this desire being the need to bring their actual, non-amputated/two-legged body into accord with their perceived, amputated/one-legged body.[2]

GD, on the other hand, is a condition in which an individual desires to be, be seen, or be treated as another gender, or to remove and replace one's sex characteristics with that of another gender (APA). This desire originates from the mismatch between the assigned-at-birth gender of the individual with the gender that they identify as (ibid.). GD patients, therefore, also appear to be suffering from an identity issue: their *actual* sexed body conflicts with that of their *perceived* sexed body.[3] So, for example, a GD assigned-female-at-birth patient may strongly express the need to take testosterone pills and have surgery in order to remove their breasts and to remedy the conflict between their actual, female-sexed body and their perceived, male-sexed body.

So far, the two concepts that have been discussed—BIID and GD—have been medical concepts: they attempt to capture a certain phenomenon within a medical context. However, I also believe it important to briefly discuss the concepts of 'transgender' and 'transsexual', which are arguably more akin to social concepts, and to distinguish them from GD in particular.

First, while it is tempting to simply equate the notion of 'transgender' with that of GD, such an equivalence would be fallacious. If an individual is transgender, that simply means that the individual identifies with (or presents, or lives, as) a gender different from the one which they were assigned at birth (ibid.). So, an individual could be transgender but not be suffering from GD, insofar as they do not experience distress, conflict, or any other sort of intense mental suffering alongside their gender identification (or presentation).

Depending on how familiar one is with either feminist philosophy or LGBTQ+ communities, this understanding of the term may be somewhat surprising, since it seems that the medical

and social history of transgender people has often included reference to their mental distress. To some, the notion of 'transsexual' is more familiar, as it has been featured in the medical terminology, and in daily conversation, for far longer than 'transgender' has. Roughly, if an individual identifies as transsexual, this means that the individual has undergone some medical intervention or procedure in order to construct their body in such a way that either merely goes against their sexed birth assignment or that is in alignment with their identified (non-birth-assigned) sex (Bettcher 2007, p. 46). Given this understanding of the term, one may be tempted to instead equate 'transsexual' with GD, since they both seem to denote a medical phenomenon—or, at the least, that identifying as transsexual requires that one suffer from GD. Unlike the previous attempt at equating between 'transgender' and GD, I am hesitant to outright reject this particular theoretical move, since the varied and dynamic history of 'transsexual', along with the current contestation of its meaning within LGBTQ+ circles, seems to invite more interpretation to the term than to 'transgender'—at least with respect to its medical connotations.

For the purposes of the following argument, however, it will not be necessary to delve deeper into this issue on the meanings of these two terms, since the argument relies on neither in order to make its point. However, these terms will come into play further in the essay when I discuss the social dimension of the argument, since that *will* involve a discussion of the potential *social* issues that may arise from accepting this analysis.

SECTION II: AN ARGUMENT FOR BIID-RELATED MEDICAL PROCEDURES

It seems that one can argue *for* the permissibility of BIID-affirming medical interventions by appealing to the apparent similarities between BIID and GD and the permissibility of doctors to treat GD with body-altering procedures. That is, one could make the

following argument:

(1) It is morally permissible for medical doctors to treat GD with body-altering medication and technology.

(2) The medical situation of BIID patients is similar in kind to the medical situation of GD patients.

∴ It is morally permissible for medical doctors to treat BIID patients with body-altering medication and technology.

Given the validity of the argument, which, if any, of the premises is false? Well, premise (1) appears to be true. Doctors are allowed to prescribe, for instance, hormone replacement therapy (HRT) to sufferers of GD who desire to have the secondary sex characteristics of another gender (APA). They are also able to perform gender-affirming operations—i.e., surgical procedures that make the bodies of GD patients resemble that of their identified gender—in order to alleviate the suffering brought upon by GD (ibid.). This, of course, does not entail that doctors always obey the demands of GD patients, just as they do not unquestioningly agree to any other medical procedure that a stubborn patient may seek. For the most part, however, GD patients are treated in the manner that the patients themselves desire to be treated. So, premise (1), I grant, is true.

Premise (2), on the other hand, is more dubious, primarily because of its emphasis on the "medical situation" of both types of patients. The motivation behind accepting premise (2) as true originates from the aforementioned parallels between BIID and GD patients. For instance, both BIID and GD patients experience a conflict between what they perceive to be the case about their body and what is actually the case. Moreover, BIID and GD patients both experience psychological distress as a result of this internal conflict and seek to remedy this distress via medical intervention, with the desired procedures typically involving a change in the bodily features of these patients. It is also uncontroversial

to claim that both types of patients often declare their interest in these body-changing medical procedures autonomously, and that they do so with the knowledge of the consequences—intended or unintended—of these procedures. And, if doctors were to go through with these procedures, both BIID and GD patients would experience euphoria at this bodily alignment.[4] Indeed, if these features exhaust the features of the symmetrical situation referred to in premise (2), then this premise also appears to be true. I deny, however, the exhaustivity of these features, since there seems to be at least one more relevant factor worth considering: the state of the patient after the medical procedure.

SECTION III: INITIAL RESPONSES AND OBJECTIONS

In the case of BIID, patients who undergo their specified medical procedure, and find that their BIID is alleviated, no longer need to deal with the incredible psychological distress that they used to experience. They now, however, must deal with being physically disabled and must learn to live with this self-imposed difficulty. So, if Ted, a sufferer of BIID and a part-time university lecturer, used to bike to work, but underwent a BIID-influenced leg amputation, he would need to secure a different, disable-friendly mode of transport. And, if Julia, a BIID patient and single mother, was the sole caretaker and provider for her family, but willingly became disabled, she may need to hire a nanny or two in order to keep her house, and her children, under adequate control.

The same cannot be said in the case of post-medical-procedure GD patients. Insofar as the medical procedure requested by the sufferer of GD is one that has, as its goal, the sexed body of the gender other than the one they were assigned at birth, the GD patient's medical procedure will not intentionally bring about disability. Whether it be a bilateral mastectomy, breast augmentation surgery, or hormone replacement therapy, none of these

three will cause the sufferer of GD to become physically disabled: at worst, it will only physically impede them for a short period of time as they recover from the bodily changes (Scheefer-Van Boerum, Salibian, Bluebond-Langner, and Agarwal 2019).[5] Therefore, given this apparent difference in *kind* between the state of BIID and GD patients post-medical intervention, it follows that the medical situations of both types of patients are dissimilar, and that the conclusion of the previously mentioned argument—i.e., the moral permissibility of body-altering medical interventions for BIID patients—fails to be established.

I now want to consider two different *prima facie* objections to this response. The first objection is that it is false that post-medical-intervention GD patients never suffer adverse or unintended side-effects that may end up physically disabling them. One can easily conjure up scenarios in which a gender-affirming operation can go wrong, or how a side-effect of hormone replacement therapy may unfortunately end up occurring, despite the low probability of the occurrence of such events. While this is true of GD-related medical interventions—perhaps more so than other medical procedures and operations due to the relative infancy of the subspecialty—it is not *intrinsic* to these operations that they bring about physical disability. To make the point clearer, one can conceive of a future state of medicine in which side-effects are completely nullified, and acknowledge that, in such a world, GD-related medical interventions would not bring about such permanent physical hindrances. In such a scientifically advanced world, however, the same could not be said about BIID-related medical interventions, since it is intrinsic to these operations that they *do* bring about these hindrances.

Secondly, one may agree that GD-related medical procedures do not intrinsically bring about physical disability, but that it is intrinsic to these procedures that they nonetheless bring about a sort of sex-based, or reproductive, disability. For instance, one may argue that the infertility that results from some GD-related

medical interventions, such as hormone replacement therapy or some gender-affirming operations, permanently hinders the individual, and, as such, is not entirely unlike the situation of BIID patients. While this objection does appear to be more substantive, at least *prima facie*, than the previous one, I do not think that it can recover the truth of premise (2) for two reasons.

First, it would need to be demonstrated that becoming physically disabled is neither better nor worse than becoming infertile. Insofar as either one of these hierarchies is reasonable to assume, then it becomes difficult to maintain that the situation of GD and BIID patients is similar in kind. If this similarity fails to hold, then the argument fails to be sound. Second, even if one grants that becoming infertile *is,* in some sense, akin to living with other physical disabilities, one can still argue that it need not be intrinsic to GD-related medical procedures. Returning to the "future world of medicine" scenario, one can easily conjure up a scenario in the near future where, for instance, GD-related medical procedures for patients that desire a female-sexed body also involves uterine transplantation, which would therefore preserve the fertility of the patient. So, since it is merely because of the current state of medical technology that brings out infertility in patients undergoing GD-related medical procedures, and because we are granting that infertility is some sort of disability, then it is not intrinsic to *all* such procedures that they bring about a disability, since eventually no such procedure may do so. It therefore seems that the original argument—the argument that attempts to establish the permissibility of BIID-related medical procedures via reference to the permissibility of GD-related medical procedures—still fails to be successful. Or does it?

Section IV: Ableism and The Social

If the above analysis were to end with my brief rebuttals to these two objections, it would seem that the original argument does

indeed fail, and that individuals like Neil, Bayne, and Müller would need to do more work in order to establish their pro-BIID-medical-intervention position, at least with respect to this specific take on the issue. However, to do so would be to accept an ableist attitude towards the disabled, and it would involve ignoring the sense in which this issue is closely intertwined with certain social and political frameworks. In this section, I will elaborate on the former claim.

The analysis presented in Section III *in its current state* is ableist precisely because it positions the disabled body as something that is categorically different than the abled body. This is evident by the use of the term 'intrinsic' in describing disability and by claiming that this therefore points to a difference in kind between BIID and GD patients. However, this model of disability is one that needs further argument to establish, since, upon further investigation, it is not clear that one should understand disability as something entirely intrinsic to an individual.

In his (2016) article, Joel Michael Reynolds attempts to put forth a critical theory of harm that is informed by the discussions occurring in critical disability studies and philosophy of disability circles. As he evaluates the ways in which the notion of harm is used in anti-BIID-medical-intervention arguments, he notices that a fair amount of the available literature, which is not much to begin with, on the ethical justification of medical intervention for BIID patients is implicitly ableist (Reynolds 2016, p. 40). Reflecting on a line of reasoning similar to the one I put forth previously, Reynolds remarks that "[it] is ableist, to be clear, insofar as it assumes, without evidence or argumentation, that the "standard" able-body is, *ceteris paribus*, in and of itself better than the non-standard, disabled body" (ibid.). Reynolds further comments that this way of conceiving of disability ignores the social dimension of such a condition, and that, once one accepts the distinction between *impairment* and *disability*, it becomes harder to establish this hierarchy of bodies, and, I would argue, to establish the anti-BIID-

medical-intervention argument.

Roughly, in philosophy of disability circles, "impairment" simply refers to the particular embodied state or condition of an individual, while "disability" refers to the social dimension that accompanies a particular impairment (Reynolds 2016, p. 39). So, for instance, one would say of an individual that is paralyzed from the waist down that their impairment is the loss of leg function, while their disability is the difficulty with getting around without the use of a wheelchair, or their unequal access to certain facilities, or the stigma they must deal with when attempting to enter the workforce, and so on. Once one accepts this distinction, the next move would be to claim that the reason why some may be against BIID-related medical procedures is because they conflate these two terms and mistakenly believe that BIID patients seek *disability*, not *impairment* (Reynolds 2016, p. 39).

Notice that, while this does seem to suggest that the issue with BIID-related medical procedures is not that they bring about an intrinsic physical impediment, but rather that they bring about an embodied state which carries with it problems that are ultimately social issues, it does not establish this position *on its own*. This is because it would still need to be demonstrated that these problems solely are social in nature: one would need to establish a purely social, as opposed to a medical, model of disability.[6]

The problem of deciding between social and medical models of disability is one that is still being debated in the literature, which makes it somewhat more difficult to apply it to the current discussion at hand. However, here I would take a stance similar to the one Richard Dean (2018) takes in his essay on the issue of "curing" autism within the neurodiversity movement. In the essay, Dean admits that the strength of the arguments for and against rejecting cures for people on the autism spectrum ultimately depends on whether or not one should view autism as a medical or social issue (Dean 2018, p. 129). If one accepts a medical model of disability, then a disability is a problem intrinsic to an individual, and would

therefore be best addressed via medical means (ibid.). On the other hand, if one accepts a social model of disability, then the problems that arise from a disability are due to the way society treats and accommodates such people, and that, therefore, it would be best addressed by changing society, not the individual (ibid.).

How can one determine whether a certain disability ought to be viewed as a medical or social issue? Well, in the case of making this determination for autism, Dean decides to employ Anita Silver's (2005) "historical counterfactualizing test" to determine whether society's treatment of a particular disability, such as autism, is justified. One could run this thought experiment by imagining the majority in a society to have the disability in question, and then seeing if the proposed treatment is also adopted by the members of this society. If it is adopted, then the treatment is just, but if it is not adopted, then it is unjust, and believing it to be otherwise would be indicative of one's ableist attitudes. Furthermore, if the treatment is justified, and if it only involves changes to society, then, according to Dean, this would seem to lend support to the idea that the disability in question is a social, not a medical, issue, and should be treated as such (Dean 2018, p. 130).

Having now laid this out, what does this test suggest for the disabilities that BIID patients desire? Well, one initial difficulty concerns the fact that there are multiple disabilities that fall under the BIID umbrella, and, as such, merely running the test once would not suffice to solve the issue. However, since amputation and paralysis seem to be the most cited disabilities among BIID patients, those are the two that I will focus on.

So, what would a society consisting mostly of people with amputations look like? Well, one could imagine that there would be a fair amount of motion detection technology implemented in order to help people with, say, upper-limb amputations enter rooms and buildings, and that there would be more effort put into creating mobility assistive equipment for people with lower-limb amputations. Similarly, a society consisting mostly of people with

different degrees of paralysis would probably feature a similar reliance on such mobility assistive equipment, and in assistive technology in general. Besides this reliance on technology, it appears difficult to imagine what disadvantage these particular conditions would grant to an individual that could not be addressed via societal changes and restructurings.[7]

Given these thought-experimental results, adopting a social model of disability, with respect to the above-mentioned conditions, would seem to be justified. Moreover, if we wanted to bring in the impairment/disability distinction, if one accepts the social model, then one can respond to the "difference in kind" response to the pro-BIID-medical-intervention argument by rejecting that such a difference exists: an impaired body, while one that instantiates one of many particular embodied conditions, is not intrinsically disadvantaged when compared to a non-impaired body, since the disability that accompanies the impaired body only does so contingently. Notice, too, that this allows one to respond to the rebuttals to the two objections entertained in Section III. If one refers to some imagined "future world of medicine" in order to defend against the idea that the infertility that can result from certain GD-related medical procedures is not intrinsic to them, then one could also refer to this same possible world and claim that, in such a world, the social problems tied to amputation and paralysis would not exist, and so the disability (understood here in terms of the impairment/disability distinction) that can result from BIID-related medical procedures are also not intrinsic to them.

It therefore seems to be the case that the original pro-BIID-medical-intervention argument—i.e., the argument given in Section II—is sound after all, and that medical professionals should comply with the request of BIID patients on these grounds. However, before I end, I do want to briefly discuss an issue that seems to be lying in the background of these discussions surrounding marginalized groups in society, since this is not something unique to the BIID-GD analogy; indeed, this problem

can crop up in other areas of philosophy.

SECTION V: REFLECTING ON TRANS PHILOSOPHY AND IMPROVING BIOETHICAL DISCOURSE

Trans philosophy, as a sub-discipline of trans studies, involves the theorizing *of* trans people *by* trans people and trans-studies-informed people in a way that is sensitive to the oppression that trans people face in society (Bettcher 2019). This arose from the "philosophizing of trans phenomena" in the mid-20th century, which saw the theorizing of trans people be done by solely non-trans people, and which paid no attention to the oppression that their theorizing could bring onto the trans population (ibid.). Were it not for philosophers like Sandy Stone, Riley Snorton, and Talia Bettcher, the medical account of transsexuality would perhaps still be the assumed account of all trans experiences, since it was their work and insight *as trans people* that helped break this all-encompassing model. And, if one is not trans, then any theorizing that one does that involves this demographic ought to at least be acquainted with the relevant trans philosophical literature, in order to at least guard against charges of transphobia and intellectual dishonesty.

The reason why I bring up these features of trans philosophy is to emphasize the need for a similar style of theorizing within bioethics on the BIID issue. If one is ignorant of the ways in which ableism can be implicitly pushed through the backdoor of an argument, then it can lead to the further marginalization of people with disabilities. This is why it is important that work that explicitly engages with one's ableist assumptions and attitudes be included in the discourse, since not all bioethicists are aware of disability-based oppression and resistance—there is often no social need for them to be.[8]

On the issue of oppression, I do want to make one last clarificatory note. Some may take what I have said in this essay and

14

equate my style of reasoning with the sort of reasoning employed by Rebecca Tuvel in her problematic (2016) "In Defense of Transracialism" paper. In her paper, Tuvel attempts to make an argument for the acceptance of so-called "transracial" people in society by making an analogy to the social acceptance of transgender people. In other words, if transgender people are seen as valid and accepted in society, and if the identity claims of transgender people are similar in kind to so-called "transracial" people, then ought not so-called "transracial" people also be seen as valid and accepted in society?

While a complete analysis of Tuvel's paper exceeds the scope of this essay, I do want to highlight one reason as to why her paper is problematic, and how my paper avoids this same charge. In her paper, she attempts to argue for "transracialism" in a somewhat theoretical manner; she does not critically engage with the literature on either race studies or trans studies, focusing on instead performing her philosophical treatment of both social groups in a vacuum. This is problematic because both groups, in virtue of being *social* groups with social and political ties, demand that at least some of one's theorizing involve acknowledging the sense in which it may either invite oppression or support resistance for the groups being theorized about.

This being said, my analysis is unlike Tuvel's for two reasons. First, while it does involve discussion of certain groups, namely, that of transgender/transsexual people and impaired people, the main focus of the pro-BIID-medical-intervention argument is on the medical notions of BIID and GD, not the social notions of transgender/transsexual or disability. However, I of course acknowledge that the medical notions of BIID and GD have close ties to the social notions of transgender/transsexual and disability, which leads to my second reason: my analysis is not ignorant of the relevant trans philosophy and philosophy of disability literature. Indeed, had I stopped at Section III, then the ableist charge would have been well-founded, and my analysis would have fallen

under the category of oppressive essays. However, it is because I am against ableism, as well as other forms of oppression, that I believe that the pro-BIID-medical-intervention argument is ultimately sound.

CONCLUSION

I began this essay by first making clear the notions that I would be using, namely, BIID and GD, and making clear how I would be using them in the pro-BIID-medical-intervention argument. I then gave the pro-BIID-medical-intervention argument that I see being implicitly referred to in the literature, and I evaluated some responses and objections to the argument. I then focused my attention on the ableist assumptions of some of the responses, and I argued for a social model of disability with respect to two of the disabilities that fall under the scope of BIID. With this social model in hand, I then reevaluated the pro-BIID-medical-intervention argument and determined that, if one wants to take the anti-ableist stance, one ought to accept the argument as sound. However, I then discussed the importance of being socially aware of the oppressive and resistant dynamics that are at play in discussions that overlap with trans studies and disability studies, and, through a comparison with one such problematic type of argument in the literature, attempted to establish my analysis as an instance of this cross-disciplinary work done correctly. Of course, this is not to say that it could not be improved or expanded upon, since I do think that further work on both disability and trans studies will likely continue to affect the strength of this argument. This much, however, is certain: one can accept the consequences of my argument while also remaining a strong ally to marginalized groups.

Notes

1. While I do think that the argument is hinted at in Müller's (2009) article, there are moments in other articles (such as her other (2009) paper) where she seems to take an anti-BIID-medical-intervention stance similar to the

one entertained here in Section III. Regardless, the point being made here is merely that pro-BIID-medical-intervention arguments are discussed in the literature, and, as such, are worthy of further analysis.

2. I believe it worth briefly mentioning that the scientific research concerning BIID is still relatively nascent. As such, it is not the only theory that attempts to make sense of the phenomenon of individuals seeking medical intervention for elective amputation or paralysis. For a discussion of some of the other related theories in this area, see Müller (2009).

3. I merely use 'sexed' here in order to emphasize that it is the body *as a sexed body* that is fueling this apparent crisis in bodily identity.

4. Some BIID patients, for instance, expect feelings of completion and inner satisfaction upon undergoing their desired medical procedure (Blom, Denys, and Hennekam 2012, p. 2).

5. Of course, from any surgery there is a risk of complications that could eventually result in death. However, putting that possibility aside, this point stands.

6. The reason that it would need to be a purely social model is so that it could defend against arguments that assume there to be an intrinsic difference (i.e., a difference in kind) between an abled and a disabled body. Insofar as at least some of it is not social, then it seems to leave one's position susceptible to this line of attack.

7. One may object to this position on grounds that it ignores the further significant medical difficulties that people with amputations sometimes face after their surgery, perhaps even for the rest of their lives. While I agree that this is an issue in the current medical context, it does not seem unreasonable to think that, in a society where the majority of the population were physically impaired, these medical difficulties would either be drastically reduced or eliminated entirely. I am open to further discussion on this point, however.

8. This is because, unless one is a member of a marginalized social group, it is unlikely that one will be aware of all the ways in which one's theorizing may disproportionately and negatively impact such groups. Furthermore, because one is not on the receiving end of these disproportionate and negative effects, then, in most cases, one has no immediate social need to address it.

Bibliography

Bayne, Tim, and Levy, Neil. (2005) "Amputees by choice: Body Integrity Identity Disorder and the ethics of amputation," *Journal of Applied Philosophy* 22(1), pp. 75-86

Bettcher, Talia. (2007) "Evil deceivers and make-believers: On transphobic violence and the politics of illusion," *Hypatia* 22(3), pp. 43-65

Blom, Rianne, Denys, Damiaan, and Hennekam, Raoul. (2012) "Body Integrity Identity Disorder," *PLoS ONE* 7(4), pp. 1-6

Scheefer-Van Boereum, Melody, Salibian, Ara, Bluebond-Langner, Rachel, and Agarwal, Cori. (2019) "Chest and facial surgery for the transgender patient," *Translational Andrology and Urology* 8(3), pp. 219-227

Dean, Richard. (2018) "Neurodiversity and the rejection of cures," *Disability in Practice: Attitudes, Policies, and Relationships*, Oxford Scholarship Online, pp. 115-133

Müller, Sabine. (2009) "BIID—*Aqua fortis* for scientific explanations of psychic phenomena?," *The American Journal of Bioethics* 9(1), pp. W3-W4

_____ (2009) "Body Integrity Identity Disorder (BIID)—Is the amputation of healthy limbs ethically justified?," *The American Journal of Bioethics* 9(1), pp. 36-43

Reynolds, Joel Michael (2016) "Toward a critical theory of harm: Ableism, normativity, and transability (BIID)," *Philosophy and Medicine* 16(1), pp. 37-46

Silvers, Anita. (2005) "People with disabilities," in: *Oxford Handbook of Practical Ethics*, pp. 301-327 (Oxford: Oxford University Press)

Tuvel, Rebecca. (2017) "In defense of transracialism," *Hypatia* 32(2), pp. 263-278

"What is gender dysphoria?," (February 2016). *The American Psychiatric Association* [I use (APA) to refer to this citation in the essay]. https://www.psychiatry.org/patients-families/gender-dysphoria/what-is-gender-dysphoria

WHAT IT'S LIKE TO HAVE AN ALIEN THOUGHT: THE EXPLANATION/ ENDORSEMENT PROBLEM IN THE DELUSION OF THOUGHT INSERTION

Steven Diaz

INTRODUCTION

It is not uncommon for people to hold bizarre beliefs. There are people who maintain that the earth is flat despite being shown images from space that depict the earth as spherical. There are also people who believe that President Obama is not a native United States citizen, even though he released his birth certificate as proof of his citizenship. Whether these cases count as delusions is up for debate, as delusions are typically a symptom of an unhealthy mind. There is no mental disorder that elicits more curiosity from both philosophers and the general public than schizophrenia.

From a philosophical standpoint, the delusion of thought insertion has received a considerable amount of attention. Thought insertion can be understood as the delusion that one's thoughts are not one's own. The delusion is classified as a first-rank symptom of schizophrenia and typically warrants a diagnosis of schizophrenia. However, thought insertion also occurs in patients who experience mood disorders, atypical psychoses, and in some cases, can occur within people who do not have a mental disorder—in other words, within the general public or as non-clinical cases (Mullins and Spence 2003; Rössler *et al*, 2007). What makes thought insertion so interesting is that it challenges many of our philosophical assumptions, for instance, that having introspective access to a thought means one will recognize the thought as their own.

The purpose of this paper is to articulate what I take to be the right way of thinking about thought insertion and to show that

there is not an obvious answer as to whether the delusion is an endorsement or an explanation of the abnormal experience. While the endorsement/explanation dispute has received attention with regard to other delusions, the subject appears to have received less attention when it comes to thought insertion. In the limited cases where philosophers do focus on the dispute, my intuition is that they fail to consider all the possible characterizations of the abnormal experience. A further issue is to what degree—if any—can the contents of the delusion differ from the abnormal experience and still be an endorsement of the experience. Additionally, there is also an issue of the conditions in which the delusion is formed, where these conditions may present further complications in deciding whether the delusion is an endorsement or explanation of the experience.

In the first section of this paper I will discuss what thought insertion is. I will offer some reports of thought insertion, highlight some important features of thought insertion that are sometimes glossed over within the philosophical literature, and contrast thought insertion with other delusions and hallucinations that schizophrenics sometimes experience. In the following section, I will briefly argue for an empirical account of thought insertion, an account which takes the delusion as a response to an abnormal experience. I will then, in the third section of this paper, articulate different ways of thinking about the experience associated with thought insertion as a set up to the problem. Following this, I then argue—in the fourth section of this paper—that it is not obvious whether we ought to adopt an endorsement view of thought insertion or an explanation view of thought insertion. Finally, I suggest some possible ways of addressing the issues I present in the previous section, before offering some concluding remarks.

WHAT IS THOUGHT INSERTION

Thought insertion can, tentatively, be understood as having two

main features: the first is the abnormal experience associated with a particular thought—in the literature, this is called an *alien thought*. It is important to point out that not every thought is taken to be an alien thought; only *some* of the subject's thoughts are associated with the abnormal experience. While the alien thought may sometimes have graphic or violent content like "Kill Bill," this is not always the case—sometimes the alien thought may be something like "It is cold in here." The second feature is the response to the alien thought—this is typically regarded as the *delusion*.[1] For instance, given the alien thought, one might believe that "this is not my thought." It is also important to point out that it is not always the case that the alien thought is attributed to a specific person, object, or agency; sometimes the thought is taken as belonging to someone (or something) else, but the subject is unsure as to whom the thought belongs to.

Reports of thought insertion widely vary, and this can sometimes make it difficult to distinguish thought insertion from other delusions and hallucinations:

> [S]he said that sometimes it seemed to be her own thought 'but I don't get the feeling that it is.' She said her 'own thoughts might say the same thing', 'but the feeling isn't the same', 'the feeling is that it is somebody else's. (Allison-Bolger 1999, #89 in Hoerl 2001)

> The patient might report that their thoughts are not experienced in the usually silent manner but rather heard spoken aloud with the patient's own voice… or seen as subtitles in a movie, requiring the patient to listen to or read [the] thoughts in order to know what [the thought is]. (Henriksen *et al*. 2019, p. 6)

> Thoughts are put into my mind like "Kill God". It's just like my mind working, but it isn't. They come from this chap, Chris. They're his thoughts. (Firth 1992, p. 66)

I didn't hear these words as literal sounds, as though the houses were talking and I were hearing them; instead, the words just came into my head—they were ideas I was having. Yet, I instinctively knew they were not *my* ideas. They belonged to the houses, and the houses had put them in my head. (Saks 2007)

[After the interviewer asks whether the patient experiences alien thoughts, having denied experienced thoughts being 'inserted' into their mind]: Yes, sometimes it is like they are… when the thoughts are kind of solemn thoughts or, how to put it, then I can get the feeling that they have been sent from another place, from elsewhere. Because, if they are not mine, and they are solemn thoughts, then they must be something special.… They are very different from my usual thoughts and are thoughts that other people don't think. (Jansson & Nordgaard 2016, p. 37)

As you can see, the reports vary in different respects. It is easy for philosophers to sometimes confuse thought insertion with auditory verbal hallucinations, cases where one "hears" a voice that is not their own, either internally or externally. However, we should resist this mistake, as some of the reports of thought insertion either suggest that the thought was not heard or that the thought was represented visually. Philosophers will also sometimes mistake thought insertion for the delusion of thought control. In the case of thought control, the subject claims that their own thoughts are under the control of someone else—i.e., someone else is forcing the subject to think certain thoughts. The difference between these two delusions is that, in the case of thought control, the subject still identifies the thoughts as their own, whereas, in the case of thought insertion, the thoughts belong to someone else. A further distinction that is made within the philosophical literature (which I find to be helpful) is the difference between thought insertion and thought broadcasting. In the case of thought broadcasting,

a subject believes that their thoughts are made public; they are under the impression that their thoughts are no longer private but accessible for others. Thought insertion can be understood as the opposite, where the subject believes they have (unwanted) privileged access to the thoughts of others.

With these distinctions and features of thought insertion in mind, we can now briefly articulate why we should adopt an *empiricist* view with regard to the delusion of thought insertion. The reason for only briefly discussing this is, on my understanding of the philosophical literature on thought insertion, it is widely accepted that the delusion is a response to an abnormal experience. So, I will not focus too much on defending this position; however, it is important to discuss it since the central issue of the endorsement/explanation dispute has to do with whether the phenomenal (or experiential) content is vague or *rich*—it is an issue of whether the content of the delusion is the same as the content of the abnormal experience.

WHY THE DELUSION IS A RESPONSE TO AN ABNORMAL EXPERIENCE

I think the natural intuition most people have when they learn about thought insertion is that the delusion is a response to an abnormal thought. However, within the debate on delusions in general, there is disagreement over whether a delusion is caused by an abnormal experience or whether an abnormal experience is caused by the delusion. It should become clearer what I mean by this at the end of this section.

In the case of delusions, one can adopt either an *empiricist* approach or a *rationalist* approach. A proponent of an empiricist approach argues that the delusion is a response to an abnormal experience. Whereas, a proponent of the rationalist approach will argue that a modification to the belief system is what causes the delusion (Bortolotti 2010). Put differently, the rationalist might

argue that a delusional belief is formed (and adopted), and given that the subject holds this delusional belief, the delusion affects their experiences and other beliefs. A helpful way of understanding this, I take it, is offered by Max Coltheart and company. The issue comes down to what is the source of the delusion (Coltheart *et al*. 2010). On both the rationalist and empiricist view, there can be some abnormality—either some abnormal experience, abnormal cognitive process, or abnormal neural process. The issue is that the empiricist approach takes it that the delusion is a response to some abnormal experience; the rationalist might, however, argue that the delusion was formed by bad unconscious inferencing or by some abnormality in the brain.

In the case of thought insertion, if the delusion is not caused as a response to an alien thought, then the delusion is the alien thought, and the alien thought is caused by some other abnormality. There are reasons for rejecting this idea. First, if one takes delusions to be belief-like, then an alien thought cannot be a delusion.[2] As Lisa Bortolotti points out, subjects who suffer from thought insertion can not only fail to ascribe the thought to themselves (they do not take themselves to be the thinker of the thought) but also fail to endorse the alien thought (they reject it). For example, the subject may have the alien thought "Kill Bill" or "No one likes you," yet disagree or fail to act on these thoughts (Bortolotti 2010; Bortolotti and Broome 2009; Bortolotti and Miyazono 2015). Furthermore, many of us have—at some point—had a strange or intrusive thought. We might think "I'm going to kill him" (and presumably, immediately reject this thought). Strange thoughts do not, on their own, make someone delusional. In the case of thought insertion, the delusion is the belief "This is somebody else's thought."

I take it that the best way—and the way that lines up best with our commonsense intuition—is to think about thought insertion as the delusion being a response to a thought associated with an abnormal experience. Assuming one agrees with this charac-

terization of thought insertion, we can ask whether the delusion is an endorsement of the abnormal experience or whether it is an explanation of the abnormal experience. In order to answer this question, we need to know what the abnormal experience is like.

WHAT IS IT LIKE TO HAVE
AN ALIEN THOUGHT?

While I take it that most people would agree that the delusion of thought insertion is a response to an alien thought, I imagine that it is not obvious to most people what the abnormal experience is. Without having gone through thought insertion, it is difficult for someone who has not had an alien thought to imagine what the experience is like.

Within the philosophical literature, there are at least three different characterizations of the experience that I have identified. While I am personally skeptical that any of these characterizations are the correct characterization, for the purposes of this paper we can treat all of them as potentially the correct characterization of the abnormal experience:

1. A lack of *a sense of agency* (Graham 2004; Stephens and Graham 2000; Campbell 2002)

2. A lack of *a sense of ownership or mine-ness* (Henriksen *et al*. 2019; Bortolotti 2010; Bortolotti and Broome 2009; Pacherie and Martin 2013)

3. The presence of a *sense of volition* (Currie 2000; Currie and Ravenscroft 2002)

In terms of the first proposal, proponents of this view argue that all our thoughts are associated with a sense of agency and a sense of subjectivity.[3] In other words, *all* our thoughts are associated with a feeling that we produced the thought and a feeling that the thought is privately accessible. In the case of thought insertion,

what happens is that we have some thoughts which lack a sense of agency—i.e., alien thoughts. While the subject produces the thought, they fail to have the experience that they produced the thought; it is due to this lack of experience that the thought seems alien. On the second proposal, the idea is, roughly, something like *every* thought is associated with a sense of ownership (or mine-ness).[4] In the case of thought broadcasting, subjects *feel* as though the thought is publicly available but still identify the thought as their own —it still *feels* like their thought. In the case of thought inser-tion, the subject still experiences thoughts as something private—whether the thought is alien or not—but *feels* as though some of the thoughts are not their own. In other words, alien thoughts lack mine-ness and are therefore confused as being someone else's. The third proposal is the least popular, but worth including as it shows that the abnormal experience may not be the result of something missing, but something added to our thoughts. On this proposal, imaginings are taken to be associated with a sense of volition—we can will when we imagine and will what we imagine. The subject confuses an imagined proposition for a normal thought, and the alien thought is just an imagined proposition.

We can visualize the difference between the three proposals by the following diagram:

	Non-Delusional Subject		Subject with Thought Insertion	
Type of Content	Propositional	Phenomenal (Experiential)	Propositional	Phenomenal (Experiential)
Proposal 1	"Mars is Red"	1. Sense of Subjectivity 2. Sense of Agency	"Mars is Red"	1. Sense of Subjectivity
Proposal 2	"Kill God"	Mine-ness	"Kill God"	
Proposal 3	"Bill is Tall"		"Bill is Tall"	Sense of Volition

I am arguing that whether an endorsement view or explana-tory view seems correct will in this case depend on how we choose

to characterize the abnormal experience. This is because the issue is whether the delusion is an endorsement of some "rich" experiential content, or whether it is an explanation or interpretation of some "vague" experiential content. I agree with Ian Gold and Jakob Hohwy that "the alien quality of the delusional experience is part of its content, and it is the content of the experience that is the locus of the delusion…" (Gold and Hohwy 2000, p. 162). The delusion is, primarily, a response to the abnormal experience associated with the thought.

This suggests (to me at least) that if the delusion is understood as being a belief or belief-like, then there is a phenomenal belief involved. The subject distinguishes some of their thoughts as alien, such as the thought "Kill God." In distinguishing the non-alien thoughts from the alien thoughts, I think it is reasonable to suppose that the subject has a phenomenal belief with some demonstrative content like *This thought* and some phenomenal content or property like "feels like (such-and-such) experience." For example, the subject comes to believe "*This* thought ("Kill God") feels *unfamiliar*" or believe "*This* thought ("The grass is cool") feels *volitional*." What the experiential content is will be important for whether we ought to take the delusion as an endorsement of the content, or whether the delusion explains the content.

ENDORSEMENT OR INTERPRETATION?

Within the debates of delusions in general, there is an issue of whether a delusion should be thought of as an endorsement or an explanation of the abnormal experience. On the endorsement view, the contents of the delusion match the contents of the experience. The delusion is seen as an endorsement of the experience, where the subject takes the experience at face value. On an explanatory view, the contents of the delusion do not match the contents of the experience. The experience is taken to be vague and the delusion acts as an explanation of the content (Bortolotti 2010).

In order to see the difference, it may help to consider the issue of whether we endorse or explain our normal perceptual experiences and whether other delusions, such as Capgras delusions, are an explanation or an endorsement. In his defense of the endorsement view of thought insertion, Michael Sollberger points out that our beliefs about our perceptual experiences are usually taken to be an endorsement of that experience. For instance, if I see a red cup on the table, then I believe that there is a red cup on the table; our perceptual belief takes our perceptual experience at face value (Sollberger 2014). In his argument for an explanatory view of delusions, Max Coltheart elaborates on what an endorsement view of Capgras delusions would be like. Patients with Capgras delusions believe that a specific individual—generally someone close to the patient, like a spouse, family member, or dear friend—has been replaced with an imposter. In the case of someone without the delusion, they may have some perceptual content "This is my wife" when seeing their wife and endorse this content. In the case of someone with Capgras delusions, the patient may see their wife but have the perceptual content "This is not my wife" and endorse this content. Yet, Coltheart argues, the problem is how one could acquire the perceptual content "This is not my wife" without inheriting it from an already existing delusional belief that the person in front of the patient is indeed not their wife (Coltheart 2005).

In this case, the Capgras delusion is an explanation for the abnormal experience. However, as Sollberger points out, thought insertion presents different challenges than perception—for example, the delusion is not concerned with an external object (Sollberger 2014). Furthermore, thought insertion is taken as a response to an abnormal experience, whereas for Coltheart and company, Capgras delusion is taken to be a response to an abnormality originating within the brain (Coltheart *et al*. 2010).[5] Since the delusion of thought insertion is taken to be a response to an abnormal experience rather than as a result of some brain abnor-

mality, the endorsement/explanation dispute ought to be understood as whether the delusion is an expression or endorsement of the experiential content or whether the delusion is an interpretation or explanation of the experiential content.

Within the literature on thought insertion, the predominant view appears to be an explanatory view. However, the issue has not been given much attention, and my suspicion is that many philosophers only consider the first proposal when thinking about the dispute. At first glance, it is difficult to see how the delusion "Chris has inserted 'Kill God' into my mind" is an endorsement of *feeling subjective* or *doesn't feel produced by me*. So, it seems reasonable to suspect that the delusion is an interpretation of some vague or poor content.

To borrow a term from David Gray, the problem that the endorsement view faces is that if the delusion was an endorsement of the experiential content, then the experiential content is *highly intricate* (Gray 2014). If one adopts an endorsement view, then it seems that the abnormal experience is incredibly complicated yet obvious enough that it can be endorsed. For instance, on Matthew Parrott's view, the experiential content would not only have to include things like *Chris is thinking "Kill God"* but also include things like *seems like my thought, but it isn't;* since some subjects' reports express ambivalence, the contradiction has to be part of the experiential content (Parrott 2017). On Gray's view, if the experiential content is *alien thoughts are being inserted into my mind*, then the problem for the endorsement view is that there are no *literal* instances of thought insertion, so how could one imagine an experience that is unimaginable? In other words, the subject lacks the proper frame of reference (Gray 2014).

Before responding to these criticisms, I want to present Gray's account of thought insertion, as I take it that it provides a useful way for thinking about the endorsement/explanatory dispute for thought insertion. For Gray, in order to lower the explanatory demand, we need to introduce a proto-delusion. The

alien thoughts and the non-alien thoughts are introspectively distinguishable. Since we typically take it for granted that our thoughts are our own, we generally do not have to attempt to identify who a thought belongs to. The fact that the schizophrenic has to perform an act of identification may generate a proto-delusion (in addition to some distress), such as "A thought I am introspectively experiencing is not mine." Since thoughts always have thinkers, and the schizophrenic believes the thought is not their own, a reasonable response would be to believe "a thought in my head is someone else's," and, so, the delusion is formed. Once the schizophrenic has developed the delusion "a thought in my head is someone else's" it is reasonable that they may attribute the thought to someone else and form a more sophisticated delusion, such as "This is Chris's thought." The notion of a proto-delusion is introduced since the thought "a thought in my head is someone else's" is still too sophisticated; we need a proto-delusion to lessen the demand on the experiential content. However, if the delusion could be formed solely on the basis of the abnormal experience, then, according to Gray, the proto-delusion just is the delusion (Gray 2014).

Both Gray and I believe that the patient may revise or further develop their delusion between the time in which it is formed and the time in which it is reported. I think the conservative approach to thought insertion is a minimal conception of the delusion—the patient has at least a minimal conception at the formation of the delusion. For example, in considering Parrott's argument, it is important to remember that not all the reports of thought insertion express a particular person, object, or agency as responsible for the alien thought. So, we should not assume that, for instance, *Chris* is part of the experiential content or part of the delusion at its formation—the subject may later, after forming the delusion, identify Chris as responsible for the alien thoughts. The same can be said for ambivalence. Whether a subject's report expresses some ambivalence does not suggest that the delusional belief

expresses some ambivalence at its formation. Contrast Parrott's conception with Gray's account; a subject may initially form the delusion "This thought belongs to someone else," but this delusion does not contain any ambivalence in it. Parrott's criticism may suffer from the problem that Henriksen and company point out, that some philosophers make the mistake of taking the report at face value or read the report too literally (Henriksen *et al*. 2019). The issue is whether the delusional belief is an endorsement of the experiential content, not whether the report about the belief is identical with the experiential content.

Gray's criticism may also put an unfair demand on the endorsement view. There are no real cases of thought insertion, so it obviously cannot be the case that the experience is the same as *literally* having thoughts inserted into one's mind. No one has ever experienced *literal* thought insertion. But why does an endorsement view have to hold the position that there must be instances of *literal* thought insertion in order for the delusion to be an endorsement? One can hallucinate a pink elephant and still believe they are seeing a pink elephant. Delusions, insofar as one takes them to be belief-like, are false beliefs about reality. If the delusion is an endorsement of the experience, the delusion does not have to express something true about the world.

The other issue is whether anyone can imagine what it is like to have such an experience, given they have no frame of reference for their imagination to draw from. For Sollberger, Gray's criticism may be an instance of experiential chauvinism—just because people without thought insertion cannot imagine such experiences, this does not mean these experiences are unimaginable to people with experience of thought insertion. Furthermore, it is compatible with the endorsement view that the experience has non-conceptual content and the delusion has conceptual content (Sollberger 2014). This is important since the patient may develop a way of *personally conceptualizing* the experience, and, since there is no *public conception* that is at odds with the personal

concept being formed, whether the patient lacks a frame of reference may not be problematic for a proponent of the endorsement view. The issue is not, for instance, whether the experience expresses some obvious concept, but whether the delusion is an endorsement of the experience.

This brings us to the first issue with trying to access whether we ought to adopt an endorsement view or an explanatory view: how are we conceptualizing the experience? For instance, if the experience is conceptualized as a lack of ownership, then it seems reasonable that the delusion "this thought is not mine" is an endorsement of an experience that lacks *mine-ness*. Similarly, it seems less reasonable that the delusion "This thought is not mine" is an endorsement of the experience of *volition*. To put this point differently (using Gray's framework), the issue is whether the phenomenal belief is a proto-delusion or the delusion. It might be the case that "This thought ("Kill God") feels *not-mine*" is a delusion, but "This thought ("Kill God") feels *volitional*" is a proto-delusion—where the delusion is an interpretation of the proto-belief.

However, I take the issue to be more complicated than this. A second issue can be elaborated on when we consider Sollberger's argument that the delusion and the experience do not have to match perfectly. For Sollberger, the experience may have non-conceptual content and the delusion conceptual content. We can ask whether the delusion and experience can differ in other respects as well. For instance, what if there is a proto-delusion and a delusion, but the proto-delusion and delusion only slightly differ from one another. Consider, for example, the case where the proto-delusion is "This thought ("Kill God") feels *not produced by me* (or *non-agential*)" and the delusion "This thought was produced by someone else." Let us assume, for the sake of argument, that the experience has some conceptual content of *not-produced-by-me*. As Gray points out, we have some reasons for suspecting that this sort of case supports an explanation view; the delusion is an

interpretation of the proto-delusion. However, it is not obvious to me that the delusion in this case is not taking the experiential content at face value. The formation of the delusion may involve background beliefs such as "thoughts have thinkers" and, given this background belief (in addition to the abnormal experience), the proto-delusion ends up being "This thought was produced by someone else." In other words, the proto-delusion stage is essentially skipped, and the delusion is an endorsement of the experience plus some background beliefs. The second issue deals with *what counts as endorsement*. Does the content of the delusion have to perfectly match the experiential content, and, if not, to what degree can the two differ and still count as an endorsement?

A third issue for the dispute—of whether thought insertion ought to be understood as an explanation or as an endorsement of the abnormal experience—is an issue about the conditions under which the experience happens and the delusion is formed. In discussing naïve introspection, Eric Schwitzgebel points out that there are some reasons to think that we are not as good at introspection as we think we are; or, that introspection may not always be the reliable source of knowledge that we assume it is. While many philosophers believe that we cannot be mistaken about experiences under canonical condition—such as seeing a red square in our direct line of sight, or focusing our attention on an intense pain—it is less clear whether or not we can be confused about experience under non-canonical conditions—we may be confused when it comes to deciphering whether a sensation in my arm is a mild pain or whether it is an itch, or seeing a red square in the peripheral field of vision (Schwitzgebel 2008). We can first ask whether the abnormal experience is closer to the example of the red square or whether it is closer to the example of the mild pain/itch, and if this should sway us one way or the other in terms of the endorsement/explanation dispute. For instance, the intuition may be that if the abnormal experience is similar to red, then this lends support to the endorsement view since the experience

appears to be taken at face value; whereas if the abnormal experience is closer to the mild pain/itch, then this lends support to the explanation view since the experience appears to be something we interpret.

However, I once again do not think this is obviously the case. For example, imagine it is possible that we can be confused about our experiences. Now consider the following: let us imagine that it is true that the abnormal experience is *not-agency*. Let us also imagine that the subject is confused and believes the experience is *not-mine* and forms the delusion "This thought ("Kill God") is *not-mine*." While the subject is confused about their experience and gets it wrong, they appear to be endorsing what they take the experience to be. The delusion is a false phenomenal belief.

A different articulation of this same issue is whether the delusion needs to form immediately, or whether the subject can have an alien thought and form the delusion later on. It is not clear to me which side of the dispute we ought to adopt. If we assume that the delusion does not form immediately in response to the presence of an alien thought, then the intuition might be that the delusion is not an endorsement of the abnormal experience. On the other hand, if the delusion's content matches the experience, regardless of how long it took the delusion to form, one might be inclined towards the intuition that delusion is an endorsement of the abnormal experience. It might be the case that we can take some time to deliberate on our experience, where that deliberation could still be construed as an endorsement of the experience.

IS THERE ANY WAY TO ADDRESS THESE ISSUES?

Without clarification on these issues, it makes it difficult to assess whether we ought to adopt an endorsement view of thought insertion or whether we ought to adopt an explanation view of thought insertion. However, I do not think it is impossible for use to decide

this matter, so we should not abandon all hope in light of these issues.

Rather than assuming what the characterization of the abnormal experience is, proponents of either view can construct arguments which follow one of two approaches. The first approach is to consider all the possible characterizations when constructing the argument. The argument has to feature the possibility that the abnormal experience is *not-agency*, *not-mine-ness*, or *volition*— and even then, it might be helpful to consider experiences outside of these three, since there are people like me who are skeptical that any of these experiences are the correct one. The second approach is to construct an argument that definitively rules out potentially problematic experiences. For instance, a proponent of the explanation view might rule out the experience of *not-mine-ness*, whereas a proponent of the endorsement view might rule out the experience of *volitional*. If the experiences cannot be definitively ruled out, then the argument should at least acknowledge which characterization it is assuming.

A further step that would help is for those involved in the dispute to decide what counts as endorsement. If the delusional content must be a perfect match with the experience, then this restricts the endorsement view. If, on the other hand, the delusional content does not need to perfectly match the experience and can differ from it greatly, then this expands the endorsement view.

Finally, we can settle issues on the conditions under which the delusion forms by acquiring more empirical evidence. We can, for example, identify people who are predisposed to schizophrenia and develop methods that would allow them to report on their experience either during the state of psychosis, where they first experience alien thoughts, or methods where they report directly after the state of psychosis.

CONCLUSION

The aim of this paper was to clarify what I took to be the right way of thinking about the delusion of thought insertion and to highlight the difficulties with determining whether we ought to adopt an endorsement view or an explanation view of thought insertion. I suggested that we ought to think of the delusion as a phenomenal belief, and I expressed three issues that face those involved in the endorsement/explanation dispute: first, the issue over how to characterize the abnormal experience; second, the issue over what counts as endorsement; and third, the issue over the conditions in which the delusion is formed.

In this paper, I suggested possible ways of dealing with each of these issues. I also pointed out that while the predominant view is the explanation view, the dispute has been somewhat ignored and it is not obviously the case that the explanation view is correct. While proponents of the explanation view offer various objections to the possibility of *highly intricate* experiences, if the abnormal experience is an impaired *sense of ownership*, then the delusion could be understood as an endorsement of the experience. If we agree that the abnormal experience is the locus of the delusion, then—assuming that delusions are belief-like—whether the delusion is an endorsement of the abnormal experience suggests that the delusion should be understood as involving a phenomenal belief—either the phenomenal belief is a proto-delusion or it just is the delusion. This is, in part, because it would be a mistake to assume that the alien thought is the delusion in question.

In closing, thought insertion is a philosophically interesting delusion. It presents challenges both to assumptions held in areas of philosophy, such as philosophy of mind and epistemology, but also presents challenges for those interested in delusions. In other words, thought insertion is still something that seems philosophically strange.

Notes

1. Not everyone writing on thought insertion agrees that the delusional element is the response to the alien thought. For instance, Gold and Hohwy argue that the abnormal experience itself is irrational, and that this is the delusional element (Gold and Hohwy 2000)

2. Within the literature that focuses on delusions in general, not everyone holds that delusions are doxastic (that the delusion is a belief or, at the very least, belief-like). For example, on Currie's view, the delusion is understood as an imagining (Currie 2000; Currie and Ravenscroft 2002)

3. This view is sometimes referred to as the two-concept view. Among those who endorse the view, there is disagreement over what the two concepts should be called. While some authors will refer to the second concept as a *sense of ownership*, I have chosen to call it a *sense of subjectivity* in order to avoid confusing the reader. On the two-concept view, a *sense of ownership* (or subjectivity) is not quite the same thing as a *sense of ownership* articulated by the second proposal.

4. It is not always clear that proponents of a *sense of ownership* or *mineness* are always in agreement over how to characterize this experience. For instance, on Henriksen and company's view, a lack of for-me-ness or mine-ness appears to be an intense form of detachment that occurs during psychosis; whereas on Bortolotti's view, it is not clear whether a sense of ownership is the recognition that some of our thoughts feel unfamiliar or whether it should be understood as the abnormal experience.

5. As Coltheart and company point out, both patients with Capgras delusion and patients with damage to the ventromedial region of the brain have a lower autonomic response to the perception of familiar faces. For Coltheart and company, the delusion "this is not my wife" is generated as an inference to the best explanation by the brain given that the brain takes this to be the best explanation on the basis that it has the highest probability (within a Bayesian framework) (Coltheart et al. 2010).

Bibliography

Bortolotti, Lisa. (2010) *Delusions and other irrational beliefs* (Oxford University Press)

Bortolotti, Lisa, and Broome, Matthew. (2009) "A Role for Ownership and Authorship in the Analysis of Thought Insertion," *Phenomenology and the Cognitive Sciences* 8(2), pp. 205-224

Bortolotti, Lisa, and Miyazono, Kengo. (2015) "Are Alien Thoughts Beliefs?," *Teorema: Revista Internacional de Filosofía*, pp. 135-148

Campbell, John. (2002) "The Ownership of Thoughts," *Philosophy, Psychiatry, & Psychology* 9(1), pp. 35-39

Coltheart, Max. (2005) "Conscious Experience and Delusional Belief." *Philosophy, Psychiatry, & Psychology* 12(2), pp. 153-157

Coltheart, Max, Menzies, Peter, and Sutton, John. (2010) "Abductive Inference and Delusional Belief," *Cognitive neuropsychiatry* 15(1-3), pp. 261-287

Currie, Gregory. (2000) "Imagination, Delusion and Hallucinations," *Mind & language* 15(1), pp. 168-183

Currie, Gregory, and Ravenscroft, Ian. (2002) *Recreative Minds: Imagination in Philosophy and Psychology* (Oxford University Press)

Gold, Ian, and Hohwy, Jakob. (2000) "Rationality and Schizophrenic Delusion," M*ind & Language* 15(1), pp. 146-167

Graham, George. (2004) "Self-Ascription: Thought Insertion," in: Radden, Jennifer, 2006 ed. *The Philosophy of Psychiatry: A Companion* (Oxford University Press)

Gray, David Miguel. (2014) "Failing to Self-Ascribe Thought and Motion: Towards a Three-Factor Account of Passivity Symptoms in Schizophrenia," *Schizophrenia Research* 152(1), pp. 28-32

Henriksen, Gram, Mads, Parnas, Josef, and Zahavi, Dan. (2019) "Thought Insertion and Disturbed For-Me-Ness (Minimal Selfhood) in Schizophrenia," *Consciousness and Cognition* 74, p. 102770

Martin, Jean-Remy, and Pacherie, Elisabeth. (2013) "Out of Nowhere: Thought Insertion, Ownership and Context-Integration," *Consciousness and Cognition* 22(1), pp. 111-122

Mullins, Simon, and Spence, Sean. (2003) "Re-examining Thought Insertion: Semi-Structured Literature Review and Conceptual Analysis," *The British Journal of Psychiatry* 182(4), pp. 293-298

Parrott, Matthew. (2017) "Subjective Misidentification and Thought Insertion." *Mind & Language* 32(1), pp. 39-64

Rössler, Wulf, et al.. (2007) "Psychotic Experiences in the General Population: a Twenty-Year Prospective Community Study," *Schizophrenia Research* 92(1-3), pp. 1-14

Schwitzgebel, Eric. (2008) "The Unreliability of Naive Introspection," *Philosophical Review* 117(2), pp. 245-273

Sollberger, Michael. (2014) "Making Sense of an Endorsement Model of Thought-Insertion," *Mind & Language* 29(5), pp. 590-612

Stephens, G. Lynn, and Graham, George. (2000) *When Self-Consciousness Breaks: Alien Voices and Inserted Thoughts* (The MIT Press)

SUPERNATURAL CLAIMS AND THE THEORY OF RECOLLECTION AS PRESENTED IN THE *MENO*

John Hurley

INTRODUCTION

Plato's theory of recollection (*anamnesis*), as expressed in the *Meno*, presents interpretive challenges that have been discussed since Plato's own time. Current views of the theory are discussed in a recent book by Gail Fine (2014) and in an article by David Bronstein and Whitney Schwab (2019). These include claims that Plato thinks that: human beings are born with innate but hard-to-excavate states of knowledge (*cognitive condition innatism* or just *condition innatism*); or that the human soul, which preexists the biological person, contains knowledge that is lost at birth (*prenatalism*); or that we have "innate cognitive contents" that may become the content of occurrent states of knowledge or opinion (*content innatism*).[1]

The theory is formulated in the *Meno* in a way that includes a sequence of three related supernatural claims. First, the souls of the dead exist for a period in Hades; second, these souls return to the earth in new bodies; third, our souls have existed for an infinite time before our births. I will consider the coherence and plausibility of the interpretations listed in the first paragraph if the support of these supernatural claims is denied. In particular, I will argue that, if we substitute some naturalistic (i.e. not supernatural) mechanism for the preexistence of the soul as a way to explain the origin of the knowledge we recollect, we are making a clear break with Plato's theory. In this case, though, our position is irreconcilable with prenatalism but consistent with condition innatism, and

especially with content innatism. I will also discuss some implications of these readings for the scope of the theory in the absence of preexistence.

THE THEORY OF RECOLLECTION
IN THE *MENO*

The Meno contains what may be the earliest statement of the theory of recollection (*anamnesis*), which I shall refer to as "the theory" or TR. TR is offered as a solution to "Meno's Paradox," the argument that:

> Premise 1: For any instance of possible knowledge, one either knows or does not know.

> Premise 2: One cannot inquire into what one already knows.

> Premise 3: One cannot inquire into what one does not know.

> Conclusion: One cannot inquire into anything.

Premise 3 is itself supported by the premise that, if one inquires into what one does not know, one will not be able to recognize the truth if it is found. Many possible attacks on the paradox claim that, in modern terms, it is invalid due to equivocation in the uses of "know" in the premises. Others are based on the observation that, even without equivocation, some of the premises are false for some interpretations of "know"[2] (Fine 2014, chapter 3). Socrates, in fact, elaborates TR as an attempt to show that Premise 3 is false and that the argument above is, in the terms a modern logician would use, unsound.

According to TR as it appears in the *Meno*, the human soul (*psuche*) is the entity in which knowledge resides, and the *psuche* preexists the biological individual. Further, knowledge exists within the soul before birth, which is not clearly distinguished from conception. Various readings of the *Meno* view TR as claiming that, at birth, knowledge either is lost or takes on some attenu-

ated form. Learning and inquiry consist of either discovering this knowledge within ourselves or simply learning it again. TR does not only claim, like many later philosophical views, that certain fundamental dimensions of understanding are *a priori*, but that such specific instances of knowledge as how to make geometric calculations have been present in the soul since before birth.

Socrates gives flesh to TR by proposing that the *psuche* experiences an infinite sequence of reincarnations, with some time elapsing between a soul's release from one body and its implantation in a new one. During this time gap, the *psuche* is exposed to true knowledge. According to most, but not all, readings of the theory, it maintains this knowledge, or a derivative of it, in some occult form in its future lives.

Several aspects of the *Meno* present obstacles to acceptance of TR. First, Socrates proposes an account of metempsychosis (or, to use a term more common in modern English, reincarnation)[3] that hinges on supernatural beliefs most modern readers are likely to reject. The account begins with a distractingly specific myth, of a type of which Socrates himself is skeptical in dialogues including the *Euthyphro*. Second, Socrates demonstrates recollection with a dialectical exchange concerning geometry whose persuasiveness is controversial. Finally, the *Meno*, whose headline question is whether virtue (*arête*) can be taught, leaves unresolved the issue of whether virtue is a form of knowledge, so it is unclear whether TR comes into play at all in the larger dialectical context.

On the other hand, these obstacles may not be relevant. A reader who rejects the myth could consistently accept the rest of the account, and indeed Socrates explicitly calls attention to the speculative character of the myth. Furthermore, questionable methodology in a thought experiment does not falsify a theory. Finally, the validity and soundness of the theory do not depend on whether it is really applicable to the central question of the Meno.

Noting Plato's characteristic subtlety and the multilayered structure of the dialogue genre, I think these obstacles may be a

deliberate effort to call attention to the provisional character of dialectical positions. If this is correct, Socrates is not setting down a dogma or asking us to believe a myth; he is testing a provisional account and inviting future philosophers to try to find something better. The dialectic of earlier dialogues is full of the interlocutors' proposed ideas, which Socrates demolishes; perhaps here Plato is inviting us to apply his method to his own thoughts or those of the character Socrates. Furthermore, the framing narrative of the *Meno* includes oblique references to Socrates' execution about three years after the dramatic date, giving its hero an extraintellectual motivation for believing in immortality. Plato uses similar devices in other narratives to explain the positions of interlocutors whose ideas are shown to be inconsistent or unfounded; it is at least plausible that he undermines Socrates himself here in this way. Whatever the status of the myth, the account of metempsychosis, and the geometry lesson, though, Plato was certainly committed to TR. Versions of the theory appear again in dialogues that appear to have been written later than the *Meno*, including the *Phaedo* and the *Phaedrus*.

My discussion of TR under various interpretations will circle around a certain black hole. It is tempting to assume that the knowledge covered by TR consists of the Platonic forms, the abstract ideas that Plato, or at least the character Socrates takes to be the foundation of reality. TR might provide a straightforward answer for a key question raised by the theory of forms, namely whether and how it is possible for us to attain knowledge of the forms. On this account, the knowledge to which souls are exposed between lives would consist of the forms. Our project in seeking knowledge during life would consist of seeking knowledge of the forms, an idea appealingly consistent with views expressed by Socrates in the *Republic*. We would do this by trying to either reach the memory of this knowledge within our souls or simply to learn it again. Perhaps we would draw on our experience of the approximations or, in Plato's most famous metaphor, projections,

of the forms which we perceive.

This satisfying account of the linkage between TR and the theory of forms, though, is not offered in any obvious way in the *Meno*. It is not clear whether the *Meno's* references to the essence (*eidos*) of virtue refers to a form in the sense that appears in the fully developed theory. The *Meno* may have been written before the theory of forms was fully developed, although Gail Fine sees clear ideas of the forms in the dialogue (Fine 2014, pp. 42-45). At minimum, though, some kind of objects must exist in a plane other than the world of our experience for the theory of recollection to make sense.

TR is an opening contribution to the long-running interplay between, on one hand, rationalism and innatism, and on the other, empiricism. Accordingly, it has received serious and sustained attention from philosophers from Plato's own time to today. One recent discussion of the theory pits an interpretation by Fine (2014) against one by David Bronstein and Whitney Schwab (2019). Both sides agree that the most common modern reading of the theory as found in the *Meno* involves *condition innatism* (or *cognitive condition innatism*): "human beings are born with latent innate states of knowledge" (Bronstein and Schwab 2019, p. 392). Both reject this view. Fine argues instead for the position she calls *prenatalism*: recollection works by learning again knowledge the soul had prenatally but lost at birth. Bronstein and Schwab, in contrast, find in the *Meno's* account of the theory *content innatism:* "human beings are born with innate cognitive contents that can be, but do not exist innately in the soul as, the contents of states of knowledge."

This paper examines the several ways a reader could reject supernatural claims in the *Meno* and their implications for each of the three readings of TR listed above. The most important question involves the tenability of various interpretations of TR for someone who rejects any reasonably literal reading of the preexistence of the soul. Rejections of other parts of the supernatural are

also considered.

THE MYTH, METEMPSYCHOSIS, AND PREEXISTENCE OF THE SOUL

Socrates presents his account of the mythological basis of TR at 81a-d:

> They [the priests, priestesses, and poets from whom Socrates has learned this myth] say that the soul of man is immortal. At one time it comes to an end—which people call dying—and at another it is reborn, but it is never destroyed and so, one must live life as piously as possible, since to those from whom:

> Persephone accepts compensation for her ancient grief,
> To the sunlight above, in the ninth year, their
> Souls she again returns.
> From them come noble kings,
> Increasing in strength and wisdom, men both swift and great,
> And for the rest of time called hallowed heroes by men.

> When the soul is immortal and has been born many times, and has seen the things here and the things in Hades—everything, in fact—there is nothing that it has not learned, and so it's no surprise that it can recollect about virtue and other things that it knew before. And because all of nature is of the same kind and the soul has learned everything, nothing prevents a soul that has recollected just one thing—which people call learning—from rediscovering everything for itself, so long as it is brave and doesn't give up.[4]

In the discussion below, I will use "the myth" to refer to the story involving Persephone, inclusive of its account of metempsychosis. However, metempsychosis might be achieved in some

other way, so I will also discuss this idea on its own. Finally, I will discuss the idea of the preexistence of the soul (which I shall call PS) independently of the rest of the account.

The *Meno* is not unusual in ancient philosophy in its intertwining of argumentation with supernatural claims, or even in its inclusion of a very specific mythological explanation. There are also analogues of this pattern in the relationship of many works of medieval and modern philosophy with Christian and other religious beliefs. Disentangling these two strands in the *Meno* is broadly similar to the procedure a naturalist-minded reader (i.e., one who rejects explanations that appeal to entities outside nature) applies in reading a wide range of other philosophical works critically.

Few modern readers of Plato believe that Persephone reigns over Hades. However, a substantial subset believes the human soul persists after death, but would reject metempsychosis and PS. I will seek here to show how such positions might, alongside a purely naturalistic position, fit in a spectrum of partial or full rejection of the supernatural account of TR in the *Meno*.

I want to make a clear distinction between reasonably literal interpretations of religious or mythological accounts and highly metaphorical ones. Symbolic interpretations quickly dissolve the integrity of supernatural beliefs. A reading of a supernatural claim should certainly be regarded as incoherent if it is actually inconsistent with a plain reading of the account it claims to interpret (for example, "karma doesn't require reincarnation; it just means that you get what you deserve in this life.") However, it is often possible to substitute some naturalistic reading that elides or ignores part of the supernatural idea, avoiding the potential falsity of the original claim (for example, "karma is just a claim that you should behave well for your own good"). Humanities-based education may prime us for this kind of purely symbolic understanding, but it is not a sound procedure when the objective is to understand supernatural beliefs as they actually are. The habit

of denaturing such claims makes it easy to misrepresent them in ways that empty them of meaning that has not simply been projected by the interpreter. That renders arguments built upon such interpretations uninformative and/or tautological. Perhaps less obviously, applying highly metaphorical readings to supernatural beliefs strips them of their original context, in which one claim might depend on others. For example, belief in salvation through faith depends on belief in the divinity of Jesus and in resurrection. This structure is demolished if a highly-metaphorical reading of salvation through faith is adopted. The result leaves the reader unable to understand the relationship of salvation through faith to the broader structure of Christianity.

For the purposes of this essay, I rule an interpretation of a supernatural claim that changes an incidental part of the account (for example, that souls wait in future mothers' ovaries rather than in Hades, or that souls are combined or divided during metempsychosis), an admissible loose reading. On the other hand, a reading that might have a different truth value from some necessary feature of the original claim breaks with the claim. I rule this a rejection of the theory and a proposal of a new one that does not depend on the suspect claim. The adaptation of parts of an old theory into a new, different one is a perfectly respectable procedure, and later in this paper I shall discuss some ways to create naturalistic theories that integrate aspects of TR. Such theories, though, are *descendants* of TR, not *interpretations* of it.

Accordingly, in reference to supernatural beliefs, I will use the term "believe" and its synonyms to mean "believe that a literal reading of the essential aspects of the claim is true," not "be able to construct some symbolic reading to which one could assent."I will use "disbelieve" and similar terms to mean "believe that the claims, taken in any reasonably literal sense, are false," rather than "be unable to find any reading that allows one to assent to the account."

If we reject plain readings of mythological or other super-

natural accounts, we must ask how theories like TR fare without them. Of course, at most points on a spectrum from religious literalist to firm naturalist, we are not equally disposed to reject every supernatural belief. The more detailed the claim, the more readily a skeptic rejects it. Where there can be no direct evidence, it appears to be a sound practice to doubt the specific more than the general. My disbelief that a vengeful god has commanded me to stone people for homosexuality is more firm than my disbelief that any kind of entity exists outside nature. This is not a loosening of naturalism, but an acknowledgement that it is more likely to be wrong with regard to general claims than detailed ones.

The text of the *Meno* suggests that Plato himself anticipated that his readers were more likely to doubt the myth than the rest of the theory, and more likely to doubt metempsychosis than the preexistence of the soul. As I will argue below, Socrates is tentative about the myth and, to a lesser degree, about metempsychosis, but he means what he says about PS, and he does not expect us to accept TR without it.

It should be noted that the rendering of *psuche* as *soul*, a word that suggests religious beliefs concerning immortality more strongly than psuche did, may cloud the waters. As noted above, Socrates believes that the *psuche* survives bodily death, but he argues this position in the *Meno*. Plato does not beg the question with his choice of terms.

FOUR WAYS TO REJECT SUPERNATURAL CLAIMS IN THE *MENO*

There are three points at which we could break with TR's supernatural claims; we can reject the myth, and/or reject metempsychosis, and/or reject preexistence. If these points of rejection were independent of each other, there would be eight possible forms of rejection. However, in two cases, the points of possible rejection are not independent. Since the myth includes an account of rein-

carnation, positions that accept the myth but reject reincarnation are inconsistent. Further, if metempsychosis is a correct account of the origin of particular humans' souls, PS is true. Thus, positions that accept metempsychosis but reject PS are also internally inconsistent. After rejecting the inconsistent positions, we are left with four ways to consider supernatural claims in the *Meno*. The first is the believer's stance:

AAA) Accept the myth, accept metempsychosis, accept preexistence.

The remaining three positions are the ones that take partly or fully naturalistic stances:

RAA) Reject the myth, accept metempsychosis, accept preexistence.

RRA) Reject the myth, reject metempsychosis, accept preexistence.

RRR) Reject the myth, reject metempsychosis, reject preexistence.

These four positions are all coherent, and they have varying repercussions for TR.

AAA, the believer's position, is obviously unthreatening to the theory. For my purposes, the remaining three positions are more interesting.

If specific assertions are easier to reject than vague ones, the myth is the easiest target among the three supernatural claims. Further, another way to understand the dependencies among the claims is to say that rejecting PS entails rejecting metempsychosis, and rejecting metempsychosis entails rejecting the myth, so that anyone who rejects any of the claims must reject the myth. This is reflected in the fact that all three of the coherent and partly or fully naturalistic stances above contain "R" in the first position; they all reject the myth.

RAA, which rejects the myth but gets back on track with the theory by accepting metempsychosis, is easily tenable. The world's religions offer many consistent accounts of reincarnation, and we could also certainly devise one of our own. RAA, though, offers only a fleeting challenge to TR. As we shall see below, Plato anticipated rejection of the myth and correctly found it unthreatening.

RRA, which rejects the myth and metempsychosis but accepts preexistence of the soul, requires either an account of PS that does not involve metempsychosis or an acceptance of PS without an explanation. RRA seems also to have been anticipated by Plato, but Socrates does not try to save TR in the face of RRA. Instead he says at 86b that, for instrumental reasons, we should act as if TR is true even if we do not accept his argument for immortality. However, RRA, with its acceptance of the preexistence of the soul, does not threaten the essential claim of TR, the persistence of knowledge or cognitive content in the *psuche* across multiple lifetimes.

RRR, the purely naturalist position, then, is the only one of these positions that entails rejection of TR. Preexisting knowledge obviously requires some temporal point at which the *psuche* existed before birth, and the *Meno* does not clearly distinguish birth from conception.

A reader might adopt a purely symbolic reading of PS. In particular, a modern person might view prenatal knowledge as a genetic inheritance and construe the *psuche* in such a loose way that it would embrace a history of evolution. Such a view may be useful as a theory about the world, but, as I argue above, it sheds little light on TR. For this discussion, I schematize any reading that does not require actual survival of the soul after death and actual reincarnation as a rejection of PS, not a loose reading of it. Thus, for my purpose, someone who substitutes some naturalistic explanation for PS is a tenant of RRR or an attempted tenant of one of the inconsistent positions which reject PS but accept one or

both of the other claims. Similarly, a reader might accept metempsychosis only as a symbolic substitute for some other way to accomplish preexistence that is impervious to explanation, incomprehensible, or just unknown. According to my method, such a reader rejects metempsychosis, making him or her a tenant of RRA or RRR.

Socrates anticipates that Meno, and by extension, readers of the dialogue, may doubt his claims, especially the myth. At 81a-b, Socrates says the myth seems to him "true and beautiful," before citing it as an expert opinion:

> Those [elaborating the myth] are some of the priests
> and priestesses whose business it is to be able to give an
> account of their practices. And Pindar says it too and many
> others of the poets, those who are divinely inspired.

Yet he immediately seems to distance himself from full commitment to the myth: "And what they say is the following; see if you think they speak the truth."

At 85c-86b, Socrates argues that the human soul is immortal. The myth states that souls are "never destroyed" and offers only one source of souls for embodiment, the stock of *psuchai* waiting in Hades. Thus, Socrates could simply say that if the myth he has just recounted is true, the *psuche* is immortal. He instead argues for immortality in a way that is intertwined with his presentation of the rest of TR. In the geometry lesson, Socrates seeks to show that an untutored person has a latent understanding of that discipline. He singles out a slave of Meno's and puts to him the problem of finding the length of the side of a square that will be twice the area of a given square. After verifying that the slave speaks Greek, Socrates uses a series of questions to get the slave to first assent to incorrect suggestions and produce incorrect answers, then assent to correct suggestions and make correct answers. The crux of his argument for immortality begins at 85e. Although no one has taught the slave about geometry at any time during his life,

Socrates has, he claims, been able to excavate true beliefs about geometry from the slave's soul.

> So: If he didn't acquire [these true beliefs] in this life, then, the following becomes immediately obvious, that he had and learned them at some other time.

> M: Apparently.

> So: And this is the time when he was not a human?

> M: Yes.

> So: If there are true beliefs within him during the time when he both is and is not a human being,[5] which each become knowledge when stirred up by questioning, then won't his soul be in a state of having learned throughout all time? Since it's clear that throughout all time he was or was not a human?[6]

> M: Apparently.

> So: Then, if we always possess the truth about things in the soul, the soul would be immortal, so that the person who doesn't happen to have knowledge now, that is, who hasn't recollected, should be brave and try to search for it and recollect?

We usually understand immortality to extend *forward* from the current point in time, involving existence after our death. If Socrates' argument here is sound and does not depend on a covert appeal to the myth, it directly shows only that the soul has existed before the life of its current host, and even then not necessarily for an infinite time. However, if our souls have existed in previous incarnations, at least some souls must survive bodily death. If some souls are lost at death, but all that are embodied have lived before, the count of human beings must decline, either episodically or steadily. We might argue that if Socrates thought

51

this was so, he would remark on this implication, since it would mean that if TR is true, the stock of knowledge is decreasing. However, this possibility seems to remain open.

In any case, Socrates has left room for a version of TR without the myth by making an argument that is only relevant if the myth is false, since it is obviously redundant if the myth is true. This is certainly consistent with the possibility that Socrates, who often speaks with literary flourish, recites the myth simply to call attention to the need for some mechanism of metempsychosis.

> At 86 b-c, Socrates states an instrumental, rather than logical, reason to accept PS:

> I wouldn't strongly insist on the other aspects of the argument, but that we would become better men and braver and less lazy if we believe it is necessary to search for what on doesn't know, rather than if we think that we can't discover what we don't know and should not look for it, for this I will fight strongly, if I am able, in both word and deed.

It might be claimed that this suggests that he thinks that even some version of TR without PS is tenable. As noted above, though, Socrates has presented TR at 85e-86b in a way that is inextricable from PS, making it integral to the theory. The diffidence about the theory at 86b-c suggests that Socrates thinks that a philosopher who is unconvinced of TR should act in the same way as s/he would if s/he believed it. Such a person should reject Meno's paradox and seek knowledge by dialectic or other methods. This does not imply that Socrates thinks the theory is tenable without PS.

THE SCOPE OF TR

The scope of TR as presented in the *Meno* is unclear. If souls have infinite time to acquire knowledge, we can't infer anything about the scope of such knowledge. They might learn about all contin-

gent facts that have ever been true, and, in a deterministic world, be able to predict all that will ever be true. Socrates himself is not explicit about the scope he claims for TR. At 81d, he speaks of a scope that allows for exceptions, stating that "searching and learning are, as a whole, recollection." At 81e, though, he speaks as if TR covers all knowledge, denying the paradox's claim that: "a person cannot search either for what he knows or what he doesn't know."

In the geometry lesson, Socrates' questions to the slave are strongly leading, and the slave readily calls up basic mathematical skills that one might doubt would be in the repertoire of a typical Greek slave. The point I want to make here, though, is that the knowledge under discussion is necessary, not contingent. Socrates needs to confirm that the slave speaks Greek; he does not try to draw out an innate knowledge of, for example, the Greek word for square. Socrates supplies the dimensions of the square he uses to begin the demonstration, which he presumably draws in the earth; he does not try to draw out occult knowledge of this *particular* square from the slave's *psuche*.

The question ostensibly at issue in the *Meno* is whether virtue can be taught. Subsequent to the discussion of TR, Socrates suggests to Meno that he examine this question by considering the record of virtuous Athenians and of the sophists, in their known or presumed attempts to teach virtue to, respectively, their sons and their students. He urges Meno to draw conclusions from empirical observation here; he does not suggest that Meno search his soul for such specifics. Socrates is unsure of whether virtue is a form of knowledge (89d), and such questions as whether Themistokles was a successful teacher of virtue may be insufficiently specified to have answers. However, either it is a fact that Meno has "heard anyone, young or old" say that "Kleophantos, son of Themistokles, became a good and wise man" (93a-94a) or it is a fact that he has not. Similarly, someone who knows how to get from Athens to Larissa (97a) has knowledge of a set of contin-

gent facts. At 97b, Socrates contrasts a person with knowledge of the road to Larissa with someone with mere true belief about the road. This example must be fresh in his mind at 97d-98b, where he compares acquiring knowledge through recollection to tying down Daedalus' statues, which otherwise have a tendency to fly away. He thinks that we acquire knowledge through recollection, and he regards contingent facts as knowledge, but he does not state clearly that such facts are within the scope of TR.

THREE INTERPRETATIONS OF TR

In this section, I will discuss each of the three modern interpretations of TR listed in the introduction. My main concern is how these readings fare in the face of RRR, the purely naturalistic view that rejects the preexistence of the soul along with the other supernatural claims. I avoid making exegetical claims about support for the innatist and prenatalist views in the *Meno*, since 1) I am not qualified to examine the texts in the original Attic Greek, and 2) such arguments would be undermined by the fact that RRR makes an explicit break with the theory anyway. Prenatalism and innatism are attempts to reconstruct Plato's views about knowledge, but they are often expressed as if they were theories about knowledge itself. In this section, I will accept that conceit and consider how they stand in the face of RRR as theories about the world rather than about the correct interpretation of the *Meno*.

Bronstein and Schwab describe a conventional interpretation of TR in which humans are born with knowledge that is simply hard to access, as a form of *condition innatism*:

> Human beings are born in the cognitive condition of knowing, in the sense that human embodied souls [sic] possess latent innate states of knowledge. The goal of recollection is to make one's latent knowledge explicit."
> (Bronstein and Schwab 2019, p. 393)

Bronstein and Schwab themselves take up the cause of content innatism:

> Human beings are born with innate cognitive contents, in the sense that human embodied souls innately possess truths that can be the contents of states of knowledge but do not exist innately in the soul as the contents of such states. The goal of recollection is to take up these contents in such a way that one knows them (again). (Bronstein and Schwab 2019, p. 394)

The two types of innatism differ in the type of innate entity that they claim to lie short of conscious knowledge and in the nature of the process by which we can convert such content to conscious knowledge. Condition innatism claims we have states of knowledge that may be latent and need to be uncovered. Content innatism claims that some truths lie deeper than that; they are innate in our *psuchai* and can become the content of states of knowledge that may in turn become usable. These truths are as yet *uncognized*; they do not require the current existence of states of knowledge (Bronstein and Schwab 2019, p. 394).

Bronstein and Schwab develop their view because they find it the best way to explain certain aspects of the text of the *Meno*. I will advocate it for a different reason. Under naturalist assumptions, a state of knowledge, as well as the existence of a *psuche*, must require a hosting brain. Cognitive contents that fall short of knowledge, on the other hand, might be derived from some other source in human biology, possibly one that exists before the formation of the brain and/or before the inception of the *psuche*.

In her comprehensive discussion of Meno's paradox and TR, Gail Fine offers a reading of TR that Bronstein and Schwab call *prenatalism*. In her view, TR does not rely on the presence of innate knowledge. Rather, souls lose their knowledge, or cognitive contents, at birth. Fine does not dub her interpretation with a name, but she does use the term *prenatal* to identify the knowl-

edge at issue (Fine 2014, Ch. 3). *Prenatal* is presumably chosen for the parallel to *innate*, but it may require clarification. Fine does not imply that her view opens up the possibility of knowledge present between conception and birth, which could perhaps have served as a substitute for metempsychosis.

Prenatalism does not offer a clear explanation of why TR views learning as recollection rather than simply learning again. If knowledge is simply lost at birth, there is no object to recall. As Fine puts it, there is no "entitative feature" to account for the persistence of knowledge and thereby support recollection. Fine thinks Plato simply did not address this key point clearly. According to Fine, this problem is located in Plato's text, not in her theory:

> either Plato rejects the entitative feature and so can't explain why we are recollecting rather than learning for a second time or he accepts it but leaves a serious lacuna in his argument by not explaining what it is. (Fine 2014, p. 165)

INNATISM, PRENATALISM, AND NATURALISM

TR, with its dependence on the preexistence of the soul, is not tenable under RRR. It is, though, possible to substitute some naturalistic account of innate knowledge for PS. As I have argued above, if supernatural beliefs are to be taken seriously, this does not create a variant of TR, but a new theory that is a descendant of TR. Like TR, such a theory would claim that learning consists of excavating material that is within our minds, not discovering information through observation or instruction. The questions that innatism and prenatalism attempt to answer about the mechanics of preexisting knowledge or content and how to access it remain relevant.

In the account given in the *Meno*, the entity that possesses knowledge is the *psuche*. Prenatalism is irreconcilable with any

view that rejects preexistence of the *psuche*, since there is no temporal position in which prenatal knowledge could have been held before being subsequently forgotten. Thus, prenatalism offers nothing to explain with a naturalistic account. The innatist views offer more promising terrain for this discussion.

The most likely naturalistic theory of recollection replaces PS with a high estimate of the quality and quantity of knowledge that is part of a genetic endowment. Indeed, as Fine notes, Noam Chomsky and others have commented on parallels to Chomsky's position that important linguistic structures are innate in humans (Fine 2014, pp. 165-168). Such a theory is consistent with innatism in general. However, it is easier to find a plausible naturalistic explanation for innate information than for innate states of knowledge. Content innatism provides more, simpler, and more plausible ways to support a naturalistic theory of recollection than condition innatism does.

If we reject PS, innatism requires us to explain how a new *psuche* can, at the moment of its inception, either hold states of knowledge or possess cognitive content that does not participate in such states. The identity of that moment of inception, however, may be specified in different ways. The myth speaks of souls' being "reborn," and seems to imply that the soul is implanted in a human at the moment of birth, but, having rejected the myth, we are not bound by its details. As far as I know, neither Socrates nor either Fine or Bronstein and Schwab allow for the possibility that *psuchai* come into being at conception or some other time prior to birth, but such a position should be accounted for here.

Since states of knowledge exist only in the *psuche*, they can't predate its creation. Innatism claims that the states are present in the *psuche* at all times during its existence, so they can't postdate it either. Thus, a version of condition innatism that does not rely on PS would require us to accept at least one of three claims:

A1) The *psuche* comes into being, with occurrent states of

knowledge, at the instant of conception.

A2) The *psuche* comes into being, with occurrent states of knowledge, during gestation.

A3) The *psuche* comes into being, with occurrent states of knowledge, at the instant of birth.

Under naturalism, though, a state of knowledge, and indeed a *psuche*, require the availability of a brain. Claim A1 is therefore not tenable for naturalists. To be consistent with content innatism, then, a naturalist descendant of TR would have to accept either A2, A3, or both; the *psuche*, with its states of knowledge, must come into being during the development of the fetus and/or flash on at birth.

Content innatism, which claims that we have innate cognitive contents rather than states of knowledge, similarly involves one or more of the following claims:

B1) The *psuche* has access to cognitive contents acquired at the instant of conception.

B2) The *psuche* has access to cognitive contents acquired during gestation.

B3) The *psuche* has access to cognitive contents acquired at the instant of birth.

Content innatism may be easier than condition innatism to reconcile with naturalism. First, unlike states of knowledge, cognitive contents that may be encoded in genetic material might predate the existence of the brain and/or *psuche*. Thus, claim B1 is available to naturalists, unlike A1. Second, B2 allows us to adopt the highly plausible view that innate cognitive contents develop with the brain, with the psuche forming at some late date during gestation or at birth. B3 is also consistent, if not as apparently plausible as B2.

Under either of the innatist accounts discussed above, the scope of the knowledge covered by TR would be limited. We might be born with knowledge or content concerned with necessary facts, such as the distinction between left and right, of general cognitive skills such as how to distinguish entities from events, or of evolutionarily-developed linguistic structures. On the other hand, while understanding of contingent facts might require innate mental faculties, it seems unlikely that a naturalistic account could support the idea that our minds have innate contents encompassing such contingent facts as whether anyone in Athens claimed that Kleophantos was a virtuous man or how to get from that city to Larissa. Such a limited scope for TR is consistent with the illustration of the theory in the geometry discussion, but it seems contrary to Socrates' broad statements about the role of recollection, as well as to the examples in the discussion about teaching virtue.

CONCLUSIONS

The account of the Theory of Recollection in the Meno makes three related supernatural claims: 1) disembodied *psuchai* wait in Hades for reincarnation; 2) experienced souls are embodied in new biological persons; 3) we inherit *psuchai* that have had knowledge in the past and, according to most readings, still have either knowledge or necessary components of it.

TR has supported several accounts of the epistemological process by which innate or prenatal knowledge is said to be recovered during life. I have not attempted here to adjudicate which of these readings is best supported as an account of the positions taken by Plato or by the character Socrates. I have investigated, though, some implications these readings have for TR, as a theory about the world, if we reject the supernatural claims involved in its exposition and justification in the *Meno*.

The theory of recollection can easily stand up to rejection of the Hades myth, which is somewhat accidental to the rest of TR.

This is not only unsurprising, but anticipated by Socrates himself in the text. The theory can also survive rejection of metempsychosis if some alternative account of the preexistence of the soul stands in its place.

Socrates himself does not expect us to accept TR without some account of preexistence, though. Indeed, any view that does not accept a literal form of the preexistence of the soul is inconsistent with the theory. A naturalistic successor to TR that simply upholds a high estimate of biologically innate knowledge in place of PS represents a fundamental break with TR as it is stated in the *Meno*. However, such a view is consistent with a condition innatist view of TR if we propose that knowledge is acquired during gestation or at birth. It is also relatively easy to reconcile with content innatism. The naturalist view, though, is not consistent with prenatalism.

Notes

1. I skip over a view known as *dispositional innatism*, also discussed by Fine, because I view it as too amorphous to fit well into this discussion.

2. Fine is careful to distinguish Meno's contributions from Socrates' statement of the paradox, but this distinction is not necessary here.

3. I will use the terms *metempsychosis* and *reincarnation* interchangeably for the sake of style.

4. Quotations from the *Meno* use the 2011 translation by Cathal Woods. The reference of "those from whom," in prose, to "Persephone accepts compensation for her ancient grief," in verse, is presented as it appears in the translation.

5. I understand the preceding phrase to mean "during the time when he is a human and also during the time when he is not a human," not that there is a time when he simultaneously is and is not a human.

6. I understand this to mean "for any instant during the infinite past, he either was a human at that instant or was not a human at that instant."

Bibliography

Bronstein, David, and Schwab, Whitney. (2019) "Is Plato an Innatist in the *Meno*?," *Phronesis* 64(4), pp. 392-430 [As of Jan 10, 2020, *Phronesis'* website incorrectly dates this issue 2018.]

Fine, Gail. (2014) *The Possibility of Inquiry: Meno's Paradox from Socrates to Sextus* (Oxford: Oxford University Press)

Plato. *Meno*, tr. Cathal Woods (2011). https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1910945

THE EXPLOITATION OF PROPERTY IN RESISTING REDISTRIBUTION

Jalen Mcleod

INTRODUCTION

Poverty, with its attendant suffering and social, economic, and political ramifications, raises some of the most important questions in ethics and political philosophy. The most immediate way to address the prevalence of poverty is to redistribute wealth from those least in need to those most in need. However, there is a persistent objection to this solution, namely that those who currently own goods (or economic output) have authority over them, and that, if they do not consent to redistribution, the goods cannot justly be redistributed. This argument has sparked philosophical discussion for centuries, enticing philosophers to try to understand the nature of what it is to own something, and what criteria establish a legitimate connection between owner and property. In this article I will assume that to own something is just to have authority over it, and instead focus on the criteria issue to show that my analysis provides us with justification for redistribution.

To serve as my foundation, I will be utilizing the theory of property by Hugo Grotius, who claims that property is legitimized by the collective recognition of the community's constituents. Consequently, Grotius posits that all constituents have a claim of ownership to goods that would assist in the preservation of their life since they are a part of the collective. If Grotius is right, then on the matter of redistribution, it follows that people in poverty have claims of ownership to socio-economic goods that would help preserve their lives, regardless of if those goods are owned by someone who does not consent to redistribution. In his

(2009) paper, "Hugo Grotius, Contractualism, and the Concept of Private Property: An Institutionalist Interpretation," Marcelo de Araujo opposes Grotius' view. Araujo argues that rights cannot be derived from people's individual interests, and that rights are only derived from the institution of law. Thus, for people in poverty to have legitimate claims of ownership to goods, their interest must be institutionalized. In order to ward off Araujo's objection and thus advocate for the redistribution of wealth as a right for people currently in poverty, my goal in this article will be to strengthen Grotius' account of the nature of property by supplementing his view with my own analysis.

In section 1 of this article, I will show that the nature of the institution of property has no inherent adjudication process and thereby requires another institution to fulfill this role of determining ownership. When economics serves this adjudicative role, ownership is determined by economic transactions. Thus, merely being involved in or contributing to this transaction establishes a legitimate claim of ownership for an agent.

However, it seems that the above argument does not support Grotius' view because people in poverty are not involved in economic transactions regarding most goods. I will argue that this is not the case. To illustrate the involvement of all individuals in economic transactions, section 2 of the article will show that people's individual interests for property or socio-economic goods can become a right through the institution of economics. My thesis is as follows: a person establishes a legitimate claim of ownership to a good by contributing to its value. That is, in having an interest in a good, a person adds value to that good by contributing to the collective demand for it, and thereby contributing to increasing the good's worth/value in economic transactions. This contribution the agent provides to the good necessarily involves them in the transaction and thereby entitles them to some level of authority over the good. Furthermore, because all persons in an economic community contribute to the collective demand for the

socio-economic goods in that community, all such persons have legitimate claims to ownership of those goods, and therefore have a right to redistribution.

A practical objection can be made even after warding off Araujo. Some communities have grossly inadequate available stocks of goods; thus, equitable distributions will not do much to improve the poverty of the community, and may even make them worse off. Therefore, I will argue in section 3 that there is only one economic community—namely, the global community—which is constituted by all persons and all economic goods. There are no distinct economic communities delineated by nation-state lines because, metaphysically, all the ways of delineating economic communities fail. That is, given our view on the nature of property, there are only two ways to delineate economic communities: namely, the goods of the community, and the agents of the community, both of which fail. Of course nation-state borders are arbitrary, but my argument will show further that nation-state lines have no meaningful basis and are unjust given the relevant circumstances I will discuss. If I am correct, then the aforementioned objection is resolved since all global citizens will have legitimate claims to ownership and a right to redistribution of all goods globally, as there is only one economic community.

SECTION 1: THE INSTITUTION OF PROPERTY

Throughout the history of the philosophy of property, philosophers have conducted their analysis of the subject using genealogy, that is, by telling a historical story to illustrate the genesis of property in order to elucidate its nature. Prominent accounts are those of John Locke, who claims that property became legitimized by first occupancy and labor contribution, and of Thomas Hobbes, who writes that property is a product of sovereign arbitration, which was needed to maintain peace and facilitate economic operations. Most other well-known accounts take on some variation of Hobbes,

while being less descriptive and more normative or abstract. Immanuel Kant, for example, claims that property is legitimized through collective social consent, which ultimately rationalizes to the best possible appropriation for everyone involved, while, according to Jeremy Bentham, property is a product of the institution of law (Waldron 2020).

As I see it, all of these accounts are plausibly true in different socio-economic communities throughout the world at different times. None of these accounts seem like universal truths (neither descriptively nor normatively) in every part of the world throughout all time. None of them disprove the others, and any combination of them even seem compatible in that they could have occurred together to legitimize property in a given community. The same applies to Grotius' view. However, I think Grotius' view gets us closest to the fundamental nature of property. My goal is to put forth an account of the nature of property, within the framework of Grotius' theory, that is both descriptive (in many cases) and normative.

Grotius' view is that property is not a naturally occurring phenomenon; it is a manmade institution which is derived from the collective recognition of the community's constituents. He claims that an institution has no inherent power; it can only derive power from the will of its constituents. Thus, if institutions are necessarily collective, then what makes the products of those institutions legitimate is the collective involvement/recognition of those involved.

Then, in regard to the institution of property, it is the people of a society that collectively legitimize who has a right to own what they own. But since most people will be born into existing collectively recognized distributions of property, and thereby have little authority in changing these distributions, Grotius claims that individual interest—if in the interest of preserving one's life—can be legitimate because their interest in this regard is a natural right in virtue of being a person. And natural rights take priority over

manmade rights because natural rights are God-given, thus being superior to manmade institutions (Araujo 2009).

Here is where Araujo opposes Grotius, claiming that interests cannot become a right via natural endowment. The interest must be institutionalized to be a right; that is, it must be collectively recognized to be legitimate. Araujo cites Jeremy Bentham's notion of rights here, which grounds rights in the institution of law (some kind of collective/societal recognition of governing rules). Even if the interest is of life preservation, all rights must be legitimized through law (Araujo 2009).

With Araujo and Bentham, I support the view that rights are derived from collective recognition—institutionally through law or any collectively recognized institution. But, like Grotius, I think our theory of property should justify people's individual interests entailing a right if the interest is reasonable. Thus, to strengthen Grotius' position, I will show that people's individual interests for property or socio-economic goods can become a right if their interests are institutionalized in economics (the institution of resource production, consumption, creation, and exchange).

Let us provide our own genealogy of property here to illustrate our view. At the genesis of the human species, it was likely the case that individual persons and very small groups claimed goods as they came across them. As time passed these individuals and small groups likely began to encounter each other and express interest in the same goods, thus causing conflicts over who would be the authority of the goods. Once this conflict occurs, there is now a need for some kind of institution, which, ideally, can ethically adjudicate among the conflicting sets of interests and establish a legitimate owner of the goods. Notice that this adjudication process precedes and determines who the owner will be. Thus, the nature of the institution of property is always derived from this adjudicative institution. The nature of the institution of property is such that there must be another institution serving as a source for adjudication/determination of ownership because it has no

inherent adjudication process.

When economics is serving as this adjudicator, we can interpret the conflict in terms of which entity gets to *consume* the goods in conflict, rather than which entity gets to *own* the good. But we can see that there is no distinction between the referent of a consumed good and an owned good in this case; though these are two different concepts, they both refer to the same object here. Thus, it is clear that the adjudicative process of economics, namely, transaction (relating to exchange or interaction between agents), is what will also adjudicate the owner.

Remember, if institutions are necessarily collective, then what makes the products of those institutions legitimate is the collective involvement. Thus, the mere involvement in a transaction and/or the contribution an agent provides to an economic transaction should entitle them to a claim of ownership. Let us now illustrate in detail how agents' interests in economic transactions is contributing to/involving them and thereby establishing a legitimate claim of ownership for them.

SECTION 2: THE EXPLOITATION OF OUR ECONOMIC TRANSACTIONS

Returning to our genealogy, I will put forth a thought experiment to show the involvement of collective agents in economic transactions. At this point in our story, the aforementioned first two entities in the initial conflict have reconciled and established a cooperative economic institution to manage their goods. Let us call them an economic community. Now comes along a third and fourth group of agents who have interest in consuming the same goods, but do not have as much bargaining power or coercive power to engage in a fair transaction with the economic community. After deliberation, the economic community concludes that, because both of these new groups have little to offer to the community, allowing them both into the community would

entail less economic output to be enjoyed by its current members. However, by instead allowing one group in, the community would see greater economic benefits. Group 4 has slightly more to offer in terms of bargaining power than group 3, and, knowing this, the economic community uses the interest of group 3 as leverage in the transaction with group 4 in order to get the most benefit possible in the transaction. Group 3's interest increases the worth of the community, because the two new groups must continuously outbid each other to consume what the community has to offer. If there was only one new group with no competition, at the very least, the worth of the community would only be influenced by two entities as opposed to three.

Group 3 is contributing something to the community via the transaction, but in the end receives nothing for their contribution, while the community receives benefits due to group 3's contribution of value to the goods, which arose merely from being involved in the transaction. Thus, group 3 established a legitimate claim of ownership to the good that was not respected and should receive some return for their contribution—otherwise, they are being exploited.

This exploitation still thrives today in similar fashion; mostly all people have an interest in socio-economic goods, such as high-quality education, healthcare, housing, and are thereby contributing to the worth of those goods in economic transactions while seeing no return for their contribution. If there is collective demand for some good, then all the sets of interests have legitimate claims of ownership to the good.

Of course, these kinds of contributions may seem unavoidable; the fact that people will have interest for goods in transactions cannot be helped, nor can the fact that it is in the best interest of people to make the best possible deal in their transactions. Thus, how do we avoid exploitation? How do we arbitrate over goods with competing sets of interests? An obvious answer is some kind of socialist institution, but a theoretical concern for

our theory is that there could be no private property like houses, cars, and televisions if some other agent has interest in them. Yet, it seems that the things we intuitively view as private (e.g. houses, cars, televisions) are less likely to have high collective demand, and the things that intuitively seem more like collective goods (e.g. healthcare, education, energy) are more likely to have high collective demand. This is the case because when someone justly adds the most value to something, mainly through purchasing it, other interests for that specific object will usually cease or be revealed as nonexistent.

For example, I think I can safely say that there is not anyone else who desires to have my specific television. There may be those who do not have a television but desire one, or there may be those who desire the model of television I have and plan to purchase one, but that do not desire to have my specific television if there are others available. However, there will be instances of something like a house located on a very desirable location, or some other scarce object whose quantity is exceeded by the quantity of desires, in which we will have to arbitrate who it can go to.

There are a few ways in which we can deal with this problem, none of which I will particularly advocate for here. We can say that something is automatically under collective ownership once it has collective demand, so that some kind of socialist adjudication must be implemented for reconciliation; something is private property when it is collectively recognized as such. We could instead view ownership as a spectrum with "private" on one end and "collective" on the other. On this account, the more collective demand something has the more it is pushed toward the collective side, and, conversely, the less collective demand a good has, the closer it is to the private side of the spectrum.

Even if we decide on the spectrum-of-ownership option, redistribution of socio-economic goods is still a consequence of our view. For goods necessary to human life and a civilized state, such as food, water, shelter, education, healthcare, or energy,

we can safely assume that demand will come from all people throughout all times unless the constraints of human biology change at some time in the future. Where all agents in the community are adding value to a good, it seems reasonable to suggest that no agent can outbid the value the majority of people contribute (through demand contribution and/or through their labor contribution), no matter how much purchasing power the sole agent has. Goods that are intuitively private but collectively demanded, like the luxury house or the rare item, will likely not be candidates for redistribution purposes because they are not necessary for the preservation of life.

According to this theory, an economic community is morally obligated to equitably redistribute socio-economic goods to its citizenry because agents of the community have legitimate claims of ownership to these goods. For many impoverished communities, though, redistribution will not combat their poverty and may even make them collectively worse off. However, there are clearly more than enough resources globally to satisfy the basic needs of all peoples, and our argument allows us to elucidate this even stronger claim regarding the nature of property and the nature of economics: namely, that there is only one economic community and, therefore, that redistributive allocations must be designed to meet the needs of those who are worst off among the world's people.

SECTION 3: ONE ECONOMIC COMMUNITY

Based on our argument in section 1, which claimed that the nature of the institution of economics is transactional, it is necessarily the case that economic elements are based on this transactional nature. If we are to analyze the metaphysics of an element like economic communities, then we must look to the transactions of these communities to tell us what an economic community is. When we reduce all the components of a community's transaction

to their most basic parts, we end up with only two: the goods of the community and the membership of the community. However, both these criteria of delineating economic communities fail given the current state of global affairs, and there is thus no possible way of making any delineations. Yet, there is obviously some economic community. Therefore, only one economic community exists.

Let us first show that membership fails to delineate economic communities by remembering our argument from the prior sections. A) An agent establishes a claim of ownership to a good if she is involved in any transactions involving the good. Our example of how this agent becomes involved was that their interest in the good contributes to its worth in an economic transaction. Thus, if a community conducts transactions that involve some agent, then that agent is necessarily a part of that community.

B) Since no economic community's goods satisfies the demand of all its members, every community has interest in the goods of its counterparts. Today, we see that the effects of globalization are causing nations' economies across the world to become more dependent on one another. With the increase in population sizes and advancements in technology/social norms, international cooperation has become vital to maintaining the thriving of communities. This is why trade and geography have become so essential to today's global politics.

C) If the two points made above (A&B) are correct, then agents from all communities are involved in the transactions of all other communities just by contributing to the collective demand of every community. Thus, all agents belong to the same economic community.

And now let us show that economic goods fail to delineate economic communities using the same argument. D) Two goods become a part of the same economic community if they are involved in transactions together, such as being exchanged with one another, influencing the value of one another, or being constituent parts of some produced good. E) In our globalized world,

all communities possess goods that are in some way involved in transactions together. F) If D and E, then all economic goods are a part of the same economic community.

CONCLUSION

If the three arguments I have presented are sound, then I have shown that all people have a right to the redistribution of goods because the nature of property and the nature of economics tell us that all people currently have legitimate claims to ownership of those goods. To do so, I developed Grotius' view on the nature of property, which claimed that all manmade institutions like property are legitimized by the collective recognition of the community's constituents. Consequently, Grotius posits that people in poverty have a natural right to goods if their interest in the goods is motivated by the preservation of their life. I agreed with Grotius' conclusion, but also agreed with Araujo's objection to how Grotius arrived at that conclusion—that rights are not naturally endowed, but are legitimized institutionally through collective recognition.

To strengthen Grotius' argument, I argued that people's individual interest for property/goods can be legitimized as a right when understood through the institution of economics. That is, if an individual's interest in a good is being used to increase the worth of a good, then that agent has a legitimate claim to authority over the object; otherwise, their contribution to the good's worth is being unethically exploited. This is the case because the nature of the institution of property is such that it requires a separate institution to adjudicate ownership. Economics can serve as this mediating institution; therefore, if an agent makes a contribution to an economic transaction concerning property, then that agent establishes a legitimate claim of ownership to that property.

As a result of our view, all economic communities are morally obligated to redistribute their goods according to the collective

claims of ownership the citizens have established. However, for many impoverished economic communities, redistribution may make them worse off because the community does not have an adequate amount of goods to sustain an equitable distribution. To overcome this, I argued that it is the global stock of goods that must be redistributed equitably according to the worst-off global citizens, as opposed to individual nations redistributing their stock of goods just to their own citizens. This is the case because there is only one economic community, that which is constituted by all global citizens. Delineating economic communities based on nation-state lines is nonsensical given that the nature of the institution of economics is purely transactional. Thus, only the transactional elements of goods and agents can delineate economic communities. But neither of these can do so currently, since we live in a globalized world where international cooperation is essential and extensive.

Bibliography

Araujo, Marcelo De. (2009) "Hugo Grotius, Contractualism, and the Concept of Private Property: An Institutionalist Interpretation," *History of Philosophy Quartlerly* 26(4), pp. 353-371

Waldron, Jeremy. (2020) "Property and Ownership," in: *Stanford Encyclopedia of Philosophy*. plato.stanford.edu/entries/property/

WHAT IS PHILOSOPHY, ANYWAY? DELEUZE AND GUATTARI'S PRELIMINARY ANSWER

Sakib Ibne Shahriar

INTRODUCTION

What *is* philosophy, anyway? What have all these people been doing over the course of thousands of years of history, bickering with and sputtering over one another, in person and on innumerable reams of paper?

It is easy enough nowadays to find a quick answer to this on the internet: philosophy "is the study of general and fundamental questions about existence, knowledge, values, reason, mind, and language."[1] This answer is perhaps good enough for a newcomer's *first* answer to this question. Nevertheless, it is not too interesting or satisfying at the end of the day, because it remains *too* general. For example: linguistics is also the general study of language, while psychology is the general study of the mind; sociology examines our social existence, and biology our biological existence; law, medicine, history, gender studies, indigenous studies, and science and technology studies all encounter many fundamental questions about knowledge and values (in fact, nearly *every* field faces its own fundamental questions about knowledge and values!). But, aren't all these fields different and unique from the field of philosophy today? Biologists often enter into dialogue with philosophers on common problems, but biologists do not *need* philosophers in order to be able to do biology or reflect on their own methods and activity in the field. There are no philosopher-rulers or monarchs, philosophers are not intellectual royalty: we indeed are not the arbiters of all legitimate knowledge

in the world, and we are not inherently better at thinking than anyone else in other fields of research and study.[2]

But if all that pomp and pageantry isn't the case (and it isn't), then what is the specific activity that philosophers do? What do philosophers in particular bring to the intellectual table of problems and ideas? And besides, *why* should you care about trying to give an answer to this helplessly huge question of what philosophy is, especially if you are not a philosophy student, and even if you are already pursuing a career in a subfield of philosophy? Why not get on with your own lives and problems, leaving behind a motley ensemble of academics who will continue loitering around this slippery question for the rest of their lives?

PHILOSOPHY'S FIRST CREATION: THE CONCEPT

Well, you should care about this helplessly huge question of what philosophy is, because it has its own specific task that other fields do not take on! For philosophers and non-philosophers alike, having a sense of what philosophy specifically does helps us think about our own problems, articulate our own creative activities, and use resources in philosophy for ends within and beyond philosophy, all the while learning about the context in which philosophy operates. It also helps us learn the specific tasks other fields take on that philosophy cannot. This enables us to think with greater intensity about interdisciplinary research with others. Example: biologists do not need philosophers to do biology, but they often find ideas in philosophy that help them frame problems in biology. At one point in his book *Chance and Necessity*, biologist Jacques Monod uses Henri Bergson's philosophical system to frame the biological problem of how it is that organisms are formed, and to what degree organisms are purposeful actors in their species-specific worlds.[3] On the other hand, philosophers also do not need biologists to do philosophy, but they often draw inspiration from biological functions when thinking about their own philosophical

concepts. In his book *Difference and Repetition*, Gilles Deleuze draws on both the notion of differentiation in cellular biology and embryology, and the notion of differentiation in calculus, to help build his own philosophical concept of difference.[4] In this way, each creative activity and intellectual field has something to offer to the other—resources and ideas from one field can be used in different ways in another field altogether. So, what does philosophy have to offer, both on its own and to other fields?

In their book *What is Philosophy?*, French philosophers Gilles Deleuze and Félix Guattari build a novel understanding of the activity of philosophy that gives answer to many of the questions we have posed so far. Philosophy is "the art of forming, inventing, and fabricating *concepts*" (Deleuze and Guattari 1991, p. 2, emphasis added). This definition calls for two follow-up questions: What is a concept? How do you make one? I will try and answer the first one here; I will not answer the second one here, but let us keep it in mind. A concept is not just any idea whatsoever. It is also not a scientific hypothesis about the world; there are important differences between a hypothesis and a concept that we will explore shortly.

What is a concept? For our purposes, a concept has three components: a concept makes a difference in our thought and our lives; a concept is distinct from a proposition; and a concept can always be repeated across time.[5]

FEATURES OF THE CONCEPT

1. In the most colloquial sense of the phrase, a concept makes a difference. Sometimes, it makes a difference in our lives, and sometimes it makes a difference in how someone frames a philosophical problem. A concept lets us see and live the world in a different way, redistributing the elements familiar to us in new forms (Deleuze and Guattari 1991, pp. 27-28). Concepts may remain abstract and unable to affect anything in the world when

they exist on the pages of a book, but it is us, people, who make concepts act in the world by incorporating them into concrete situations.[6]

Let us look at an example. In their book *Anti-Oedipus*, Deleuze and Guattari make a seemingly unbelievable claim: everything in the world is machines. "Everywhere *it* is machines—real ones, not figurative ones: machines driving other machines, machines being driven by other machines, with all the necessary couplings and connections" (Deleuze and Guattari 1972, p. 1, emphasis original). They do not use machines as a metaphor for what things are *like*; they give a real, actual definition of everything in the world in terms of a machine. For example, our eyes are machines, made up of a variety of tiny, biological parts, and receiving flows of visual information from the other machines around us—another human-machine, the sun-machine, even a philosophy-book-machine. Think of anything in the world, and you can conceive of it in terms of a machine made up of parts, connecting with other machines, and receiving and producing flows of its own.

Here is a common kind of objection to this concept. If someone grants that Deleuze and Guattari are not using 'machine' as a metaphor, but they feel uneasy agreeing with their definition, they can make two possible counterattacks. They can say that the affirmation "Everything is machines" cannot correspond to anything in the world. We cannot scientifically test whether or not something is a machine (unless, perhaps, we are looking at literal, industrial machines), let alone discover whether *everything* has the quality of being a machine. We cannot use this affirmation as a hypothesis, either: since there is no way to confirm (or disconfirm!) this affirmation through empirical observation, it cannot *even* count as a well-constructed hypothesis.

Both of these counterattacks are technically right. In the affirmation, "Everything is machines," Deleuze and Guattari are not referring to specific things in the world the way a science may,

and they are not creating a testable hypothesis about the world. All they are doing is creating a conceptual difference. This difference is not something that refers to anything in the world, and it does not hypothesize anything about the world. Rather, it is a way of looking at the world. Take a moment to join me in a thought experiment. Try and imagine that the items around you, the people you love, and you, yourself, are all special kinds of machines. Machines have connections with other machines: the sun's rays connect with the earth, or a person makes an acquaintance with another person. You connect with other machines, and sometimes you break away from other machines. There are many kinds of flows between machines: the utterance of a sentence, the water falling out of a sink, or the movement of cash from a check to a bank account. All of you produce your own flows, and sometimes you receive flows from other machines. To imagine this is enough to show that the *job* of a concept is to make a difference: to make sense of the world in a different way. The concept is an imagination of the world, whereas a hypothesis is a testable claim about the facts of the world. A concept is not a hypothesis, and vice versa. Again, we will have more to say on this later.

Of course, this concept of machine is useless if it is not applied in concrete situations. What are these concrete situations? There are many, but I will only speak of one: an individual life. Earlier, I asked you to imagine your world in terms of machines. This is fun enough on its own as a game of imagination, but concepts do not exist simply for fun and games. Deleuze and Guattari create this concept of machine as a partial response to, among many other questions, this one: what does it take to live an ethical life? How do we desire with other people? When your existence and mine are framed in terms of machines, we get to ask many different ethical questions that can bear on how we live our lives. As machines, when and how do we break down? Are breakdowns a normal part of who we are, and if so, how do we live them? How do I make a connection with another machine without

hurting myself or the other? What is our status as machines within the capitalist economic system? How are we being used against our own will and consent, and how do we escape that use and manipulation?

The wonderful thing about a concept is that you do not *have* to bring the concept to bear on your life. Nevertheless, you *always* have a choice to make when faced with a concept: will you let this concept make a difference in how you see the world, and in how you act and move within it? Or will you leave it behind, looking for another concept that will cut up the world in a different way than the old concept did?

2. From this preceding discussion, let us grant the following point: concepts are not propositions. Sometimes, sentences that express concepts read like propositions, but it is not accurate to consider them as such.[7]

Take a look at the following example. The proposition "Snow is white" has a referent in the world: it makes a truth claim about whether or not a certain material, snow, has a certain quality, the quality of being white. This is fine and well: what it is to be a proposition is to have an external reference, or at least claim something about an external referent beyond the proposition (Deleuze and Guattari 1991, p. 22).

If we try to take "Everything is machines" as a proposition, we run into a whole slew of issues.

> "Okay, it's a claim about literally everything in the world. Usually, in logic, claims that are this big are circumscribed in a smaller 'universe of discourse,' or a domain of reference, something that narrows down what context we're talking about, So, the everything in this proposition may refer to "everything on this table," or perhaps "everyone inside this restaurant." But Deleuze and Guattari's 'universe of discourse' is, literally, the entire universe as it stands, known and unknown. How do I test that? Clearly, every-

thing is *not* a machine, because I'm a human, and I eat, drink, and sleep, and I'm an animal among other animals. Well, maybe I'm a machine according to this specific definition of the word "machine," but how do I test that with everything in the world? I can't test it for everything, and so this proposition will never have a definite truth value. And besides, doesn't it feel arbitrary to rely on definitions that I can change at any time in order to 'successfully' say something about the world? I can change my definitions ad-hoc to accommodate anyone's attempt to falsify my claims."

Of course, all this misses the point of the concept, and it especially does not do justice to the concept of machine. When we try to think of a concept as a proposition, we remain "constantly trapped between alternative propositions" (Deleuze and Guattari 1991, p. 22). Since the concept does not refer to a series of facts, it slips out between our fingers when we try and find a referent for it. As we saw in our thought experiment earlier, our question should not be, "Does this concept succeed in referring to anything in the world?" Our question should rather be, "What happens if I start seeing myself and the world through the viewpoint offered by this concept?" The concept does not refer to anything in the world, but it can always be brought to bear in the world, changing how we act and move in it. It reorders how we live our lives. We have already seen that the affirmation "Everything is machines" is capable of doing this work, so long as we are willing to let it do that work. There is no analogous question of bringing a proposition like "Snow is white" to bear on our lives; it is either true or false, regardless of how we individually want to see the world. It does not change *how* we see the world, but it lets us claim something about properties of objects that are already in the world. A proposition does not do what a concept does, and vice versa. We would do well not to confuse the two for one another.

3. A consequence of this fact, that concepts are distinct from propositions, is the following: a concept can always be repeated across time. "If one can still be a Platonist, Cartesian, or Kantian today, it is because one is justified in thinking that their concepts can be reactivated in our problems and inspire those concepts that need to be created" (Deleuze and Guattari 1991, p. 28). The history and current field of philosophy is full of people reusing and reactivating the concepts of previous philosophers. Philosopher and artist Adrian Piper uses Kant's concepts of reason and the self in order to pose the ethical question of how we live with one another. Ethicist Iris Murdoch revives Plato's concept of the Form, using it to create a concept of perfection that can only be shown, but never attained, guiding our actions over time. Concepts like Leibniz's possible world, or Aristotle's soul, have this timeless quality to them, allowing them to be reactivated in the present time, in different contexts than their original conceptual homes. Modern-day Platonists and Thomists appeal to old concepts in order to make sense of modern problems in epistemology, ethics, and politics.

Of course, as we have seen, it is precisely *because* these concepts are not hypotheses or propositions, and because they do not refer to any *thing*, that they are meaningful in their own way and can be reactivated and brought to bear on our lives. The concept's only object is itself. As it turns out, a sentence's not referring to something in the world is a poor criterion on its own for whether or not that sentence is meaningful. What kind of sentence are we concerned with? If, even now, our beloved interlocutor remains convinced that concepts are propositions and, thus, are meaningless, we ask that they consider taking a break from that belief and trying to think about a concept in its own way anyway. We can only argue with each other for so long before we decide to work together on a common problem. Let us put our collective imagination to good use! Different sentences are meaningful for different reasons: philosophical sentences cannot be

judged against scientific criteria, and scientific sentences cannot be judged against philosophical criteria.[8]

PHILOSOPHY CREATES CONCEPTS, BUT WE HAVE NEW PROBLEMS

So, we have now established a particular concept (!) of the concept. A concept makes a difference, and it is precisely because it is not a proposition that it can repeatedly make a difference in different lives and contexts across time. A concept lets us see reality through a particular viewpoint, one that can cut up reality in a unique way, granted that we let it do so. Philosophy's task is to create these concepts, differences that we can inject into our lives. In this sense, philosophy is a deeply ethical project, even in its wildly different expressions in history. Great! We have now given one answer to the helplessly huge question of what philosophy is.

Yet, now we arrive at a different issue we must think about! (This is common in philosophy—solving certain issues tends to lead to new ones). Given this account of the concept, the concept cannot be the only element philosophy creates. The history of philosophy is rife with contestations across time, with philosophers challenging one another on what philosophy ought to be and do. If philosophy's only creation consisted in ideas that made a difference in our lives, then what would there ever be to disagree about? There would only be a question of whether one is willing to let a concept affect how they live in the world. What sense does it make to argue against another's concepts as true or false if the concept is not a proposition making a truth claim, and if the primary aim of a concept is to make a difference? Surely, this is not all that philosophers are doing when they argue with one another. But if it is not right to understand the contests between philosophers as a disagreement about concepts alone, then what comprises philosophical contestation?

There is another issue as pressing as this previous one.

As it stands, this exposition makes it seem that one can choose concepts without rhyme or reason to combine with one another, and once they have done so, they will have created a philosophy. This appearance seems suspicious. The choices of concept philosophers make do not seem to be arbitrary; they seem to be at least partially determined by a greater problem, whether that problem is in philosophy or in a different field of study. Philosophers do not create concepts "just for fun," as it were (though there is a lot of fun to have with concepts). It seems that a philosopher's concepts are unified with one another; they are made to hang together as one consistent project, and the concepts are made to communicate and resonate with one another. What does it take to unify concepts in this way? What does it take to ground the use of one's concepts?

PHILOSOPHY'S SECOND CREATION: AN IMAGE OF THOUGHT

The concept is not the only element philosophy creates. Philosophy also creates an image of thought. An image of thought determines what it means to think, and more specifically, what it means to create philosophical thought: "the image thought gives itself of what it means to think, to make use of thought, to find one's bearings in thought" (Deleuze and Guattari 1991, p. 37). This image of thought is a kind of context and home that concepts can live in; it is where a philosopher's concepts live before those concepts are brought to bear on a concrete life.

So far, we have spoken of isolated concepts making a difference. More often than not, in an author's philosophy, one finds multiple concepts, each of them strung through with arguments, conceptual bridges, analogies, and resonances. We go from Plato's concept of the soul, to his concept of the city, of representation and participation, the forms, justice and good, virtue. We need not be familiar with what each of these concepts are, for the time being; it suffices to say *that* Plato connects many of these concepts together

throughout his writings. These concepts are connected and unified in certain ways unique to Plato. It is not enough to say that they are connected simply by virtue of being signed by the author—in this case, Plato.[9] Rather, Plato creates an entire way of thinking that is capable of linking the concepts together with one another: "Concepts are the archipelago or skeletal frame, a spinal column rather than a skull, whereas the [image of thought] is the breath that suffuses the separate parts" (Deleuze and Guattari 1991, p. 36). Or, a different image: concepts are like the various colors, lines, and textures that can be used when creating art, but the image of thought is an individual composition, an artistic piece, that brings those elements together in a specific way. The composition unifies these artistic elements, and it makes their function specific to itself and the context that it creates. Even if two artists use the same colors, they will use the colors in different ways and to different ends. Two philosophers may use the same concepts, but they will use them for different purposes altogether, and they will articulate those concepts in different ways according to what is demanded by their respective images of thought.

The image of thought is what keeps a philosopher's choice of concepts from being arbitrary. No philosopher does their job from a disinterested viewpoint, even if they may try to claim otherwise. Perhaps a philosopher is concerned with who God is and how a person can have a relationship with this God; perhaps a philosopher is unconvinced of God's existence, and they are instead concerned with what it means to negotiate an ethical life in a world devoid of the divine. Both of these philosophers will have specific concepts of God, but the way in which they articulate those concepts of God will be specific to each of their respective viewpoints of the world and of what philosophy ought to think. Furthermore, it is because the two philosophers will have different images of thought that these two concepts of God are separate from another—they are not unified in the same philosophical system. It may be the case that a philosopher works

with no concept of God at all; they may be interested in a part of philosophy that has nothing to do with religion, personal ethics, or humanity. This philosopher will use and unify a different set of concepts than those of the previous two. The image of thought a philosopher lays down helps them coordinate what concepts they must create that will suit the affirmed image.

A concept makes a difference, and it does not refer to objects the way a proposition does. A concept does not make a truth claim, and thus, it makes little sense to agree or disagree with a concept. One either brings it to bear in their own lives, or they leave it behind and move on to different concepts. An image of thought similarly does not make any truth claims the way propositions do. Yet, unlike concepts, an image of thought *does* make a claim. An image of thought makes a moral claim: what deserves the title of philosophy, what kind of thought is strong enough to deserve the official stamp of philosophical thought? What should philosophy do in the world?

It is at this point, when this question of philosophy's nature, activity, and use in the world is posed, that philosophers are capable of contesting one another.

The history of philosophy is rife with philosophers advancing what philosophical activity should and should not look like, and what kinds of philosophical works do not deserve to be taken seriously. Though a philosopher may acknowledge that many pieces of writing are called philosophy in *fact*, they may not consider many to be philosophy by *right*: "what pertains to thought as such must be distinguished from contingent features of the brain or historical opinions" (Deleuze and Guattari 1991, p. 37). What pertains to thought as such must be distinguished from that which, as a contingent fact, *happens* to go by the name of "thought" in the world.

There is a certain image of thought one can find in the early history of analytic philosophy. There, philosophy's task is to analyze language, whether that is the language of logic and prop-

ositions, or the ordinary language of day-to-day life. Philosophy must analyze language and, in doing so, clear away confusions and falsities philosophers have created by affirming senseless propositions, or by getting confused by the grammar of one's language. In his book *Language, Truth, and Logic*, A. J. Ayer notes that if the philosopher wants to contribute something specific to the intellectual domain that no other field can, they "must, in fact, confine [themselves] to works of clarification and analysis" (Ayer 1936, p. 37)—the clarification and analysis of propositions, the language we use to affirm them, and the criteria we use to judge them. Any work in the history of philosophy that does not confine itself to this task must be considered senseless metaphysics. In the *Philosophical Investigations*, Ludwig Wittgenstein does not advance philosophy as an activity that takes on a single task: "There is not *a* philosophical method, though there are indeed methods, like different therapies" (Wittgenstein 1953, p. 51). In his own words, philosophy leads the fly out of the fly-bottle, and it accomplishes this with many different kinds of signs and lures.[10]

If you have been following along the conversation so far, you will notice that in this paper itself, I have been advancing my own image of thought from the very beginning, one that is not quite Ayer's or Wittgenstein's. Philosophy *does* have a specific task, a task different from those taken on by the sciences and the arts. Philosophy creates concepts, and concepts make a difference in concrete lives.

Let me at this point clarify something. By concrete lives, I do not want to necessarily confine us to talking about human lives alone. A tradition or body of thought has a concrete life of its own, a life with its own movements, blockages, and contours. When a philosopher is engaged with a tradition of thought, they may face conflicts between what they desire to do with philosophy, and what their tradition of thought allows philosophy to do. In such a situation, a philosopher may succeed in creating a concept that makes a difference within that tradition of thought, allowing that

tradition to take on new aims and carry new sense and meaning. Though my image of thought does not align, for example, with Wittgenstein's, my image can accommodate the fact *that* Wittgenstein engaged in philosophical activity—he created concepts. His concepts of the language-game, rules, meaning as use, and private language, make a difference in how we, people reading his work, see ourselves and the world in which we live. They also make a difference in how philosophers, and particularly analytic philosophers at the time, ought to think about language and meaning. Through his concepts, Wittgenstein reframes what it is to think and work with language. His concepts make concrete interventions and redistributions in the field of the philosophy of language.

I noted that my image of thought can accommodate the fact that Wittgenstein engaged in philosophical activity. There is more to this. I desire to build an image of thought that can accommodate for the following fact: though the history of philosophy is made up of people advancing *different* ideas about what philosophy is and what it ought to do, they were all nevertheless doing philosophy. Each of them is deserving of the title of philosopher, because they successfully took on philosophy's creative task. The image of thought is where philosophers are capable of contesting one another, as the image of thought makes a moral claim. Nevertheless, we are capable of creating images of thought where contestation is not our primary aim. At the end of the day, even though my image of thought differs from those of many others, I do not want to affirm mine over everyone else's. I only want it to be acknowledged that my image of thought *exists*, that it is a viewpoint that was created, and that this viewpoint is capable of saying something interesting, or different, about the world.

The image of thought makes a moral claim. Once again, let us reiterate that philosophy, in its wildly different expressions, is a deeply ethical project. At this point, this affirmation takes on a different meaning: to do philosophy—to lay down an image of thought—is an ethical project because we do not do philosophy

alone. The world is, in fact, full of philosophers advancing what philosophy ought to think, or at the very least, what philosophy must think for their purposes. When a philosopher affirms an image of thought in their work, they must negotiate this affirmation with others: *Am I affirming this image as superior to other images? Why? Should I be concerned with claiming superiority, or should I rather care more about working and thinking with others? Do I have to affirm my image of thought as correct over others in order to do work with it? What other options do I have?*

That different philosophers affirm different images of thought, and that they affirm those images in different ways, indicates that philosophers must continually negotiate their images of thought with one another. We do not have the luxury of neglecting another's philosophical thought. We have it too easy when we dismiss another's work as senseless, ridiculous, or unimportant. We also cannot spend all our time bickering with other philosophers. All of this does not mean that every philosopher is fundamentally right in their own personal way and beyond criticism. It is only a reminder that ultimately, in philosophy as in all other parts of our lives, we must work with one another. We must reckon with the work that each of us creates. We do not see an easy way out of this ethical task. We must take it on repeatedly, across all time, as we try and create new, less oppressive, and more life-affirming ways to live with one another.

Notes

1, This answer is taken directly from the Wikipedia page on philosophy. This paper is not addressed to an academic audience; it is addressed to a non-academic, philosophical non-specialist. I am using such an introductory definition because it seems to be one of the first definitions a newcomer to philosophy may come across today.

2. Perhaps we could try defining what philosophy is by looking at the word's etymology. The linguistic root of the word "philosophy" comes from the Greek "phílos", which means "love", plus the Greek "sophía", which means wisdom—thus, philosophy would mean the love of wisdom. Perhaps philosophy, at some point in history, *did* mean the love of wisdom. Yet, as we

already see, it turns out that *many* fields today beyond philosophy also love wisdom and knowledge, each in their own unique ways!

3. If you are interested in reading more, see Monod's *Chance and Necessity*, Chapter II, "Vitalisms and Animisms."

4. If you are interested in reading more, see Deleuze's *Difference and Repetition*, Chapter IV, "Ideas and the Synthesis of Difference," especially pp. 208-221. Best wishes if you decide to crack this book open—it is difficult reading.

5. As we will see later on, these three components all make the exact same point: because concepts do not refer to anything, you can always reactivate a concept in your life, no matter how old the concept is. These components become indiscernible with one another in the concept (or rather, in this case, in the concept *of* the concept!). Deleuze and Guattari: "Components remain distinct, but something passes from one to the other, something that is undecidable between them. There is an area *ab* that belongs to both *a* and *b*, where *a* and *b* 'become' indiscernible" (1991, pp. 19-20).

6. As we will see later, concepts do not even remain abstract on the pages of a book; there, they live alongside other concepts, being connected to them in different ways, and engaging in processes of change and resonance. This makes the book of philosophy a machine of sorts, one with its own way of moving in the world.

7. A good example of this phenomenon is Descartes' affirmation, "Cogito ergo sum," or "I think, therefore I am." If this is taken to be two propositions, with a premise and a conclusion, this reads as an incomplete argument, one without any supporting premises or axioms.

8. Deleuze and Guattari: "...the philosophical concept usually appears only as a proposition deprived of sense. This confusion reigns in logic and explains its infantile idea of philosophy. Concepts are measured against a 'philosophical' grammar that replaces them with propositions extracted from the sentences in which they appear" (1991, p. 22). Though this line of thought is concerned with the confusions between philosophy and logic, the gist of the message also applies to the relationship between philosophy and science.

9. There are cases of single philosophers having breaks, shifts in their thought, shifts strong enough that their works before the shift seem irreconcilable with their works after the shift. I am thinking of Wittgenstein: despite both his early works and his later works being signed under the same name, Wittgenstein uses concepts differently between his earlier works and his later works. There is a transition in his works happening between different images of thought, and this transition expresses itself in his different usages of the concepts of language, meaning, and philosophy.

10. For my more academic, specialist audience: I am aware that I am placing the later Wittgenstein's works within the history of analytic philosophy. I am also aware that his later writings may position him much closer to conti-

nental philosophy than to analytic philosophy. I do not concern myself too much with the difference between these two philosophical coordinates. I do not have much interest in hashing out the specific conceptual boundaries between analytic and continental philosophy, if there even are any important ones. It suffices to say that given that Wittgenstein's history is irremediably tied with the history of analytic philosophy and the Vienna Circle (no matter his attitudes toward either, and especially if he theoretically broke away from these very two elements), it is not overstepping bounds to include his later works as an expression in the history of analytic philosophy.

Bibliography

Ayer, A. J. (1936) *Language, Truth, and Logic* (London: Penguin Books)

Deleuze, Gilles, and Guattari, Félix. (1972) *Anti-Oedipus* (Minneapolis: University of Minnesota Press)

Deleuze, Gilles, and Guattari, Félix. (1991) *What is Philosophy?* (New York: Columbia University Press)

"Philosophy," (n.d.). *Wikipedia*. https://en.wikipedia.org/wiki/Philosophy

Wittgenstein, Ludwig. (1953) *Philosophical Investigations* (Oxford: Basil Blackwell)

WHY KANT WE JUST KILL OURSELVES?

Jen Markewych

INTRODUCTION

The vast majority of writing on suicide either deems the act immoral or the product of mental illness. Although some contemporary philosophers have taken opposing views, these views have not taken hold in society at large. At present, institutionalized views of suicide allow that a suicidal person who is viewed as a potential harm to themselves to be taken into involuntary custody by the state for "treatment." I shall present a view that suggests that there is no moral problem with suicide and, therefore, that the institutionalization of persons with suicidal ideology ought not to be automatic. On the contrary, in fact, all persons *should* consider the managing of their own death and that, in some cases, *not* committing suicide is the *actual* moral wrong. I shall begin this argument by imagining a Kantian Kingdom of Ends in which suicide is perfectly permissible and expected and then analyze whether any argument remains robust against the act. I shall then use a party analogy to show how and when one must commit suicide, and when one is actually morally prohibited from doing so. Finally, I shall conclude by showing that anti-suicide positions are *actually* immoral, using the party analogy; if one does not wish for a person to commit suicide, one has a moral obligation to create a space in which a person wishes to stay, rather than condemn them for leaving.

I. CONTEMPORARY CULTURAL CONTEXT

It is not the intention of this brief article to review all of the religious, philosophical, and scientific writings on suicide up until this point; however, it is important to understand the aforemen-

tioned views in considering the actual current climate in which one discusses suicide today and to define the act clearly and in unambiguous terms. Broadly, all the major world religions expressly forbid or condemn suicide, considering it an affront against the gift of life that God has given, a violation of rules against murder, or an affront to doctrines of non-violence. Although some religious orders have become more progressive, allowing for the possibility that a person who committed suicide may also have asked for forgiveness prior to the moment of death or been so mentally ill as to not be held accountable for their actions, there is no significant religious dogma that allows for the rational contemplation of taking one's own life.

In considering views on suicide from mental health professionals, the American Psychological Association (n.d.) assigns a circular definition to the act itself, "Suicide is the act of killing yourself, most often as a result of depression or other mental illness." In the view of mental health professionals, it is nearly impossible to be a person contemplating the management of one's own death without any mental illness. The Mayo Clinic (n.d.) goes further, discussing suicidal ideology on their website under a page entitled, "Diseases and Conditions," and begins, "Suicide, taking your own life, is a tragic reaction to stressful life situations—and all the more tragic because suicide can be prevented." Even typing the word 'suicide' into any major internet search engine returns the large-print results for the 1-800 number for the National Suicide Prevention Hotline first, before any other result. Causes of suicidal ideology are listed exclusively as related to underlying mental disorders, trauma, or addiction; a person contemplating suicide must have a treatable condition and ought to be seeking the services of a professional specializing in that form of disorder. It is *nearly impossible* to discuss rational reasons for ending one's own life with a mental health provider without drawing a diagnosis... and also likely at least a 72-hour hold in a psychiatric facility for further evaluation.

A common moral view in society regarding suicide is that it is an act of cowardice, a choice made to escape hardships that others opt to stay and face. The suicidal person can be accused of transferring their own pain and suffering onto others who are left to tend to their affairs or suffer their loss. Further, the person who has completed a suicide is viewed as having failed to fulfill their own obligation to "reach out" and prevent their own demise. Rarely is it the case that when we hear of a person who has committed suicide, do we collectively shrug our shoulders and respond with an unemotional, "Meh. Makes sense." Even in the cases where suicide is the most obvious and reasonable choice, collectively we scrub the facts in order to remove the castigation and negative views associated with the choice.

For example, on September 11, 2001, when two planes hit the World Trade Center in New York City, a significant portion of the resulting fatalities were attributed to individuals jumping from the towers and, thus, subsequently ought to be considered suicides. Individuals trapped on the upper floors of the towers *chose* to jump and end their lives instead of waiting for their probable demise or possible rescue, and yet not a single fatality from 9/11 was classified as a suicide by the Office of the Chief Medical Examiner. Despite evidence that nearly 200 people jumped from the towers, little effort was made by the 9/11 Commission to identify those individuals (Smith 2011, p. 40). Further, a public outcry in response to the alleged insensitivity of publishing photos of the jumpers all but erased those who jumped from the immediate public record. In essence, even when the choice to end one's life instead of continuing to suffer is so blatantly and obviously defensible, even when there is absolutely no moral call to push on, the stigma around suicide caused by institutionalized religious and psychological views is so great that we are inclined to scrub the facts surrounding it from history. Regardless of various philosophers' arguments surrounding suicide, it is an absolute fact that social institutions have ratified the view that suicide is wrong and

harmful in all cases. The institutionalization of these views causes greater harm than the act of suicide itself and should be subject to greater scrutiny.

II. A CASE STUDY USING KANT AND HUME

Before proceeding further, what an actual completed suicide in modern society looks like today ought to be considered without any moral hand-wringing regarding the action itself. Let us consider Joe Citizen, who decides to kill himself. For the moment, we will not consider Joe's motivations or mental health in this decision, only that he has decided that he no longer wishes to live. Due to the aforementioned institutionalized views regarding suicide, Joe must keep his decision a secret. Although he may be able to put some of his affairs in order, because of the widespread awareness of suicide prevention in modern society, Joe cannot have lengthy conversations regarding his decision or wonder aloud if it is the right decision at all. He cannot turn to his friends and loved ones to say goodbye or tend to unfinished business. Joe has to make the most important decision of his life in near silence, aware of the fact that if he shares too much with the wrong person, a quick call to 911 will have him in an ambulance and off to a psychiatric facility.

Thus Joe, who may already be suffering, knows that he likely can have no impartial conversation on the matter, and keeps it to himself. When Joe does kill himself, he will be alone. Whether by overdose, carbon monoxide, hanging, or gun shot, he will be deprived of all the comforts that society offers to those in hospice or, for that matter, to an unwanted dog in an animal shelter. Joe's death will certainly be lonely, may be painful, and worse yet, possibly ineffectual. Joe, like every other person on the planet, can have no experience in taking his own life effectively and may make a mistake. A misplaced bullet, incorrect dosage, or interruption by another may send Joe to the hospital, committing him to a

lengthy and agonizing recovery or slow slide to death. Even if Joe is successful, there still remains the issue of, well, his remains, for which he cannot properly make plans without generally arousing suspicions.

Because Joe has not been afforded the luxury of choosing his death in a forthright manner, *someone*—a parent, a child, a spouse, a sibling, ultimately likely a loved one—is condemned to the horrible reality of discovering Joe unexpectedly, swinging from a rafter in the basement, brain matter splattered in the bathroom, or floating bloated in the pool. Regardless of the ethics around death and suicide, it is not hard to argue that seeing a loved one in this condition is—in and of itself—harmful. Of course, Joe, trying to mitigate harm as much as possible, could foresee this and could also wander off into the wilderness or fall into the ocean, leaving instead the terrible weight of the unknown on his survivors. This does not even begin to address the mundane aspects of Joe's life that are left to be addressed by Joe's survivors: the disposition of his job, home, assets, etc. When considering the timeline of Joe's suicide, the moral wrongness of the action does not seem to settle on the act itself, but on the harm to Joe, his loved ones, and society by virtue of the fact that Joe could not openly discuss his decision and facilitate his death in a manner most comfortable and convenient to all involved.

But what does philosophy say about our man Joe, contemplating terminating his own existence? Of the early modern philosophers writing during the seventeenth and eighteen centuries, only David Hume and Immanuel Kant wrote on the matter of suicide at any great length. Hume wrote an essay entitled, *Of Suicide*, in 1757, although out of "an abundance of prudence," kept the essay out of his published works. It was published in 1777, a year after his death. Hume was rightfully hesitant to publish, as his essay resisted the bulk of philosophical writings up to that point which argued against taking one's own life. Kant addressed suicide nearly twenty years later in 1797, when he wrote *Fundamental*

Principles of the Metaphysic of Morals, where he argued that suicide would violate his categorical imperative and, therefore, was unethical in all circumstances. Although Hume's arguments were radical at the time, they appear more cogent today, whereas applying his own categorical imperative, Kant might reverse his views on suicide if writing in a modern context.

Hume (1757) begins his article, "Of Suicide," underscoring the importance of philosophy as a whole in banishing the superstitions and failures of thought that plague the common person. Hume also acknowledges that death puts an unreasonable fear in humans, such that cowardice prevents them from taking their own life. This cowardice, combined with some form of superstition, prevents people from taking their own lives. However, Hume (1757) sets out to "endeavor to restore men to their native liberty, by examining all the common arguments against suicide, and showing, that that action may be free from every imputation of guilt or blame; according to the sentiments of all the ancient philosophers." He argues that without superstitious inclinations, "If suicide be criminal, it must be a transgression of our duty either to God, our neighbor, or ourselves" (ibid.). Hume considers each aspect of the argument more thoroughly before concluding that, although "no man ever threw away life, while it was worth keeping," there is no moral imperative to prevent suicide (ibid.).

The first—and most common—argument that Hume considers in his meditation on suicide is whether it is somehow an affront to, or forbidden by, God. Hume was a known agnostic who founded all his arguments in empiricism and naturalism. The general zeitgeist of the time when Hume was writing was to argue that suicide somehow upsets the natural order of things and/or was taking the providence of God into man's hands. For example, John Locke (1690) argued, "For men being all the workmanship of one omnipotent, and infinitely wise Maker; all the servants of one Sovereign Master, sent into the world by his order, and about his business; they are his property, whose workmanship they are,

made to last during his, not one another's pleasure." Hume (1757), however, did not accept this argument made by Locke and others and wrote:

> Every event is alike important in the eyes of that infinite being, who takes in, at one glance, the most distant regions of space and remotest periods of time. There is no one event, however important to us, which he has exempted from the general laws that govern the universe, or which he has peculiarly reserved for his own immediate action and operation.

Although he expounds on this point at length, his ultimate position is that there *can be no* action outside of the natural order that a higher power created. If a man takes action to preserve his life, it is part of the natural order; if a man takes action to end his own life, it must also be part of the natural order.

Hume then stitches together the arguments against suicide being a crime against society and the notion that suicide could be a crime against oneself. He argues that a man only has an obligation to contribute to society for so long as he is a member of said society and receiving its benefits. A man removing himself from society and therefore no longer receiving its benefits is not an action that we would call morally harmful, and this is how Hume considers suicide. Further, a person who is a harm or drain on society could be considered morally righteous in committing suicide. Hume writes, "But suppose, that it is no longer in my power to promote the interest of the public: suppose, that I am a burden to it: suppose, that my life hinders some person from being much more useful to the public. In such cases my resignation of life must not only be innocent but laudable" (ibid.). Hume extends the basic foundation of his argument that one may have no moral obligation to society to prevent suicide to one no longer having a moral obligation to oneself to persist. He writes, "That suicide may often be consistent with interest and with our duty to

ourselves, no one can question, who allows, that age, sickness, or misfortune may render life a burden, and make it worse even than annihilation" (ibid.). Hume recognizes that if a man's future contains nothing but projected misery, it may not only be morally allowable to commit suicide, but it could be argued that we have a duty to ourselves to prevent our own misery. Although Hume's arguments were fairly radical for his time (again, he did not wish to publish them for fear of response), they seem perfectly plausible today.

In contrast, in 1797, Immanuel Kant wrote, in *Fundamental Principles of the Metaphysic of Morals:*

> Now we see at once that a system of nature of which it should be a law to destroy life by means of the very feeling whose special nature it is to impel to the improvement of life would contradict itself, and therefore could not exist as a system of nature; hence that maxim cannot possibly exist as a universal law of nature, and consequently would be wholly inconsistent with the supreme principle of all duty.

Kant makes it abundantly clear in his writing that, on all counts, in all circumstances, suicide is unequivocally wrong and violates his categorical imperative.

Kant (1797) considers the potentially suicidal person solely as "a man reduced to despair by a series of misfortunes [who] feels wearied of life". If the case were this simple, perhaps Kant may be right. It may be profoundly reductionist to evaluate Kant's position on suicide in relation to history, geography, and culture, but it is not unreasonable to take a moment to consider the circumstances in which Kant was writing. One does not need to understand the nuances of 18th Century European history to know two vital facts that would have been profoundly influential in Kant's life: he was born only three short years after the Great Northern War in which hundreds of thousands were either killed or died of disease or famine; and during his lifetime and before his birth,

there were several large scale plague outbreaks that could easily decimate the population of entire regions in a very brief period of time (Black 1999). Kant lived in a world where death *could not* be staved off. It was ever present and could sweep through, taking everyone in its path with it. When considering what Kant would have been aware of in the world, one can almost hear a sneering nature that modern psychologists would blanch at when he describes our suicidal man's "series of misfortunes". Let us not forget, Kant was one of nine children, only four of whom lived to adulthood.

When considering life under these circumstances, it becomes quite plausible to hypothesize why Kant would take the position he did. If one were a thinking person and saw the devastation around him, how could one not consider ending one's life? In Kant's world, it seems quite likely that suicide would violate the categorical imperative simply because, if the maxim of "My life has become unpleasant, I shall end it now without further consideration" were universalized, it becomes difficult to imagine that this would not be a delightful option in the face of plague, famine, and invaders. Kant not only would have to make the argument that he made, but he would have to make it in the strongest terms possible; not only would suicide be a violation of his very theoretical categorical imperative, but it makes sense that he would say that one had a duty to oneself to carry on. It was not enough for Kant to make what, in essence, is a slippery slope argument against suicide; he created an *a priori* argument against suicide by saying one has a duty to oneself to not do it. It is not Kant's strongest argument, but considering the time in which he was making it and the audience to whom he was making it, it becomes clearer why he chose his line of reasoning. In Kant's world, death was ever present, likely to come for you soon in a very unpleasant fashion, and, had he successfully made a convincing argument *in favor* of suicide, Europe could have easily looked like Jonestown, which might have made Kant morally culpable—but this is

another matter.

The question to now consider is whether Kant's argument is still legitimate today and applicable to our man Joe. The answer is likely no. Kant's argument rests on an unstated, but what would then have been obvious, proposition: Death is looming imminently and unavoidably. In 2020, in at least most of the "Western world," this is mostly false. Even discounting the biomedical precipice that society is rapidly charging towards, where death may be avoided indefinitely, death is mostly something that can be fairly safely avoided with some minimal precautions. One can generally avoid war, pestilence, catastrophic trauma, and many diseases. Even when these things cannot be avoided, social welfare systems and tremendous medical advances stave off death with increasing efficiency.

In considering whether Kant or Hume's views are still applicable today, we must consider some basic facts. Currently, the average American life expectancy is estimated to be 81 years; however, that data does not take into account the exponential advances in medicine or individual history (Dobis, Stephens, Skidmore, and Goetz 2020). The number is likely to be a gross underestimation. However, it is significantly less likely that, up until one's death, *all* of those years will be spent in delightful pursuits of self-improvement or contributing meaningfully to society. There is likely to come a tipping point at which one experiences a slow and steady physical and cognitive decline during which Kant's notion of life improvement is no longer possible. If we apply Kant's view to suicide, then we all must slowly wait for death to come while becoming an increasing burden on our loved ones and society as a whole without contributing meaningfully or improving upon ourselves. Does not this then turn the categorical imperative upside down? In our current state, it is arguable that not only should one *not* not take one's own life, but that one actually *ought* to choose a time and place, put one's affairs in order, and perhaps host a lovely going away party in lieu of a wake and

funeral, before neatly disposing of oneself. Certainly, this *could* be universalized.

Hume's views, however, remain as valid today as they were 262 years ago and are significantly less radical. In a secular society, there is no argument to be made that suicide violates God's will or the natural order of things. Although one may still believe such things, this is only evidence that philosophy has not yet been the "sovereign antidote" to such "superstition and false religion" that Hume hoped it would be. However, more importantly, Hume's views are prescient in a society in which life can be extended nearly indefinitely beyond its individual quality or social usefulness.

In considering these two philosophers' views on suicide, what would happen to our man Joe Citizen if we inserted him into a Kantian Kingdom of Ends in which Humean views on suicide were the standard? Kant (1797) put forth a thought experiment in his *Groundwork of the Metaphysics of Morals*, positing a society in which all residents were rational beings who were treated as ends unto themselves and *not* simply as means to others' ends. When considering Joe Citizen's life then, we would see him as a rational being, perfectly capable of deciding when and if it were appropriate to end his own life. Joe is an end unto himself, he is unbeholden to the superstitions and fears surrounding suicide, and he is no longer chained by society's *a priori* views that suicide ought to be prevented in all cases. Joe can simply contemplate whether he no longer wishes to be alive. Further, this view would be universalized to every citizen. Let us take a stroll through Joe's neighborhood in this Kingdom of Ends to see how death is managed.

Before we can begin to see how Joe considers his own death, we must take a look at death itself and consider some realities in today's society that were not present when Hume and Kant existed. First, as it stands, death is still an inevitable outcome for us all. (It may be the case that science advances to a point

where this may not be true; the social and moral implications of this possibility are beyond the scope of this article.) In a modern, stable society, there are three broad ways in which one can die: suddenly and unexpectedly; with some, but not much notice (6 months to 1 year); and, lastly, a prolonged decline. It is difficult to imagine a society in which the first possibility is ever completely eliminated; regardless of advances in trauma care, sudden death may always lurk. A terminal diagnosis, however, gives one the opportunity to scratch items off of the proverbial "bucket list," resolve lingering issues in relationships, come to terms with death, and put one's affairs in order. Lastly, a prolonged decline at the end of life leads to an individual no longer experiencing the joys of existing, becoming a burden on loved ones, and utilizing more resources to prolong their life after their existence has ceased to bring them or anyone around them joy.

When considering the three possible ways in which our man Joe could meet his end, the second option seems most preferable. Certainly, a quick snuffing out rightfully terrifies most of us. Further, it keeps us from fulfilling our obligations to each other. I have committed to finishing this article; however, if a safe falls on me before completion, I will fail to fulfill my obligations, not to mention my more significant promises to loved ones. Presumably, our man Joe has similar obligations. However, Joe's being given notice of his death allows him to settle his obligations, make peace with his loved ones, and likely pass from this world to the next in a more peaceable fashion than the aforementioned squashing-by-safe. The final option seems the most miserable for all involved, a terrible prolonging of the inevitable that occurs every day in hospitals and nursing homes, where souls who have long outlived their joy in life or usefulness to society linger in suffering.

If death is absolutely inevitable, why would our man Joe, living in a rational Kantian Kingdom of Ends, not be expected to manage his death as he sees fit? We certainly seem to agree that if Joe is an autonomous human being, he has the right to choose

where he wishes to live, whether he wants to have children, what he chooses to study, who his significant other shall be, and what his vocation will be. Why would Joe's death, which is ***one of the only inevitable things that he as a human is bound to experience***, be left to fate when we allow for Joe to choose all other aspects of his life? If we allow for the fact that Joe certainly has a right to choose and manage his death in the same rational manner as he chooses to manage the rest of his life, suicide then becomes the moral imperative for Joe, both in his duty to himself and his duty to others.

Let us return to our Kingdom of Ends in which Joe now finds himself, instead of the rather unpleasant world that is today in which Joe's suicide has a covert and subterranean nature because of the institutionalized views of suicide prevention. To be clear, Joe may still get squashed by a falling safe in the Kingdom of Ends. This is likely tragic for Joe as, in his last moments, he wonders how that novel he was reading ended, he worries who will take care of his dog, he regrets that he never finished his will, and he agonizes that he may not have resolved a matter with a loved one. Further, those that care for Joe will be traumatized by his sudden departure and burdened by all of the matters Joe left unfinished.

This, however, is the scenario that our present society forces Joe and his loved ones to endure because of the non-negotiable nature of suicide prevention. Although a calm and rational Joe who chooses to end his life may be able to address some of the aforementioned issues, it is likely the case that his departure will have to come as a surprise to those that care about him, and this alone is traumatic. However, we seem to fail to recognize that Joe's *not* managing his own death is as traumatic, if not more so. Why ought Joe waste away in a nursing home, allowing his faculties to slowly decline, forcing his loved ones to watch his demise, while he consumes valuable resources? It seems Joe has a moral obligation to himself and his loved ones to prevent this equally miserable death.

In a Kingdom of Ends in which Joe's managing his death is not only acceptable but expected, *none* of the horrible aspects of sudden death occur and *all* of the misery of a prolonged death can be avoided. Our man Joe can simply say, "Hey Bob, I've been thinking that it's getting to be about time for me to die," and Joe can have a rational, calm conversation with those in his world about this matter without anyone pouncing on a phone to call 911. Just as with marriage, moving, or having a child, those who love Joe can still advocate for him to not commit suicide at this time and can explain the reasons why, but an acceptance of suicidal ideology removes the psychopathology, moralizing, and social stigma from the conversation. Joe could say goodbye to his loved ones, ensure his affairs are in order, and be comfortably surrounded by the people of his choosing. His methodology would be effective and the disposal of his remains pre-planned. Joe could even attend his own wake. This Kingdom of Ends is obviously both morally and practically superior to the current world in which suicide is always a thing to be viewed from a lens of pathology and to be prevented at all times.

III. A TIMELINE OF SUICIDE

To critics of this view who argue that suicide can still be a result of mental illness or an immoral act, we must consider a timeline of Joe's life to analyze whether their argument is valid. If we consider Joe's life, with T_0 being Joe's life up until he contemplates suicide, T_1 the time at which Joe considers ending his own life, T_2 the time of Joe's actual suicide, and T_3 the events in the wake of Joe's suicide, we can still certainly find moments where their arguments may be valid, *but it is not at T_1 or T_2;* their arguments rest solely on what occurs at T0 and T3.

At T_0, it may be the case that Joe is mentally ill and this mental illness can be treated. It may also be the case that Joe has simply done all of the things he has wanted to do in the world and

sees no reason to stay any longer. It could also be the case that Joe has had a miserable existence with little chance of improvement and he wishes to end his misery. Even if Joe is mentally ill, this fact does not seem like indisputable grounds for suicide prevention and could possibly constitute an argument *in favor* of it. If Joe is depressed, schizophrenic, addicted, or a victim of significant trauma, the simple argument that he *may* get better is no reason to believe that he ought to suffer through a lengthy recovery process just to see if he will. It *may* be the case that the view from the top of Mount Everest is the most spectacular thing I could ever experience in my life—that does not mean I have any obligation to endure the financial cost, physical suffering, and emotional hardship required to get there. I will take a pass. Simply because it is possible for Joe to recover from his mental illness does not mean that he ought to try, nor does it mean that his suicide ought to be prevented for the sole reason that he may be mentally ill.

When we consider how Joe could be committing an immoral act by ending his life, any immorality that occurs is not due to actions at T_1 or T_2, but at T_3, where Joe fails to meet his obligations to others or actually inflicts harm. If Joe kills himself in the fashion that one must today, Joe *has* to be found unexpectedly and likely by someone who cares about him. Seeing a loved one with a gunshot wound to the head when one comes home from work inflicts a moral injury that could be avoided if Joe had had the ability to manage his death in a reasonable way. Leaving others with unresolved issues that they could have discussed with Joe, but did not have the opportunity to, leaves a moral injury that could have been avoided if Joe could depart openly.

If Joe kills himself without Bob's being able to say, "Hey man, you said you'd bring the potato salad to my Super Bowl party. What gives?" Joe has done something wrong, but not in his suicide: Joe's moral violation occurs when said Super Bowl party happens without the all-essential potato salad. Joe's unethical action is in breaking his word about potato salad, not in choosing

his death. However, Joe can remove his moral obligations if he can openly discuss his suicide. Joe can attend to his obligations and not take on any future ones if he properly schedules his demise; as it stands today, it is difficult to give a two-week notice at work in order to kill oneself. However, again we see how the institutionalization of suicide prevention forces Joe into moral harms that he could have otherwise avoided if he could have openly discussed his end.

IV. Joe Goes to a Party

To see where there is moral harm in suicide, we must examine the current role of institutionalized suicide prevention. To do this, we shall drop our man Joe in a grand and lavish party in a castle to analogize the life that Joe experiences. Joe did not ask to be invited; in fact, with no will of his own at that point, two people dragged him to that party. The party is vast, with opulent banquets on the top floors, generally restricted to those with special access, while much of the castle is occupied by a working class, occasionally partaking in amusements, but rarely allowed in the top levels. There are also those living in locked sub-basements, with little to do but slave away in misery while listening to the faint celebrations of the glitterati from high above. Some people in the castle find friends, loved ones, and invite others to the party to celebrate, regardless of their access, while others drag more people into their stations to find company in their misery. Some people can move about freely, others stay restricted to only one room, regardless of where they enter. The party represents life in society in general, with people engaging in different activities in different places and experiencing vastly different realities from one another. Although the general view is that no one is allowed to leave the party of their own accord, there are those experiencing the event in such a way as to wonder why anyone would ever wish to leave, while others wonder when their misery will end.

So, what of our man Joe? He did not ask to be invited to the party and so he ought to be allowed to leave when he sees fit. However, if he in turn asks others to join him, he does have some moral obligation to stay and show them around. Certainly, we would agree that inviting someone to a party where they know no one and then slipping out the door when they arrive would be wrong. In essence, if Joe has children, it would be wrong to abandon them in an act of suicide before fulfilling his obligations to them. If Joe makes friends, is having a nice time, and says he is going to run to the kitchen for a bucket of beer for everyone, then simply leaves and never returns, we would say it was wrong for him to break his word to his friends. We could even argue that if Joe was in a casual conversation with a girl who he knew loved him, slipping out the back door without so much as a goodbye would be wrong for the hurt it would cause her. We can see ways in which Joe ought not to leave the party and we can see ways in which Joe at least ought not to leave the party unannounced. We can also recognize that some of these moral obligations are worthy of debate. However, we would never say that Joe has an obligation to stay at the party, no matter what, no matter where he was locked in the castle, no matter how awful the people who dragged him there were, no matter how mean the crowd is, or how miserable a time he is having. That would be absurd.

However, this is exactly what happens. People who occupy higher levels of the castle have made it their business to bar Joe from leaving in all cases. People who appear to be having a fairly pleasant time, enjoying the banquets, socializing with friends, passing through the marvelous and interesting rooms, have called down to those who are miserable, "Why don't you just come up here? The view is lovely and you can eat cake! What's wrong with you that you would want to leave?!" all the while telling everyone at the party to bar the doors and that the ones who try to escape are crazy or morally bankrupt. **It is those people who are crazy and/ or morally bankrupt.** They are the ones whom we should call out

for their wrongness. If Bernice on the top floor with her face full of cake can not understand why Joe may want to leave, she ought to get herself down to the dungeon, unlock the doors, and share her pastries, not condemn our man Joe for wanting to leave her shitty party when it is not fun for him. If we have a moral obligation to prevent suicide, it is only insomuch as we all have a moral obligation to make the party a fun time for everyone. Where we all started as guests, the longer we stay, we become de facto hosts. We are obliged to make the party more pleasant for our fellow humans if we wish for them to stay, but we have no right to bar their exit, even if it means we are stuck without potato salad.

CONCLUSION

In applying a Humean approach to suicidal ideology while crafting a Kingdom of Ends in which all rational persons who exist will eventually have to confront their own demise, we see very quickly that the morally just thing to do is to manage our own deaths in a way that fulfills our obligations to ourselves and others. There is no view that stands in the face of this that holds that the act of suicide itself or the contemplation of one's own suicide is either immoral or simply a product of mental illness. There are ways in which one can be mentally ill and suicide can remain a rational choice; there are also ways in which we can agree that the abdication of obligations in the wake of a suicide may be harmful—but that is not the suicide itself. However, the moralizing around suicide and the social institutionalization that all suicide is harmful and ought to be prevented is abjectly wrong. Anyone promoting this view has a moral obligation to improve the conditions in which a suicidal person lives, not try to prevent them from stopping their suffering. The view that suicide ought to be prevented is itself a moral harm and those promoting it cause more harm than they prevent.

Bibliography

Black, Jeremy. (1999) *History of Europe: Eighteenth Century Europe*, 2nd ed. (London: Palgrave MacMillan)

Dobis, Elizabeth, Stephens, Heather, Skidmore, Mark, and Goetz, Stephan. (2020) "Explaining the spatial variation in American life expectancy," *Social Science & Medicine*, 246(112759)

Hume, David. (1757) "Of Suicide"

Kant, Emmanuel. (1797) *Fundamental Principles of the Metaphysic of Morals*

Locke, John. (1690) *Second Treatise of Government*

"Suicide" (n.d.). *American Psychological Association*. https://www.apa.org/topics/suicide/

"Suicide and Suicidal Thoughts", *Mayo Clinic* (n.d.) https://www.mayoclinic.org/diseases-conditions/suicide/symptoms-causes/syc-20378048

Smith, David. (2011) "And They Leapt into the Unknown; Their last moments were watched around the world. But the subject of the 9/11 "jumpers" has proved so upsetting and unsettling that, a decade on, few in America are willing to mention them. Why have they been sidelined from history?," *Sunday Times* (London, England), p. 40

POPULISM AND RECOGNITION

Meelaud Michael Pourmehr

INTRODUCTION

While populism is by no means a new concept, it has undoubtedly become a buzzword among much of contemporary political theory—if not for its excessive usage in political punditry, then for its capacity, or incapacity, to account for the recent wave of right-wing reactionary political thought that has taken hold of the West. Theories on populism confront both the practitioners of populism as well as its adherents, though terminological debates have been a consistent element of the discourse since the first popular theories of populism arose in the 1960s. While the recent literature has, more or less, found some common ground on the definition of populism—often referencing, whether directly or indirectly, the work of the political theorist Ernesto Laclau—there remains substantive debate on the roots of voter behavior as it relates to those who vote for populist parties.

There are three major perspectives that attempt to provide a causal root for populism: (1) the historical analysis of populist party behavior; (2) the socioeconomic analysis of the petty bourgeoisie; and (3) the social-psychological analysis of populist voters—the last two being somewhat intertwined. Empirical analysis puts a good deal of strain on the first two perspectives, as will be discussed below, yet arguments for or against the third analysis often wither down to debates on the normative feasibility of reducing populist voting behavior to psychological tendencies of fear and resentment. Populism is often put forth as a reaction to left-wing identity politics—particularly, as a result of the rise of multiculturalism—yet there is reason to believe that populism just is identity politics *tout court*. That being said, it is a form of

identity politics that not only is reactionary, but that also has a call for recognition that operates quite differently. Populist recognition is an adulteration of the normal models of recognition, which is to say, not only does populism recognize an identity claim, but, prior to recognition, it also creates the identity at hand. Additionally, it seems to escape one of the key components of recognition, insofar as it tends to begin and end at moral issues, often leading to very little objective change in the lives of the voting base. This paper will focus on the notion that populism just is a form of identity politics—albeit one lacking objective resolution. The paper will go about by (1) delving into the literature on populism in order to provide a background for the discussion; (2) lay down the foundations on the philosophy of recognition; and (3) identify the qualities of populism that lend itself to being understood as its own distinct version of identity politics.

1

As trite as it is to note that attempts to pin down a definition of populism have been befuddling, it is nonetheless necessary to identify why exactly the terminological debate remains contentious. When academics first latched onto the notion of "populism" in the late 1960s—Ghita Ionescu and Ernest Gellner's edited volume from the 1967 London School of Economics conference on populism being the first major touchstone—it was in reference to what was seen as the rising areas of conflict, that is, peasantism, decolonization, and the developments of Maoism (Müller 2016, p. 7). Populism was, of course, not the source of the anxiety that it is today, given that is was viewed mostly as an issue for non-Western states; however, the fundamental tension between democracy and populism has remained the key concern over the past few decades.

Populism first emerged near the end of the nineteenth century in the United States with the rise of the Populist Party—or, People's Party—and in Russia with the Narodniki, which roughly

translates into "peopleism", or "populism" (Canovan 1981, pp. 5-6). The former was a radical political movement comprised primarily of farmers in the United States, and the latter was a cohort of middle-class intellectuals who idealized the seemingly bucolic life of rural peasants (Hofstadter and Walicki, 1969). The common divider here is the identification of the peasantry as the core cultural and economic pillar of the nation (Mudde 2002, p. 219). What differentiates these two movements from future populist programs is the explicit emphasis on the rural working class, whereas in later movements appeals are made to the ambiguous notion of "the people"—which, in effect, allowed for the mobilization of multi-class coalitions (Drake 2009, Ch. 6).

It is in this notion of "the people" where populism, as we know it today, first finds its bearings. Unlike appealing to a specific economic class, the reference to "the people" makes for what Laclau dons an "empty signifier" (Laclau 1977). What is being referred to by "the people" is not to be taken as an empirical statement regarding a certain population, but rather one that is fictional at its root. "The people" that populists appeal to should be understood as synonymous with the "real people," what Nixon famously referred to as the "silent majority." This essentially amounts to a *pars pro toto* argument that is at its core anti-pluralist (Müller 2016, p. 20). Populists insist that they, and they alone, speak for the "real people"—though this is always presented as an appeal to "the people." The "real people" here should not be confused with the traditional notion of the population-at-large, but rather, it is a symbolic representation of a true, authentic homogeneous people, whose beliefs the actor is capable of reading and describing, perhaps even before the people themselves have fully adumbrated their beliefs. Jan-Werner Müller puts forth a particularly noteworthy example of this phenomenon by noting a statement that Trump made at a campaign rally in May 2016, wherein Trump exclaimed that "the only important thing is the unification of the people—because the other people don't mean anything."[1]

In essence, populist actors claim to capture the *Volksgeist* of the people. To be clear, however, the *Volksgeist* should not be conflated with the Rousseauian *volonté générale*, which requires some level of actual participation by the polity. Instead, the populist creates a homogeneous people through what is essentially a top-down prescriptive action, though it is painted as descriptive by the populist. The "real people" that the populist actor speaks of, and, in effect, creates, necessarily excludes all who would be opposed to the populist actor. The people, or the population, for the populist actor, are not truly part of "the people"; only those who agree with the populist actor are the true citizenry. This essentially allows populists to disregard empirical outcomes of votes, or the will of the people as represented through elected representatives. Since the populist represents the "real people", all of these empirical facts can be painted as the workings of elites that do not truly represent the people; if the people were to truly speak they would always bring the populist actor into power, since he necessarily speaks for them. This potentiality, per Müller, is an inherent feature of modern representative democracy (Müller 2016). Accordingly, populism is the "permanent shadow of representative democracy" (Müller 2016, p. 101), whereby a politician can come about exclaiming that they speak for the "real people" and, as such, are the only true, legitimate representative. Bureaucrats, or the "elite", are derided as *others* who maintain a hegemonic control over the political workings of the nation-state behind closed doors, which makes for a perennial scapegoat for populist politicians even when they are in power. Furthermore, elements of clientelism and cronyism can be shrugged off by populist voters, who rationalize corrupt acts of populists as being ultimately for the people—as has been seen worldwide with Donald Trump in the United States, the Freedom Party of Jörg Haider, and the Italian Lega Nord.

Outside the act of voting, populist voters are fundamentally passive when it comes to actual political participation and

will-formation. Populist politicians often operate through what Nadia Urbinati confusingly calls "direct representation" (Urbinati 2015, p. 2). The concept refers to the manner in which populists overcome intermediary bodies—such as the media, or even legislative branches—by directly "speaking" to the public, giving the veneer that the politician is working through an imperative mandate (Urbinati 2015, p. 4). The most concrete example of this is the founder of the Italian Five-Star Movement, Beppe Grillo, who went from comedian to popular blogger, and eventually to the Italian Parliament. In true populist fashion, Grillo acts as the only representative of the movement—which he has exclaimed as being the voice of the people—going so far as to have weekly meetings with others he groomed in the Five Star Movement, many of whom were commentators on his blog (Urbinati 2015, p. 6). Though the blog is presented as a forum for the people, in truth it operates with only a small number of devoted bloggers—aside from Grillo, that is—who make up the vast majority of its content, yet that in no way hinders the blog, and the Five Star Movement, from being presented as the voice of the people (ibid.). At the end of the day, the notion of "direct representation" ultimately produces a passive people that is prohibited from any meaningful political will-formation, or even participation outside of voting. Any potentially unsavory outcomes can be pointed back towards the electorate, since they, according to the notion of populists as representing the populist will, are ultimately to blame (Müller 2016, p. 31). However, this is often not the case, since populists always have the scapegoat of corrupt elites, often portraying them as a hindrance to their ability to perfectly represent the popular will. As such, the conspiracism is intrinsically tied to populism.

Populist conspiracy operates much in the same way as participating on Grillo's blog, which is to say there is very little room for actual deliberation or meaningful progress. Often enough, there is a sense in which conspiracies merely need to be true enough, which for conspiracy theorists is just the notion that the event was not

impossible (Muirhead and Rosenblum 2019, p. 41). Participation ultimately boils down to repetition, and through repetition comes tribal assent to a narrative that has roots not from any meaningful empirical work, but from the mere fact that the same conspiracy is repeated ad nauseum (ibid.). It is political activity, but the activity remains frozen. Thus, when Obama is derided as not being an actual American, as has been repeated for over a decade now, all that is needed for validation on behalf of his detractors is a mere repetition of the claim through as many outlets as possible, rather than any meaningful deliberation. Populism thrives on this unreflective aspect of conspiracies. It relies on conspiracy for much of its discourse, and furthermore, populist leaders give credence to conspiracy by repeating the claims themselves. And, when there are those who attempt to rebuke the claims of conspiracy, they can be cast off as being part of the elite, which in effect only helps populists to further cement them as the *other*.

2

While theories of recognition truly gained steam around the 1990s, particularly because of their ability to illuminate many of the social movements of the time—such as those of gay, lesbian, ethnic or religious minorities, and disabled peoples, just to name a few—they find their first notable iterations in the works of Hegel and Fichte. Fichte's core claim was that our autonomy becomes present to us once we are challenged, or as Fichte puts it "called upon", by another individual (Neuhouser 2008, p. 43). Hegel, for the most part, follows this same train of thought; however, Hegel notes that the ensuing struggle that comes about as a result of this meeting creates an impasse because of its inability to provide mutual recognition. In Hegel's slave/master dialectic, the victor of the struggle, while asserting their status, inevitably finds that their status cannot be recognized by the slave as a result of the slave's status as a slave. Mutual recognition for Hegel, as is most

concretely developed in his *Philosophy of Right*, requires an institutionalized order (Williams 1997, pp. 59–68).

Theories of recognition put strict emphasis on the outcomes that come as a result of misrecognition, which is the lack of feedback from other individuals, or society as a whole. Without the requisite feedback, individuals will find difficulty in attempting to embrace not only themselves, but also the activities they engage in. As Charles Taylor puts it: "[n]onrecognition or misrecognition can inflict harm, can be a form of oppression, imprisoning someone in a false distorted, and reduced mode of being" (Taylor 1994, p. 25). This violation of the individual creates a motivation to resist and brings about an attempt for recognition. This attempt is central to Hegel's theory of the progress of spirit, but so far as modern political theory is concerned, it lies at the heart of what has been popularly called, and derided, as identity politics. However, the importance of recognition appears far before individuals, or groups, engage in the political sphere.

Axel Honneth notes that it is through recognition that we are able to understand the evaluative reasons of others, a capacity which is primary to cognition (Honneth 2005, p. 58). Empirical findings note that the cognitive development of an infant relies on an empathetic attachment to their caregiver, and only in doing this does the child gain access to the world (Tomasello 1999). Honneth notes that though this recognition initially allows the young to understand all others as autonomous individuals, there lies the risk that reified normative practices can potentially cause individuals to become forgetful of recognition, even going so far as to treat others as subhuman (Honneth 2005, pp. 59-60). By setting aside recognition, one also sets aside the normative claims of other individuals.

Demands for recognition attempt to bring specific attention to how particular aspects of individual identity are often neglected by the dominant normative values of their society. It is with this background that identity politics finds its footing. As Taylor

famously argued in "The Politics of Recognition," there is a real concern that laws which are designed within the framework of what are the dominant political norms can have adverse effects against those who do not necessarily fall within its auspices (Taylor 1994). The modern struggle for recognition is often situated within the institutional confines of the modern nation-state, and as such, it calls upon specific political programs that take account of the differences inherent to groups and the individuals that comprise them. As a caveat it should be noted that there is a tendency within identity politics to provide either essentialized or deconstructed notions of identity. However, Taylor disputes these in favor of a discursive mode of identity, wherein individuals put forth a deep toleration of other individuals and groups, such that they could come to find new, meaningful identifications by way of interactions with disparate groups (ibid.).

Returning to the point at hand, identity politics within an institutional framework calls for the codification of respect and esteem, and this can be anything from redistribution, such as affirmative action, or—to take a lesser known example, at least in the states—the Quebec sovereignty movement, which Taylor was intimately a part of. With all that said, identity politics has been derided from the right, which often slanders identity politics as Postmodern Neo-Marxism—an identification that rests on conflating two lines of leftist thought that are ultimately incompatible. It has also been derided from those on the left who critique identity politics for being an integral part of neoliberal capitalism, lacking any teeth to attack political economy, or for its inability to account for intersectionality. While these latter critiques are all worth inspecting, it is not the point of this paper to do so. Rather, this last section will take a look at how right-wing populism, which often derides identity politics as pretentious and filled with neologisms, squarely operates as a political movement that is focused on identity, often to the detriment of any meaningful utilitarian goals—that is, goal-directed orientations, such as improving

social or economic circumstances. Importantly, however, populist recognition only scratches the surface for providing recognition of particular ways of life; successful populists operate by creating a constituency of which only they can be the true representatives.

3

Ideally, voters are expected to operate on some sort of background knowledge, which contains a degree of knowledge and critical capacity, as well as a penchant for engaging in public discussion with the hope that they may arrive discursively at some sort of standard as to what can serve as right and just political action (Habermas 1962, p. 212). The notion that this is almost never the case for the majority of the electorate is surely a point that has been repeated endlessly. At the same time, the belief that there is a horizontal stream of political opinion is largely fictitious. As is most often the case, the actual function of political will-formation, at least for the majority of voters, is much more vertical in its direction. Where actual political discussion occurs, it more or less functions as a mutual confirmation of ideas with little room for political will-formation outside of the bounds of already pre-disclosed preferences. This potential pitfall of this cohort is its saving grace as it manages to avoid the pernicious effects of a manufactured public sphere of discourse, which, to be fair, are most often the machinations of the higher strata themselves (ibid., 214). The Post-World War II political era functions roughly within this guise. The electorate is presented with a manufactured public sphere with a limited horizon of discourse, wherein lively debate ensues among punditry and candidates, yet it ultimately operates at the ocular level for the vast majority of the electorate (ibid., pp. 215-217). This era of representative democracy is referred to as "undemocratic liberalism" by Cas Mudde. It has become particularly prevalent since the first iteration of neoliberalism, what Will Davies calls "combative neoliberalism," which was the period

of time spanning from 1979 to 1989 wherein neoliberalism was presented as a combative force against enemies of liberal democracy (Davies 2016, p. 122). With the fall of the USSR and the arrival of normative neoliberalism—Fukuyama being perhaps its most notable theoretician—this undemocratic liberalism laid down the foundations for the populist reaction, which Mudde calls an "illiberal democratic response to undemocratic liberalism" (Mudde 2007, p. 4). This illiberal democratic response is the anti-pluralism of populism, which hopes to create a holistic conception of the state as comprised of a homogeneous, authentic people. And though populism is often presented squarely in opposition to neoliberalism, it tends to embrace and even exaggerate many of its worst tendencies, the key exception being neoliberalism's acceptance of multiculturalism. Populism fits quite neatly within modern representative democracy because it ultimately relies on the veneer of democracy in order to avoid becoming full blown authoritarianism, even if populist politicians act in much the same manner as authoritarians. The discussion that this section began with brings attention to how representative democracy works dynamically in composing the political identity of a large part of its electorate. Populism, as the shadow of representative government, does not shy away from this aspect and, if anything, fully embraces it, not only in its symbolic representation of "the people", but even through its statecraft.

Right-wing populist parties, particularly in Europe, are frequently predicated on xenophobic grounds, most often in the form of Islamophobia, such that it becomes one of the drivers of policy and voter agitation. This, in effect, is an attempt on behalf of populist politicians to create the homogeneous people they claim to speak for—this is, in fact, the stated goal of Viktor Orbán and his Fidesz Party, which has held a super-majority in Hungary for almost a decade.

The particularly virulent disdain that populists maintain for immigrants is predicated on the notion of Western identity

as morally superior to that of minorities, particularly when they are of different religious faith. Recognition, at least for Taylor, means putting forth attitudinal respect, insofar as our recognition of the other provides a degree of affirmation that what others articulate to be good is viewed as meaningful (Taylor 1989, pp. 92-95). Populism outright rejects the recognition of the other, often portraying the other as subhuman, which Honneth notes to be the "forgetting" of recognition (Honneth 2005, pp. 59-60). The good is understood as proper to the identity of the "real people", and it is often an overtly moral claim. Some of the more piercing analyses of populism are right to acknowledge that populism is constructed upon a moralistic imagination of politics: a battle between the good—"the people", which, as mentioned above, is ultimately fictional—and the bad—"the other", which is often an amalgamation of disparate ideas that are repeated *ad nauseum*.

Populist identity, at least so far as right-wing populism is concerned, is always a reactionary assertion. Marx famously wrote, after watching the conservative reaction to the revolutions of 1848, that in epochs of revolutionary crisis we "anxiously conjure up the spirit of the past" as a way of comforting ourselves as we face the unknown. Much talk has been made of how populist leaders look to the past in order to excite the masses, even going so far as to make the notion into campaign slogans. Of course, these calls to the past are often fuzzy at best. Where the intersubjectivity of memory allows us to make room for consensus on past events in the face of the fallibility of memory, it is just as capable of instigating false memories, wherein a nostalgia is produced that glazes over all concrete matters in order to appeal to a time preferable to now—that this would be particularly powerful as a tool to instigate a white identity movement should come as no surprise. At the same time, it operates as a nostalgia for an imagined future—both in the sense of what was lost during the neoliberal turn, as well as an appeal to that which was promised, yet now seems unattainable to younger generations. Populism cannot help

but view any movement in the cultural milieu, that is multiculturalism, as being disastrous to the cultural hegemony and the past. Writers like Éric Zemmour laments the recent cultural changes of his native France. In his hyperbolically titled *Le Suicide français* he runs the gamut by attacking everything from feminism, halal food, no-fault divorce, bans on smoking in restaurants, illegal immigrants, and just about everything else you would expect from a man who wishes to make the far-right into the only right-wing political party in France and openly longs for the cultural unity of his youth (Zerofsky, 2019).

Populist politics in the United States were something of an oddity for several decades, at least until the Tea Party Movement thrust itself into the limelight. The group itself lacked any cohesive party structure or leadership, and much of its early funding and promotion was provided by the right-wing think tank, Americans for Prosperity, that eventually lost oversight of the movement after a few weeks, recognizing that the situation had grown far out of the control of the organization. The Tea Party laid down the foundations for what would become the Trump presidential bid, and most all of the talking points that took center stage during the 2016 election were first explicated by the Tea Party. In the wake of the mortgage crisis and the subsequent bailout by the Obama administration, the Tea Party was presented as a movement calling for lower taxes, as well as for a reduction in government spending; however, that shrouds much of the actual campaigning of the group. The rhetoric of the Tea Party was not so much about the bailout but, in true neoconservative fashion, a derision of those who had taken subprime loans and unintentionally helped bring about the financial crisis. The Tea Party not only explicitly pointed to racial minorities that had taken these loans, but in a departure from the bailout, it began its cultural crusade by first alleging that women were guilty of wrongfully accusing men of rape in order to avoid the consequences for sleeping with loathsome partners (Marley, 2012). As Adam Kotsko notes, though the rape apologist

candidates lost their elections, this was perhaps the first indication that we would one day see a president that jovially recalled his own attempts at sexual assault (Kotsko 2018, p. 105). Libertarian bromides eventually took a backseat as the Tea Party firmly became a part of the culture wars, particularly when its adherents lost their collective minds at the notion of transgender individuals using the respective bathroom of the gender they identify as.

However, even when it was concerned with fiscal policy, the Tea Party was somewhat paradoxical, insofar as empirical studies found that across the board Tea Party members endorsed Social Security while firmly opposing other Social Welfare policies. Social Security was, however, initially helped to create the notion of "white citizenship" (Olson 2004, xix). White citizenship is not to be understood as an overt claim of superiority, but the result of New Deal policies that inherently advantaged white individuals by way of social programs that specifically targeted whites—this stems not only from Social Security's initial design that excluded agricultural workers and stay-at-home mothers, but also in the designs of highways and the redlining of suburbs (Disch 2012, p. 140). However, the New Deal was presented as both neutral and universal, and social security was put forth as a reward to wage labors that gained their independence through their work—note the contradistinction to what was until this point the dominant view of wage labor, that it was wage slavery, wherein the worker was dependent upon their employer (ibid., p. 139). This view of social security as being earned by independent individuals carried on into the Tea Party, who had at this point explicitly stated their resentment for those who depended on the government for help, the so-called "welfare queens". The neat rhetorical trick of the Roosevelt administration helped to blind the most advantaged from the fact that most social programs are actually substantially geared towards them (Mettler 2010, p. 809). As is the case, empirical findings show that whites disapprove of social programs only when they come to benefit minorities, but not when they benefit

whites (Disch 2012, p. 143). As Disch argues, this disavowal of social programs that benefit minorities is a key aspect of the Tea Party, not a quirk (ibid.).

The denigration of minorities as dependents has been a key constituent of populist rhetoric, both aboard and in the United States, whether it is comprised of the view that immigrants at the southern border are going to come and take jobs from hard-working independent Americans—that is, white Americans—or of when right-wing nationalist parties in Europe disparage immigrants fleeing war torn countries in the Middle East as seeking a free-ride into Europe. That the plight of the Middle East is largely predicated on the interference by Western states seems to be lost on most commentators, or worse yet, overtly brushed aside. Populist rhetoric feeds upon the other insofar as it needs some sort of object that it can point to as a causal mechanism for the failings of their own nation-state. The convenience of immigrants lies in the fact that they, like "the establishment", are an invisible other that need not be a tangible object in front of populists. Populism only requires a perceived injustice. If it is convincing enough, then it can be repeated. Concrete resolution, however, is not quite the modus operandi of right-wing populism. As mentioned earlier, populism operates by scapegoating. Corrupt elites are what are driving the country to ruin, even when the populist party is in power. One would expect that when a populist party gains power that the voters would view themselves as (1) being represented, and (2) receiving recognition, in the manner elucidated in the previous section. To the first point, populist voters take representation in a rather different direction than what is to be expected of in a liberal democracy. The populist actor operates much like the Leviathan of Hobbes, such that the populist actor comes to act as the body politic of the people, who, once again, are the 'real people'. This produces a passive voter base, not in the sense that they are placated to the point of no longer railing against corrupt elites, but rather, insofar as actual substantive utilitarian

change is concerned, there seems to be very little movement—what does get promoted is often vacuous vanity projects that yield negligible results, such as extensive border walls. Often enough, this result comes about from one of the key rhetorical moves of populists: the notion of the populist running the nation like a business. Sometimes this is announced outright, such as when Berlusconi would refer to his management of Italy by proclaiming the nation as *azienda Italia* (Müller 2016, p. 31). That one of Trump's major selling points for his voters was his 'success' as a businessman should not be lost here. This does seem to be one of the major points of distinction between left-wing and right-wing populism, so far as recognition is concerned at least. Left-wing demands often coalesce around collective attitudes that are goal-oriented—these goals often concern social and economic realities. Unfortunately, this does allow for the lamentable result of potential buy-outs, in the sense that delivering on certain goals can mollify the left—avoiding this particular pitfall was one of the particular goals of the Marxist conception of permanent revolution. However, right-wing populism, which operates almost entirely on moral ground, forgoes objective goals because of the very fact that regardless of who is in power, populism manages to provoke its adherents through constant agitation, often by way of unfalsifiable conspiracy.

Recognition, so far as it concerns populism, is rather tricky because of the fact that populism often engages those who are part of the cultural hegemony of any particular nation. If one were to view populism in the United States as a type of white identity movement, there is an obvious difficulty as whiteness is suffused into the nation's culture. One may ponder that since it is so hegemonic it may be imperceptible. This leads to calls for recognition only once other cultures begin to make their way into the mainstream—for example, think of the outcry against minorities, such as Muslims or LGBT communities, being introduced into popular culture. Pontification as to why populist voters feel as if they are

not being recognized aside—and obviously, we can be sympathetic to the loss of once well-paying working class jobs in rural America—how the actual phenomenon of the calls for recognition manifests itself is what is most notable at the current juncture.

Perhaps one of the most interesting phenomena of the Trump presidency has been the seeming deluge of rallies that the Trump campaign organizes. Why rallies have become so magnetic at this period of time is not so difficult to imagine. Given the undemocratic nature of modern liberalism, rallies, or political gatherings in general, seem to be one of the only venues whereby one can recognize themselves as part of a political struggle. Recognition of solidarity becomes one of the only ways in which respect that has been lost can be regained, and rallies do exactly that: they are a show of solidarity (Honneth 1995, p. 163). Whether or not the outcomes of these rallies in any way objectively change the lives of the participants is ultimately only represented by having their candidate winning. That is not to say, however, that there are no substantive reasons for why rally participants may feel that they, or their ways of life, have been disrespected, nor is it a claim that populist voters are inherently motivated by a sense of what is morally good or bad. Rather, right-wing populist rallies succeed because of the conditions of neoliberalism, which have, for the most part, diminished the pathways to actual democratic participation. The chasm that is left over is fertile ground for atavistic forces, such as those of xenophobia, racism, and the like. Identity then begins to revolve around the beck and call of populist leaders, who alone are capable of any change. That this change does not seem to appear is, as mentioned above, a non-issue. What remains is a shallow form of recognition, in which some self-respect is regained, but substantive objective change is foregone.

CONCLUSION

The most salient point of right-wing populism as a form of iden-

tity politics, one that revolves around a truncated form or recognition, is that it does not rely on any actual measurable change in the lives of its adherents outside of the basic fact that their candidate is in power. If the current pandemic is any indication, often enough voting for populist leaders, many of whom rely purely on rhetoric and charisma, leads to the detriment of their voters. Surely enough, the number of conspiracy theories that seek to ameliorate the befuddling actions of many populist leaders grows by the day. The right-wing version of permanent revolution comes in the form of permanent conspiracy, whereby objective change is disregarded in favor of a moral battleground wherein populists are seen as combatting an unseen evil. This alone is capable of maintaining power for populist leaders.

Populist voters, meanwhile, continue to engage in culture wars, the very objective of which is the misrecognition of the other. This, of course, becomes conflated with the prospect of a hidden corrupt elite, who not only hinder the goals of the populist leaders, but who also work in tandem with the maintenance of a mainstream culture that promotes the representation of minorities. Representation then, besides reinstating mono-culture, becomes the act of having your candidate in power, and this candidate is ultimately the driving force at the level of ideas. Even when key concepts come from the outside, by lending their voice populist leaders capture the Geist of their voters. What remains on the part of populist voters is a bland form of recognition, but only insofar as their will is being represented through their candidate, who, through the dynamic process of representation, comes to construct the will. The populist notion of "the people" loses all of its substantive capacity of recognition—that is, its capacity to promote distinct forms of life and bring about deep toleration. If anything, populism is fundamentally opposed to that goal. Populist recognition, being inherently anti-pluralist, is an identity movement that attempts to diminish outside claims to recognition by painting the other as a morally evil force. Yet, in doing so, it

only provides itself with a vacuous claim to power, one that ultimately does little in way of ameliorating the potential pitfalls of modernity.

Notes

1. See "CBS Weekend News," *Internet Archive*, May 7, 2016, https://archive.org/details/KPIX_20160508_003000_CBS_Weekend_News#start/540/end/600

Bibliography

Canovan, Margaret. (1981) *Populism* (New York: Harcourt Brace Jovanovich)

Disch, Lisa. (2012) "The Tea Party: A 'white citizenship' movement?," in: L. Rosenthal and Trost (Eds), *Steep: The Precipitous Rise of the Tea Party*, pp. 133-152 (Berkeley: University of California Press)

Drake, Paul. (2009) *Between Tyranny and Anarchy. A History of Democracy in Latin America*, 1800-2006 (Stanford: Stanford University Press)

Hofstadter, Richard. (1969) "North America," in: G. Ionescu and Gellner (Eds), *Populism. Its Meaning and Characteristics*, pp. 9-27 (London: Weidgenfeld and Nicolson)

Habermas, Jürgen. (1962) *The Structural Transformation of the Public Sphere* (Cambridge: MIT Press)

Honneth, Axel. (1992) *The Struggle for Recognition: The Moral Grammar of Social Conflicts* (Cambridge: MIT Press)

_____ (2005) *Reification: A New Look at an Old Idea* (New York: Oxford University Press)

Kotsko, Adam. (2018) *Neoliberalism's Demons: On the Political Theology of Late Capitalism* (Stanford: Stanford University Press)

Laclau, Ernesto. (1977) *Politics and Ideology in Marxist Theory* (London: NLB)

Olson, Joel. (2004) *The Abolition of White Democracy* (Minneapolis: University of Minnesota Press)

Marley, Patrick. "Rep. Roger Rivard criticized for 'Some girls rape easy' remark," (October 10, 2012). *Milwaukee Journal-Sentinel.* http://archive.jsonline.com/news/state-politics/state-legislator-criticized-for-comments-on-rape-hj76f4k-17358796i.html

Mettler, Suzanne. (2010) "Reconstituting the submerged state: The challenges of social policy reform in the Obama era," *Perspectives on Politics* 8(3), pp. 803–824

Mudde, Cas. (2002) *The Ideology of the Extreme Right* (Manchester: Manchester University Press)

_____ (2007) *Populist Radical Right Parties in Europe* (Cambridge: Cambridge University Press)

Muirhead, Russell and Rosenblum, Nancy. (2019) *A Lot of People Are Saying* (Princeton: Princeton University Press)

Müller, Jan-Werner. (2016) *What is Populism?* (Philadelphia: University of Pennsylvania Press)

Neuhouser, Frederik. (2000) *Foundations of Hegel's Social Theory: Actualizing Freedom* (Cambridge: Harvard University Press)

Taylor, Charles. (1994) *Multiculturalism: Examining The Politics of Recognition* (Princeton: Princeton University Press)

_____ (1989) *Sources of the Self: The Making of Modern Identity* (Cambridge: Harvard University Press)

Tomasello, Michael. (1999) *The Cultural Origins of Human Cognition* (Cambridge: Harvard University Press)

Urbinati, Nadia. (2015) "A revolt against intermediary bodies," *Constellations* 22(4), pp. 477-486

Walicki, Andrzej. (1969) "Russia," in: G. Ionescu and Gellner (Eds), *Populism. Its Meaning and Characteristics*, pp. 62-96 (London: Weidgenfeld and Nicolson)

Williams, Robert. R. (1997) *Hegel's Ethics of Recognition* (Berkeley: University of California Press)

THE PRIMACY AND LOGICAL NECESSITY OF ACTION IN HUMAN CARING

Peter Van

INTRODUCTION

Whence our caring? Whence *human* caring? Most broadly, this paper is concerned with precisely this question. But the purpose of this paper is in particular to clarify what human caring involves. So, what does human caring involve? Jaworska (2007) argues that human caring in both children and adults involves at least a particular kind of minimally complex cognition, while Seidman (2010) argues that human caring involves, at least in the case of rational adults, deliberative cognition in addition to minimally complex cognition. But both accounts do not emphasize what role *behavior* (action) plays in human caring. For this reason I will argue that caring in both children and adults involves not just cognition but behavior, but also that, more pointedly, any cognition involved (simple or complex) is behavior. In this paper, I give an argument for the logical necessity of action, drawing upon the notion of the primacy of action. Given that primitive behavior just is instinctive, thought-free, and reflex-like action or reaction, I argue here that primitive behavior is the *prototypical* form of behavior from which most if not all more complicated forms of behavior *stem*. Given that such primitive behavior is the *root* of most if not all more complicated forms of behavior, it becomes clear that the kinds of cognition which Jaworska and Seidman argue to be involved in human caring just are some examples of these more complicated forms of behavior. Broadly speaking, the insights to gain in this paper are that caring is something that we humans *do*, and that the *root* of such human caring is what we do—and have *always* already been doing.

129

In Section 1, I will consider first Jaworska's account of human caring in terms of what role minimally complex cognition (namely, attention and perception) plays. I will also point out, in this section, that an account in terms of cognition overlooks both the plain but nevertheless illuminating truism that we are ourselves animals and the action is primary. Next, in Section 2, I will consider Seidman's account of human caring in terms of what roles attention, perception, and practical deliberation play. What I want to show in these first two sections is that the failure to make explicit that human cognition is human behavior conceals human behavior itself. Then, in Section 3, I will motivate an understanding of ourselves that acknowledges the aforementioned truism, where this truism is precisely what will set up the context in which we want to situate our reception of the logical necessity of action. The utility of such an understanding consists in recovering the animal in us and restoring an appreciation of the primal origin of human behavior—including, in this case, human caring. Finally, in Section 4, I will give *the argument for* the logical necessity of action, elaborating upon the notion of the primacy of action. In order to show that both attention and deliberation are both logically necessarily based on pre-deliberative action and, by the very fact of this basis, *action*, I will argue that the roots of attention and deliberation are action (behavior).

1. COGNITION IN JAWORSKA'S ACCOUNT OF CARING

In her article "Caring and Internality," Jaworska argues that caring is an emotional attitude which is made of complex emotions (each of which in turn can be understood to be a mental state) and which supports various mental states (Jaworska 2007, pp. 560-561). Also, Jaworska conceives of caring as involving and requiring of the person who cares, *at the least*, minimally complex cognition (Jaworska 2007, pp. 559, 561). First, to point out that caring

involves cognition of a certain complexity, she remarks that "a caring response *presupposes* the agent's grasp of a cognitively complex picture of the relevant circumstances" (pp. 559, 561, my emphasis). But more than that, she conceives of caring as not only involving, but also requiring of the subject, at the least, an apparatus just cognitively sophisticated enough for the subject to *perceive* the other person as someone *important* to her. As Jaworska puts it, "a subject cognitively sophisticated enough to employ, at least implicitly, the concept of importance will *comprehend* the object's importance and this can inspire further cognitive activity" (Jaworska 2007, p. 561, original emphasis). Given the complexity of caring as an attitude, the cognition which it involves and requires is also complex, though *minimally* complex—minimally complex because this *minimal* requirement makes her account of caring compatible with "unsophisticated creatures such as *young children*" (Jaworska 2007, p. 544, my emphasis). Indeed, this account of caring is meant to account for the fact that such marginal creatures as young children are capable of caring, of genuinely caring.

Imagine, for instance, the following scenario—an example we will revisit below. A young child named Iris, upon hearing that her mother is crying in pain, *glances* over to see that her mother is in pain and her mother's pain-behavior. Seeing that her mother is in pain, she fetches a bandage for her mother, so that she might help alleviate the pain. Now, what Iris does here—namely, her turning around, glancing over, fetching a bandage—amounts to one such caring response. But what makes this caring response a *caring* response, if only with respect to the minimal requirement of minimally complex cognition? What cognition is involved here? That Iris responds at all to her mother's agony is telling. For one, she attends to and therefore perceives both her mother and her mother's pain-behavior, which is to say that at least attention and perception are involved. But more than that, her response, or her responding at all to the pain-behavior, also tells us that she

understands her mother to be *someone to respond to*—that there is something about her mother (namely, the agony she exhibits) that is *something to respond to.* Iris perceives her mother as someone to respond to and her mother's agony as something to respond to. That is, she grasps and is just cognitively sophisticated enough to grasp the fact that her mother is in pain—*i.e.*, the relevant circumstance. It is in this respect that Iris can be said to *grasp* the cognitively complex picture of the relevant circumstances; to employ, at least implicitly, the concept of importance; and to perceive (in the sense of comprehend) her mother as not just someone important, but someone important *to her*. When the child responds to her mother's behavioral expression of agony, then the child shows that she is capable of recognizing, if only in virtue of minimally complex cognition, the mother's expression of agony *as* something important to respond to (and the mother herself *as* someone important to respond to). And, in virtue of perceiving such an importance of her mother as the object of her concern, she can exhibit further cognition or more complicated kinds of cognition (such as, say, deliberative cognition).

Now, in contrast to young children like Iris, Jaworska points out that (most) animals lack the relevant cognitive capacities (such as the minimal apparatus considered thus far) for caring, or *genuine* caring, and that it is for this reason that *caring* distinguishes us persons from (most) animals (Jaworska 2007, p. 564). Since certain animals are, unlike Iris and other children like her, *not* cognitively sophisticated enough to employ, even implicitly, the concept of importance, these animals will *not* comprehend the object's importance; this failure to employ the concept will *not* inspire further cognitive activity in these animals (Jaworska 2007, p. 564). Granted, Jaworska makes this distinction with respect to the striking differences between us persons and most animals specifically in *caring*. Given that most animals lack the apparatus required for them to deploy, at least implicitly, the concept of importance and thereby perceive the object as important, these

animals can be said not to meet even the minimal requirement of sufficient cognitive sophistication in caring (Jaworska 2007, pp. 561, 564). As Jaworska makes her point nicely:

> The animal, like an infant (and in contrast to a two-year-old child), does not have, even implicitly, the concept of importance, and grasp of importance is necessary for a caring attitude to inspire the agent to organize (unify) his actions and intentions around that which matters. (Jaworska 2007, p. 564)

It is nevertheless my contention that this distinction between persons and animals throws shade on certain insights that we could gain by instead acknowledging what is *not* distinct about us, *i.e.*, what we have in common with animals. We make a mistake when we distinguish persons from animals: we first conceive of *human beings* as persons and then note that there is something which distinguishes us from animals. More pointedly, the reason this mistake throws shade on our relation to animals is that it makes it look *as if* we persons are *not* animals ourselves. That we *are* animals ourselves gets or risks getting overlooked, even if someone were to submit to it when this is brought to their attention. And insofar as the truism that we are ourselves animals serves or can serve to throw light on human caring, making the above distinction is all the more misleading. What does this truism illuminate for us in our considerations about human caring? I will address this question in Section 3. There, I will take the truism as a point of departure for motivating a more serious appreciation of it and clarify its role in coming to better understand human caring. I merely wanted to call the above distinction into question so as to bring out precisely what is at issue here. Now I would like to turn to the next of our considerations.

2. COGNITION IN SEIDMAN'S
ACCOUNT OF CARING

In the previous section, we considered a case in which the caring subject (a young child) engages in *minimally* sophisticated cognition—namely, some *basic* form of attention and some *basic* form of perception. Now I would like to turn to a case in which the caring subject (this time a rational adult) engages in rather *fairly* sophisticated cognition, which necessarily depends upon or builds on the minimally sophisticated cognition considered previously. In his paper "Caring and Incapacity," Seidman (2010) argues for an account of human caring that involves attention, perception, belief, and practical deliberation—the last of which is the fairly sophisticated cognition. Here I will only be responding to the involvement of attention, perception, and deliberation.

Seidman argues for an account of human caring in terms of cognition, which states, in sum and in terms of attention and perception:

> A person's concern with a given end, e, organizes [both] her [attentive and] perceptual efforts, so that she is on the lookout for e-related features of her environment—features of her environment which pose a threat or present an opportunity for her goal of realizing or sustaining e. (Seidman 2010, p. 311)

In order to argue this, Seidman runs his original argument (Seidman 2010, pp. 309-310) as follows, for this latter case which we are now considering. Consider Sarah. Since Sarah cares for the object of her concern (i.e., her son), she deliberately pays close attention to the object, the various ends focused around the object, and their surroundings so as to pay close attention also to anything which potentially endangers either the object or the welfare of said object (Seidman 2010, pp. 309-310, 314). That is to say, since Sarah cares about her *son*, she pays attention to her *son* and concerns

herself with her son's welfare, among certain other ends focused around him. But here's the relevant difference, compared with Iris: "These concerns [are said to] organize [both] her [attentive and] perceptual efforts, and put her *on the lookout*"(Seidman 2010, p. 309, original emphasis) for anything which might endanger either her son or his welfare, among other ends focused around him. Compared with Iris, who merely *responsively* engages in a caring reaction, Sarah here does more than that; she (Sarah) also *deliberately* engages in a caring reaction and deliberately searches her son's surroundings for potential danger. Moreover, compared with Iris (a young child), who engages in *minimally* complex forms of attention and perception, Sarah (a rational adult) engages in more complex forms of attention and perception.

But more than these more complex forms of attention and perception, Seidman argues that human caring in rational adults like Sarah also involves a form of deliberative cognition—namely, practical deliberation. Now, practical deliberation just is, according to Seidman, a cognitive capacity to think over which courses of action among those currently available to her, as they *appear* to her while she considers them, and to be able to, more importantly, "*exclude* from her deliberation courses of action incompatible with [ends focused around the object of her concern]" (Seidman 2010, p. 314). Let us again compare Sarah with Iris. Compared with Iris, who exhibits *no* deliberation in her minimally sufficiently sophisticated caring reaction, Sarah *does* exhibit such deliberation in her fairly more sophisticated caring reaction. Just as Sarah engages in more complex forms of attention and perception than does Iris, Sarah also engages in a *fairly* complex form of *deliberative* cognition. Both in the case of attention and perception and in the case of deliberative cognition, Sarah exhibits *a more complicated form of* a caring reaction than that which Iris exhibits.

Now, it is worth noting at once that what this translates to is this: A rational adult (and a rational animal at that) can engage in more complicated forms of *cognition* which are based on less

complicated or more basic forms of *cognition*. Which is to say, in terms of *cognition*, a rational adult (or, rather, a rational *animal*) can exhibit more complicated forms of caring reaction which are based on less complicated or more basic forms of caring reaction. But it is worth noting further that the point just made in terms of cognition can be said more rightly, as will later become clear, in terms of *action* or behavior. Deliberative *imagination* is, according to Seidman's account, also involved when Sarah thinks over (considers) the various possible courses of action available. In order not to exceed the scope of this paper, I will not respond to this involvement of a certain capacity to imagine. But let me now turn to some criticisms of Seidman's account thus considered. The following discussion will anticipate the argument for the logical necessity of action later.

Although Seidman speaks of deliberative cognition as the cognitive capacity to think about *what* to do, he does not point out that this deliberative capacity or disposition is itself a capacity or disposition to *do* something—namely, deliberate (think). Although Seidman speaks of deliberative cognition as a capacity to think over which possible course of action (as opposed to certain other possible courses of action) to take, he does not point out that such deliberation is *itself* already a course of action which someone takes in order to consider one's options (the various possible courses of action). I want to say: this disposition to *think* about what to do is itself a disposition to *do* something—namely, think about what to do; that is to say, thinking is already *doing*. Thinking is something that we human beings *do*. But more than that, in failing to acknowledge that thinking is itself doing and that, in involving thinking, caring involves *doing*, it remains difficult to regain clarity that "at the roots of thinking is … doing" (Hutto 2013, p. 284). Put another way, we are animals that, for the most part, are primarily actors and secondarily deliberators. But even when we are deliberators, we are, to put it more rightly, *logically necessarily* actors—this will become clear in the argument

for the logical necessity of action below. When we distinguish deliberation from action, if only in certain respects, it is important that we not forget that deliberation, although something over and above pre- or non-deliberative action, is itself after all *action*. But speaking of the *root* of human thinking, let us turn to the next of our considerations—the question of the *origin* of our human caring.

3. WHENCE OUR CARING? WHENCE HUMAN CARING?

Whence our caring? *Whence* human caring? I have hitherto suppressed this question, but it is to this question I would like now to turn. I now want to motivate an understanding of ourselves as animals and of human caring in the light of our being animals— the above-mentioned truism. Focusing on what distinguishes us *as persons* from other animals can keep us from recognizing and acknowledging the fact that we are ourselves animals. Whereas, importantly, focusing on what distinguishes us *as animals* from other animals lets us keep in mind that we are, humbly, animals ourselves. In giving either an explanation or a description of caring, it is important to bear in mind that the kind of caring we are giving an account of is *human* caring. So, it will be just as important to bear in mind that both the subjects and objects (if persons rather than things) of care are human beings, who, more-over, belong to the human species.

Now, it is my contention that *first* attending to the fact that we are animals will allow us to take this fact more seriously as such a fact, no matter how plain this truism may at first strike us. Focusing on what is *distinct* about our (human) caring leads us to lose or at least risk losing sight of what is *not* distinct about human caring: of our *continuity with* our primate (non-human animal) ancestors. (It is important to note the distinction between what is and is not unique about us human beings, because what is

distinctive about human beings builds on what is *not* distinctive about human beings—*i.e.*, what we share in common with our primate ancestors). If we focus all too much on the uniqueness of the human being, we shall mislead or at least risk misleading ourselves into thinking that we are discontinuous with our ancestors, when we are in fact of course *not* discontinuous with them. I do not wish to deny anyone any personhood, but I do, however, want to bracket (for at least the duration of this paper) any such personhood from the human being in order for us to recover the animal in us. The reason we want to recover the animal in us is to regain clarity of what we inherit from our ancestors in the light of both our genealogical ancestry and *our continuity with* that genealogical ancestry.

If in distinguishing ourselves from other animals we fail to acknowledge that we are *ourselves* animals, then we risk losing sight of *the origin of* the human expression of care and *the depth of* this origin. It is for this reason that I contend that acknowledging our ancestry can illuminate the ways in which human beings care—indeed, the ways in which *we* do, the ways in which *one* does.

To think really honestly about ourselves (and one another) in terms of *what* we are, we will have to conceive of ourselves (and one another) as animals, even if we are more than just animals—which we certainly are. I am not saying that we are *just* animals. Nor am I saying that we should treat ourselves and one another as animals, *as if* we were just animals. What I am saying is this: when we conceive of ourselves as being humans or human persons, then we risk forgetting to conceive of ourselves as being animals. An animal is what I am. When we distinguish ourselves from other animals, then we risk forgetting and sometimes even *effectively* forget that we are ourselves animals that descend from our animal ancestors—namely, our primate ancestors. More importantly, it is in this regard that I concur with Moyal-Sharrock that the most important insight that Wittgenstein offers would have to be this:

to approach an understanding of human beings, of *ourselves*, with respect to "the animal in us" (Moyal-Sharrock 2013, p. 263). As Wittgenstein himself remarks:

> I want to regard man here as an *animal*; as a primitive being to which one grants *instinct* but not ratiocination. *As a creature in a primitive state.* (Wittgenstein 1997, OC, §475, p. 62e, my emphasis)

For at least the rest of this paper, it will be important, as we consider human caring, to keep in mind 'the animal in us'—the animal that has, indeed, *always* already been there in us.

The manner in which we humans care about one another, I suspect, stems from the manner in which our primate ancestors "cared" about one another. That is to say, the way in which we, *strictly* speaking, care about each other derives from the way in which our primate ancestors, *loosely* speaking, "cared" about each other. And this point is compatible with the point that Jaworska wants to make (and rightly so), which is that there is something about human caring which distinguishes our kind of caring from "theirs" (Jaworska 2007, p. 564). That which is quintessential, *if* indeed quintessential, about human caring stems from both that which is *not so* quintessential about human caring and (again, I suspect) that which was characteristic of our primate *ancestors*. I want to say: human caring has an ancestral origin, and it would throw light on the way in which we care for one another if we approach our understanding of human caring *in the light of* such an ancestral origin. But although it is pivotal to acknowledge this primal origin, it is not within the scope here to illuminate the primitiveness of the "caring" which our primate *ancestors* exhibited. So, I would like now to turn to, at least, the primitiveness of the caring which *we* exhibit. For the rest of this paper then, the notion of primitive behavior shall be restricted to *us* human beings—that is, primitive behavior just refers to the "instinctive, thought-free, and reflex-like" actions and reactions which *we* exhibit (Moyal-

Sharrock 2013, p. 263).

4. THE PRIMACY AND LOGICAL NECESSITY OF ACTION

I will now argue that human caring is a mental state which involves at the least pre- or non-deliberative primitive behavior and then involves further, only in virtue of this, any other more complicated forms of behavior, including but not limited to deliberation, which stem therefrom. In order to argue this, let me start with the claim that caring involves, at the least, *primitive behavior*, before I argue for the claim that caring can involve further more complicated forms of behavior which emerge from and depend upon the former. Consider one of Wittgenstein's original philosophical remarks:

> It is a help here to remember that it is a primitive reaction to tend, to treat, the part that hurts when someone else is in pain; and not merely when oneself is—and so to pay attention to other people's pain-behaviour, as one does *not* pay attention to one's own pain behaviour. (Wittgenstein 2007, Z, §540, p. 95e, original emphasis)

Although Wittgenstein himself does not make this remark, nor any other remark, with respect to the phenomenon of human caring, I want to interpret it with respect to human caring.

Recall Iris, who is a young child. When Iris turned around so as to *glance* over at her mother who was in pain, then she engaged in *one such* primitive reaction. Now, we can of course imagine otherwise that what she did instead was merely uttered aloud in her concern for her mother, "Mom?" or "Mother?" And we can of course imagine still otherwise that what Iris did instead was pat her mother on her back or set her hand on her shoulder, say, so as to comfort her. In any case, these are examples of different kinds of primitive reactions, as Wittgenstein notes:

> The primitive reaction may have been a glance or a gesture, but it may also have been a word. (Wittgenstein 1958, *Philosophical Investigations*, p. 218e)

But this *glance*, I want to say, is not merely one such primitive reaction; it is a prototypical form of attention and thus a prototypical form of perception also—an instinctive mindedness, if you will. This instinctive glance is the prototypical form of attentive behavior from which most, if not all, the more complicated forms of attentive behavior stem. To say that this glance is the prototype (the *root*) of most, if not all, other more complicated forms of attentive behavior is to say that more complicated forms of attentive behavior (including Sarah's) *stem from* the instinctive glance (like Iris's). Now, given that the more complicated forms of attention *stem from* the instinctive glance, they are logically necessarily based on the instinctive glance. And if the more complicated forms of attention are logically necessarily based on the instinctive glance, then they are logically necessarily themselves (at bottom) instinctive glances. Although the complex forms of attention are something over and above the simple glance, it is a matter of logical necessity that the former are, if only at bedrock, the latter. When Iris turned around and glanced over to her mother *as* the object of her concern, this turning around and glancing over was an example of a *prototypical* form of attending to said object. In contrast, when Sarah was "*on the lookout*" (Seidman 2010, p. 309, original emphasis) for any potential harm to the object of her concern and the ends (e.g., welfare) focused around said object, this *deliberately* being on the lookout—in addition to an instinctively glancing around and about—was *a more complicated form of* attending to said object. A similar argument for the logical necessity of action can be made in the case of deliberation. It is to this argument that I now want to turn.

Drawing upon Wittgenstein's original work and his notion of the primacy of the deed (action) (Moyal-Sharrock 2013, pp.

262-266), I will mount an argument for the logical necessity of action. Primitive behavior just is (once again), for Wittgenstein, instinctive, thought-free, and reflex-like action or reaction. Moreover, such primitive behavior is the prototypical form of behavior from which most if not all other more complicated forms of behavior stem. To say that the more complicated forms of behavior *stem* from primitive behavior just is to say, as Hutto would put it, that a more complicated one "emerges from and depends upon" a less complicated and more primitive one (Hutto 2013, p. 284). For instance, in the event that someone (else) is in pain, the more complicated behavior of a given creature's deliberative reaction to the other's pain-behavior stems from the given creature's primitive behavior of a pre- or non-deliberative, i.e., an instinctive, reaction to the other's pain-behavior.

For most if not all ways of thinking, given that a way of thinking *stems from* a primitive reaction, in which case this primitive reaction is the *root* of the thinking, the way of thinking is logically necessarily based on a way of acting. If the way of thinking is logically necessarily based on a way of acting, then the way of thinking is *itself* logically necessarily a way of acting. What Wittgenstein is saying is that, as Moyal-Sharrock emphasizes: thinking is not just based on acting but *is*, by the very fact of its being based on acting, more pointedly, *acting* (Moyal-Sharrock 2013, p. 264; cf. Wittgenstein 1980, CV, p. 31e). Basically, it is a matter of logical necessity that doing things deliberately derives from (is based on) doing things rather instinctively, that is, doing things *without* thinking. It is a matter of logical necessity that doing things after giving one's options some thought derives from (is based on) doing things *without* giving one's options (I mean, *e.g.*, courses of action) some thought. This is to say that it is logically necessary that doing what one does premeditatedly derives from doing what one does *without* any such premeditation. The deliberative response is at bottom a behavioral response. And as Moyal-Sharrock succinctly sums up this notion of the logical

142

necessity of action, in her exegetical work on Wittgenstein, "[a]ction … is, for Wittgenstein, at the logical foundation of thought" (Moyal-Sharrock 2013, p. 264). Similarly, as Hutto puts this same point, in his exegetical work on Wittgenstein, "[a]t the roots of thinking … is doing" (Hutto 2013, p. 284). Indeed, given these two formulations, the one and the same point can be formulated thus: to say that the *root* of our thinking and thought is action is to say that the *logical foundation* of our thinking and thought is action.

Now, recall that Seidman understands human caring to involve a certain kind of thinking—namely, practical deliberation. Note, importantly, that practical deliberation does not exhaust the number of ways of thinking; it is *just one* of many. Since practical deliberation is a way of thinking, practical deliberation is *itself* logically necessarily a way of acting. So, when we speak of human caring as involving practical *deliberation* (a way of thinking), then we are also speaking of human caring as involving practical *behavior* (a way of acting or behaving). As Wittgenstein remarks on the primitive reaction:

> But what is the word 'primitive' meant to say here? Presumably that this sort of behaviour is pre-linguistic: that a language-game is based on it, that it is the prototype of a way of thinking and not the result of thought. (Wittgenstein 2007, Z, §541, p. 95e, original emphases)

Given that practical deliberation is but *one* such way of thinking among many different ways of thinking, the argument just given above, more generally, for various ways of thinking, applies to the particular case of practical deliberation. Contemplating what courses of action are available to choose from and which course of action to take is but one. But since ways of thinking are logically necessarily ways of acting, *contemplating* what courses of action are available to choose from is *itself* one such course of action. Granted, this contemplating here does not count toward

the number of possible courses of action contemplated upon. Both in general and in human caring, the practical deliberation about which course of action to take is itself action. Which is to say, deliberating about *what* to do to safeguard the object of care is *itself* already *doing* something for the object of care, even if deliberating about *how* to safeguard the object of care is not yet, strictly speaking, to *safeguard* it. Navigating the numerous possibilities in this manner, with respect to the object of one's concern, is *itself* logically necessarily a way of *behaving* in response to the object of care.

The deliberative response (or, rather, reaction) in Seidman's account (Seidman 2010, pp. 312-314) is an example of a *more complicated form* of the pre- or non-deliberative primitive reaction. To say that this deliberative response here is *not* a primitive reaction would be misleading, even if our intuition that the deliberative response is something over and above the instinctive one tells us otherwise. For to say, more charitably, that this deliberative response here is *not just* a primitive reaction (*not just* an instinctive or reflex-like response) is to say that this deliberative response here *is* in fact a primitive reaction (an instinct or a reflex), only it is *more than just* this primitive reaction (this instinct or reflex). And this can be said likewise in the context of human caring. For to say that the deliberative *caring* response here is *not just* an instinctive response (bearing in mind that this primal instinct is an *action* in response to an object of concern) is to say that this deliberative *caring* response here *is* in fact an instinctive response, only it is *more than just* this primal instinct. As Wittgenstein remarks:

> I really want to say that scruples in thinking begin with (have their roots in) instinct. Or again: a language-game does not have its origin in *consideration*. Consideration is part of a language-game. And that is why a concept is in its element within the language-game. (Wittgenstein 2007, Z, §391, p. 70e, original emphasis)

Regarding the complex mental or psychic life of a given subject (a given human *animal*), it is my contention that in order to gain insights into her mental life, her psychic life, we will have to attend at least in part to her behavior, her actions—what she *does*. The "sophisticated form of mentality," to borrow a phrase from Hutto (2013, p. 284), in human caring stems from the less sophisticated ones—including, at bedrock, an instinctive mindedness.

Conclusion

In this paper, I motivated an understanding of *ourselves* as animals (alongside fellow animals) and of human caring in the light of our being such animals. As I said above, the utility of such an understanding consists in recovering the animal in us (*i.e.*, the truism) and restoring an appreciation of the primal origin of human behavior both in general and in human caring in particular. After motivating such an understanding, I gave an argument for the logical necessity of action in both the cases of attention and deliberation. Once I mounted the argument, I showed how minimally sophisticated and fairly sophisticated kinds of cognition (namely, attention and deliberation, respectively) can be understood in light of it. It should be noted, however, that what we considered in this paper does not of course exhaust the number of considerations about human caring. We are far from fully understanding what (in more detail) human caring involves and from what our human caring derives. One might say: What is considered here in this paper is *itself* a root for further thought on human caring. But if the arguments laid out in this paper are right, then acknowledging the primacy and logical necessity of action forces us, as I contend, to approach an understanding of human caring in terms of action and in the light of this primacy, this necessity. It is with this that I leave the reader to ponder, for herself, this human phenomenon.

Bibliography

Hutto, Daniel D. (2013) "Enactivism, from a Wittgensteinian point of view," *American Philosophical Quarterly* 50(3), pp. 281-302

Jaworska, Agnieszka. (2007) "Caring and internality," *Philosophy and Phenomenological Research* 74(3), pp. 529-568

Moyal-Sharrock, Danièle. (2013) "Wittgenstein's razor: The cutting edge of enactivism," *American Philosophical Quarterly* 50(3), pp. 263-279

Seidman, Jeffrey. (2010) "Caring and incapacity," *Philosophical Studies* 147, pp. 301-322

Wittgenstein, Ludwig. (1958) *Philosophical Investigations* [PI], tr. G. E. M. Anscombe (3rd edition) (New Jersey: Prentice Hall)

_____ (1972) *On Certainty* [OC], G. E. M. Anscombe and von Wright (Eds), tr. Denis Paul and G. E. M. Anscombe (New York: Harper & Row)

_____ (1980) *Culture and Value* [CV], G. H. von Wright (Ed), tr. Peter Winch (Chicago: University of Chicago Press)

_____ (2007) *Zettel* [Z], G. E. M. Anscombe and von Wright (Eds), tr. G. E. M. Anscombe (40th anniversary, bilingual edition) (Berkeley and Los Angeles: University of California Press)

CONTRIBUTORS

Steven Diaz. B.A. Philosophy, University of California, Riverside. Philosophy of Mind, Philosophy of Psychology, Philosophy of Psychiatry, Philosophy of Cognitive Science, and the Philosophy of Education.

David Fonth. B.A. Philosophy, California State University, Los Angeles. Philosophy of Mind, Metaphysics, Philosophy of Language, Feminist Philosophy. After earning his M.A. in the spring, David will be pursuing a Ph.D. at Rutgers University.

John Hurley. MA in History, Harvard University, MS in Computer Science, California State University at Los Angeles. Ancient Philosophy, Ethics. John is a Lecturer in Computer Science at Cal State LA.

Jen Markewych. M.P.A. Public Administration, California State University, Los Angeles. Ethics, Social Philosophy, Public Philosophy. Jen plans to continue her work to bring philosophy to the field of crisis response while intending to apply to Ph.D. programs in the Fall of 2021.

Jalen McLeod. B.A. Philosophy, California State University, Los Angeles. Philosophy of Economics, Political Philosophy, and Philosophy of Education. Jalen plans to be a Congressman and then President of the United States.

Meelaud Michael Pourmehr. B.A. Political Science, California State University, Northridge. Ethics, Social and Political Philosophy, Philosophy of Law. Meelaud intends to apply to joint JD-Ph.D. programs in the Fall of 2020.

Sakib Ibne Shahriar. B.A. in Psychology, Minor in Philosophy, California State University, Los Angeles. History of Philosophy, Social and Political Criticism, Trans Philosophy, Metaphilosophy, Interfaith Theologies, Moral Philosophy, The Philosophies of Gilles Deleuze and Simone Weil. In the near future, Sakib wants to read more fiction, bake a wide variety of breads, practice writing in different, non-philosophical styles, and listen to more experimental dance music.

Peter Van. B.A. Psychology, University of California, Riverside. Philosophy of Language, Philosophy of Action, Philosophy of Psychology and Wittgenstein.

MASTER OF ARTS IN PHILOSOPHY
CALIFORNIA STATE UNIVERSITY, LOS ANGELES

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

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

PHILOSOPHY IN PRACTICE
SUBMISSION INFORMATION

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

For more information on *Philosophy in Practice*, please contact:
mshim@calstatela.edu