Philosophy in Practice

Volume 11 – Spring 2017

California State University, Los Angeles
Department of Philosophy

Philosophy in Practice

VOLUME 11 – SPRING 2017

Contents

Acknowledgments

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

California State University, Los Angeles
Philosophy Faculty

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

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

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

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

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

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

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

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

EMERITUS PROFESSORS

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

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

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

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

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

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

Professor Spotlight: Talia Mae Bettcher

Talia Mae Bettcher was born in Calgary, Alberta, Canada, and lived in various places throughout the country during her childhood. As an undergraduate, she attended York University, originally planning to pursue a B.Ed. Initially, her main interests of study were English and French, but around the same time she came across philosophy.

She signed up for a critical thinking course, which she then applied in studying the philosophy of religion. Despite excelling in this course, Talia found the subject matter personally discomforting. At the end of the course, her professor and advisor suggested she major in philosophy. Talia initially rejected the idea; however, in the years ahead, she found herself gradually gravitating towards philosophy. While Talia had originally planned on becoming a high school teacher, the discipline and rigor of theoretical philosophy lent new focus to her studies and she wound up majoring in philosophy and linguistics. After receiving her bachelor's degree, Talia decided to pursue her PhD in philosophy at UCLA.

Midway through her studies at UCLA, she approached her department about gender transitioning. At the time, there were no legal protections for trans people in California. However, the department proved very supportive.

During this time Talia had been doing work on the self and first person self-reference. Initially focusing on the work of Wittgenstein and Anscombe, she turned to investigate early modern conceptions of the self. She wrote her dissertation on Berkley and Hume, arguing that Berkeley had contributed a great deal more to how we should understand the self than had previously been thought. It was not until later on that she came to realize that her early work on the self was partly motivated by the desire to make sense of the complicated experience of transitioning.

Towards the end of her Ph.D. program, Talia had begun teaching some classes at Cal State LA, and was eventually offered a position. The department's plurality and strong feminist roots made it particularly appealing. While she began her work as a Berkeley scholar (publishing several articles and two books), it was not long until she turned her attention to the philosophy of gender.

Talia had been informally philosophizing about trans people and the obstacles and violence they confront since she was a graduate student. From even before her transition, she had been actively involved in Los Angeles trans activist communities and her experiences there provided the starting point for her reflections. However, it was the murder of Gwen Araujo in 2002 and L.A. trans community response to it that inspired her to begin publishing articles on trans issues. Since the initial publication of *Evil Deceivers and Make Believers* in 2007, which examined transphobic violence, Talia has continued to work in the area of trans philosophy. She is currently working on a new book about what transphobic violence can show about the concepts of intimacy and personhood. The more she has delved into the philosophical issues underpinning transphobia and violence, the more she has been returning to critique the starting points of those early modern conceptions of the self. While the work is highly theo-

retical, it is also grounded in real-life experience and motivated by the desire to make the world a better place for trans people. Unlike some professional philosophers, then, it is clear that Talia pursues philosophy as a deeply personal project.

— F.C. et al

Enabling Libertarian Free Will Through Human Enhancement

Taylor A. Dunn

Overture

It has recently been argued that whether or not we have free will is an empirical question to which there is not yet an answer. In particular, the claim is that neuroscience will eventually give us the final word on the matter. Assuming this is the case, there are several plausible answers neuroscience may give us about free will. It could be that our brains are structured exactly as needed for us to consistently act freely, or we may find out that our neurophysiology has nothing like what is needed for free will. However, the most plausible and probable answer neuroscience can give us is that our neurophysiology is close to enabling libertarian free will, but every action is at least partially determined in some way such that we cannot be said to possess free will. It is worth considering, given the likelihood of the close-but-not-quite answer, how we could manipulate our neurophysiology in order to *enable* libertarian free will. I believe that if given the opportunity to enable free will, there are no good reasons not to.

Whether or not the choice to enable free will seems straightforward, there are significant arguments made in the relevant literature for why we ought not adopt any human enhancement. However, it is not obvious if a free will-enabling biotechnology should be considered a full-blown human enhancement rather than a corrective therapy. It is generally agreed that our pre-philosophical intuition is that we do have free will—particularly, the phenomenology of choice is very robust—and were we to find that we in fact did not have free will, a procedure or drug which enables free will might seem like it is simply adjusting reality to align with our experience, rather than granting us superhuman

capabilities. But, based on the distinctions made between the concepts 'therapy' and 'enhancement', I will argue that a free will-enabling biotechnology is a human enhancement. As a result, free will-enabling biotechnologies are open to the general objections to human enhancement. And, there are further concerns specifically regarding enabling free will via enhancement. Undergoing such an enhancement may negatively impact the phenomenology of choice too drastically such that it would impede decision-making, or produce widespread social and ethical problems regarding moral responsibility. The difficulty, therefore, is to contend with a list of general objections regarding human enhancement as well as the particular problems of a free will-enabling enhancement.

In Section I, I will review Mark Balaguer's thesis on free will and what, as a result, would be required of our neurophysiology for us to have free will. In Section II, I will argue for the view that the close-but-not-quite answer is the most plausible and probable answer that neuroscience could give regarding free will. Section III will focus on explicating the distinction between corrective medical therapy and human enhancement, and then demonstrate that a free will-enabling biotechnology falls under the latter category. The purpose of Section III is to center free will-enabling biotechnology in the human enhancement discussion. In Section IV, I will demonstrate that the general objections to human enhancement fail to object to a free will-enabling enhancement. Section V will consist of responses to the particular problems regarding free will-enabling biotechnologies.

I. Free Will and The Brain

First, it is worth considering what I am assuming for free will to be an open scientific problem. The assumption is that the problem of free will (1) is not solvable *a priori*, (2) is not yet solved, and (3) will only be solved after enough relevant information about the world is gathered. It must be further specified what this relevant information would be. In order to do so, I turn to Mark Balaguer's central thesis, as well as clarify and define some terms.

The standard libertarian view, which I will be using for the

purposes of this paper, argues that in order for an action to be free, the agent must be the author of the act, and the agent could have done otherwise (often called the sourcehood condition and the leeway condition, respectively).[1] Balaguer offers an event-causal picture of libertarian free will, which characterizes these conditions as they relate to appropriately non-random undetermined events (rather than the agent-causal picture, which characterizes these conditions as they relate to uncaused causes) (Balaguer 2010, p. 67). The particular decisions Balaguer has in mind when discussing free will are what he calls "torn decisions," which he describes primarily in terms of how they feel to us. A torn decision is one where there are reasons for two or more options for which there is no conscious belief as to which is best, and the decision is made without deliberating—we have the experience of feeling torn and then "just choosing" (Ibid., pp. 71-75). Balaguer formalizes his view in the following way:

> Libertarian free will exists, iff
> Some torn decisions which are made (a) are both undetermined and appropriately non-random at the moment of choice, and (b) the indeterminacy is relevant to the appropriate non-randomness, in the sense that it generates or procures the non-randomness. (*Ibid.*, p. 68)

The important component of Balaguer's view is that the indeterminacy generates the appropriate non-randomness right at the moment of choice. For a decision to be made freely, it can't be that something has caused the decision, nor can the indeterminacy be present in some way which is irrelevant or disruptive, such as appearing just before the decision is made. Thus, Balaguer states: "...[T]he question of whether libertarianism is true just reduces to the question of whether some of our torn decisions are undetermined in the appropriate way" (*Ibid.*, p. 69). So, together with the relatively weak and uncontroversial view that mental experiences have a token-token identity relationship to neural events, this question is a purely empirical one about our neurophysiology during torn decisions.[2]

Given this characterization of free will, Balaguer offers what would have to be true of the world for free will to exist, a condition which he names "TDW-Indeterminism":

> TDW-Indeterminism: Some of our torn decisions are *wholly undetermined* at the moment of choice, where to say that a torn decision is wholly undetermined at the moment of choice is to say that the moment-of-choice probabilities of the various reasons-based tied-for-best options being chosen match the reason-based probabilities, so that these moment-of-choice probabilities are all roughly even, given the complete state of the world and all the laws of nature, and the choice occurs without any further causal input, that is, without anything else being significantly causally relevant to which option is chosen. (*Ibid.*, p. 78)

It is worth noting that in Balaguer's prescription for what must be the case for free will, he does not include any current science on the matter or take on any controversial empirical claims—what is required for free will to exist is put forth in terms that leave open how our neurophysiology would have to be.[3] For Balaguer, neural events must meet the condition of TDW-Indeterminism, whatever the facts are about our brains. It is simply a matter of empirically verifying if it is in fact the case that this condition is met. Note also, that in Balaguer's view, TDW-Indeterminism is regarding torn decisions being *wholly* undetermined. However, he points out that we might consider a continuum of determinacy that ranges from fully granting free will with the wholly indeterminate at one end, to fully undermining free will with the wholly determinate at the other (this notion of the continuum will be relevant to the consideration of plausible neuroscientific discoveries about torn decisions in Section II) (*Ibid.*, pp. 76-78).

Assuming Balaguer is correct, the neuroscientist's task seems rather straightforward, and so there is a strong possibility of eventually having an answer as to whether our neurophysiology meets the condition of TDW-Indeterminism. From here we must ask, what is the most plausible discovery the neuroscientist will

eventually make?

II. What Neuroscience May Tell Us About Free Well

One answer the neuroscientist may give us is that neural events during torn decisions occur in such a way that completely exclude the possibility of TDW-Indeterminism. For example, it could be that our neurophysiology is set up such that every torn decision is wholly determined by a fixed set of neural operations that completely and unchangeably undermine authorship and control. Thus, the neuroscientist would empirically prove that free will does not exist. This answer, however, is implausible for a number of reasons. If every torn decision were wholly determined, it would require of our neurophysiology to be completely contrary to our experience of choice.[4] That is not to say that mental experiences are always accurate to our physiology, but it seems plainly unlikely that we could be wrong every time about what we are experiencing as related to our physiology. The experience of depression, for example, does not always correlate to a serotonin imbalance, but it often does. Also, in order for our neurophysiology to be unchangeably deterministic, it would be, to a certain extent, contrary to what we currently know about matter at the quantum level.[5]

Another answer is that neural events during at least some torn decisions occur in such a way that perfectly meets the condition of TDW-Indeterminism, thus empirically proving the existence of free will. However, in order to prove that torn decisions are consistently *wholly* undetermined, it would require our neurophysiology to be set up in a highly particular way—there is only one specific arrangement which would consistently ensure the right kind of indeterminacy.

Let us consider again the notion of the continuum of indeterminacy. In a given torn decision event, the level of indeterminacy dictates the extent to which the decision is made freely. A wholly undetermined torn decision event is wholly free, while a

wholly determined torn decision event is not at all free. Along the continuum there exists a multitude of torn decision events which are partially indeterminate, and therefore only partially free, and the extent of which is decided by how near to one end the event sits. As a matter of probability, given the continuum of outcomes, there is a low probability that our neurophysiology's arrangement will be such that it allows for torn decision events to consistently land at the wholly indeterminate end of the continuum. At least, the probability of our neurophysiology's arrangement being such that torn decision events always lie somewhere on the continuum between wholly indeterminate and wholly determinate is much higher. As a result of the above considerations, the most plausible and probable answer the neuroscientist will give us is that our neurophysiology is arranged such that torn decisions are always partially determined, and so empirically prove that free will does not exist.

III. Enabling Free Will and the Corrective Therapy/Human Enhancement Distinction

Although there exists some level of vagueness in the distinction between what is considered a human enhancement and what is considered corrective therapy, there are some commonly accepted properties of biotechnological devices, treatments, and surgeries which place them firmly in one or the other category. In order to see how a free will-enabling device, treatment, or surgery would be categorized, we must simply consider the properties of human enhancement and corrective therapy in order to determine the category to which it belongs.

According to Norman Daniels' work on this distinction, corrective therapy is understood to have the properties of healing illness, returning bodily function to a state prior to some kind of damage, and helping the body function in the way that other bodies typically do. Continuing with Daniels' characterization, human enhancement is understood to have the properties of improving typical bodily function, aiming to augment a desirable capacity

which we may already be equipped for, or enabling the body to perform tasks that it was not originally equipped to perform (Daniels 2000).

It may be worth considering how the notion of 'typical bodily function' plays into the human enhancement/corrective therapy divide, and to what extent this further notion is vague or debatable. Granting this vagueness or debatability, I think we can, at least for the purposes of this paper, let our intuitive sense and common use of 'typical bodily function' guide how we might classify certain treatments, surgeries or devices as corrective therapy or human enhancement. For example, we say of brains with a serotonin imbalance that they are "neuroatypical", and we say of SSRIs that they work to return the brain to a "neurotypical" state. By way of this example, we generally understand corrective therapy by way of its intended use and the outcome of its implementation. Such is also the case with human enhancement. A common example of human enhancement is gene selection and genetic engineering. Julian Savulescu, a central proponent of human enhancement, argues that the goal of genetic engineering is to augment desirable capacities for which we are already equipped, such as intelligence and empathy.[6] Historically, the desired outcome of eugenics was to create "superior" humans.

While many devices, surgeries, and treatments may exist in the penumbra between what is distinctly therapy and what is distinctly human enhancement, hopefully I have said enough to establish the bright centers of these concepts in order to place free will-enabling biotechnology in the appropriate category. If we do not have free will, it is neither due to an illness, nor as a result of damage our bodies have sustained. If we were to undergo a treatment or surgery, or have a device implanted which would give us free will, it would not be that our body would begin to function like other human bodies. Rather, if we were to manipulate our neurophysiology such that our torn decisions would change from always partially determined to consistently wholly undetermined, we would be augmenting a desirable capacity for which we are already equipped. Or, depending on the extent to which we would

need to manipulate our neurophysiology in order to enable free will, we would be enabling our body to perform tasks it was not originally equipped to perform. Despite our feeling as though we have free will, if we in fact do not possess it, the enabling of it falls clearly on the side of human enhancement.

IV. General Objections to Human Enhancement Avoided

Despite a few prominent proponents of human enhancement, there exists a widespread anti-enhancement position among philosophers. I find the objections against human enhancement compelling. But, I will argue that a free will-enabling enhancement avoids these objections. The following two objections capture many, if not all, of those most commonly discussed regarding human enhancement:

1. *Eugenics:* Human enhancements will be *de facto* unavailable to much of the population, either due to class disparity or generational implementation, and will lead to a two-tiered society where the unenhanced hold less social or moral value.[7]

2. *Autonomy:* Human enhancements, particularly those that shape the human future without the consent of future humans, undermine our autonomy and human dignity.[8]

I will now show how (1) and (2) fail to apply to a free will-enabling human enhancement. I believe *Eugenics* has a particular potency, especially due to recent Western history. This objection works from a very reasonable assumption, namely that biotechnology will not be accessible to everyone, and from this assumption, the very dangerous and undesirable outcome of making unenhanced humans an underclass seems to simply follow. For many of the human enhancement biotechnology ideas that have floated around, I believe *Eugenics* is perhaps too strong to overcome. However, in the specific case of free will-enabling biotechnology, this objection fails to do the same work as elsewhere in

the debate. The objection hinges on the notion that the enhanced are in fact, or will be considered, "superior" by social or moral standards, and the unenhanced are in fact, or will be considered, "inferior" by social or moral standards. When considering human enhancements like genetic engineering for improving cognition or improving athleticism, it is easy to see how the notions of "superior" and "inferior" might come into play. When the enhancement results in quantifiable, outwardly visible differences, it is not difficult to imagine a scenario where the enhanced view themselves as "superior" and would divide society accordingly.

But, a free will-enabling enhancement does not follow the same pattern as other human enhancements. It is not at all obvious that we would view those with free will as superior to those who do not have free will. For example, we do not, generally, view those who live under a dictator to be socially or morally inferior to those who live in a liberal democracy—the extent of a person's political freedom has no apparent effect on a person's moral value or social worth. In cases where people are under quarantine with a contagious virus, there is not a sense of moral or social inferiority toward them or a sense of superiority amongst those not under quarantine. Those who have access to the biotechnology may gain a sense of superiority just for simply having the financial access, but that class disparity and whatever feelings come from it already exists prior to the employment of a free will-enabling enhancement—in that case, the enhancement would just be another signifier of class distinction among many. The force of *Eugenics* is we know from history that the belief that we can "make better humans" is a path to genocide. However, by employing a free will-enabling enhancement, even if it only remains available to a subsection of the human population, we have not attempted to "make better humans" and thereby rendered the unenhanced "inferior." As a result, *Eugenics* fails to apply to this case.

Autonomy first appeared in Jürgen Habermas' *The Future of Human Nature*, and is an oft-cited objection to human enhancement.[9] The objection is aimed at human enhancements which would be intended for either the unborn or newly born, and for

enhancements the implementation of which would have an irreversible effect on future generations. The force of the objection is the notion that choosing or deciding the future on behalf of others is thereby violating many future human's autonomy by revoking their ability to choose, and flouting human dignity by altering the human future and human nature without consulting future humans.

A free will-enabling enhancement, however, avoids all the issues raised by this objection. Free will-enabling biotechnology does not have to be administered prior to the age of consent, unlike the enhancements *Autonomy* has in mind, such as genetic enhancements. Further, the notion of choice, that future generations wouldn't be able to choose as a result, breaks down in a world where free will does not exist except as a result of the free will-enabling enhancements in question. The notion of human dignity remains intact in a free will-absent world, but is not threatened as a result of a free will-enabling human enhancement. Arguably, a fundamental component of our notion of dignity and human nature is itself agency. Rather than object to a free will-enabling enhancement, *Autonomy* appears to endorse it, as enabling free will would enhance our autonomy and allow us to better maintain our dignity. *Autonomy*, therefore, fails to object to a free will-enabling human enhancement. A free will-enabling human enhancement is simply not the target of these objections. In the case of *Eugenics*, the objection fails to apply because it has in mind a particular kind of human enhancement which is unlike a free will-enabling enhancement. In the case of *Autonomy*, the objection fails to apply because it assumes free will exists.

V. The Phenomenology of Choice and the Problem of Moral Responsibility

Despite avoiding the more general objections to human enhancement, there are at least two serious concerns for a free will-enabling enhancement in particular. I will discuss the following objections in turn:

3. *Phenomenology:* The phenomenology of choice is vital to

decision-making, and a free will-enabling enhancement might partially or fully disrupt our ability to make choices by changing what it is like to choose.

4. *Gyges' Ring:* A free will-enabling enhancement would also enable moral responsibility. It might be that many would opt out of the enhancement in order to avoid blame and punishment.

In torn decisions, the experience of choosing is both robust and consistent. We can extensively describe this experience, we are familiar with its qualitative character, and it is not the case that the experience of choice varies wildly from torn decision to torn decision. It is entirely an empirical question whether a free will-enabling enhancement would preserve this experience of choice. *Phenomenology* objects to the employment of free will-enabling biotechnologies based on the possible threat to our experience of choice. If the free will-enabled phenomenology of choice is too drastic a difference, it could disrupt the decision process to the extent that the free will-enabling enhancement's implementation is difficult to justify—if the experience of choosing with the enhancement is actually a barrier to making a decision, then why enhance? *Phenomenology* could be expanded to also state that, if the phenomenology of choice were similar but still noticeably different, we would not desire the change even if it did not fully disrupt our ability to choose.

In response to this objection, first, I will grant that if a biotechnological enhancement intended to enable free will disrupts the experience of choice so drastically that we cannot make decisions, then the biotechnology in question is not in fact a free-will enabling enhancement and ought not be implemented. But, I will reiterate that what it will be like to make a decision while enhanced is an empirical question. There are a number of possible ways in which the phenomenology of choice could change, and I imagine them on a spectrum between no change at all and so drastically different that it becomes impossible to make decisions. However, unlike a biological system which we discover to be

arranged one way or another, a technology is created and designed with a particular purpose in mind. If the enhancement changes the phenomenology too much, it could be simply a matter of adjusting the enhancement's specifications. It is not difficult to imagine that in the early stages of a free will-enabling biotechnology, it might fail to maintain a similar enough phenomenology of choice and require improvement, or an additional, supplementary technology to be implemented which "suppresses" the difference in phenomenology. We don't know how a free will-enabling enhancement will affect our experience of choice, and we also don't know if the experience of choice is a manipulable variable in designing a free will-enabling enhancement.

The second way I want to respond to this objection is to consider the way in which new technologies often come with a learning curve, or a learning period that precedes proficiency. While we may find that the phenomenology of choice while enhanced is different, it could be that adapting and becoming accustomed to the difference is all that is required in order to make decisions. The size of the curve would depend on how different the experience of choice is while enhanced. I contend that we ought to overcome this possible learning curve, even if it is steep, as the resulting proficiency leads to libertarian free will.

The phenomenology of choice is extremely important—one of the reasons the libertarian position is so attractive is due to its preserving our pre-philosophical notion of free will, which is primarily shaped by how things feel for us. But, I insist that, barring total dysfunction in decision-making, we ought to enhance ourselves in order to enable libertarian free will. If the phenomenology of choice is not accompanied by agency, control, authorship, or any of the other things we associate with decision-making, then we are only trying to preserve an illusion by opting out of a free will-enabling enhancement.

According to *Gyges' Ring*, if we discover that free will does not exist, we would also be discovering that moral responsibility does not exist. Enabling free will, therefore, would mean enabling moral responsibility. Avoiding a free will-enabling enhancement

would result in avoiding culpability, and this might be an attractive prospect for some. Why would we want to be morally responsible and become open to blame, punishment, or retribution when we could opt out by not getting enhanced? Although many of us find the value in being morally responsible for what we do, it is not a certainty that all of us do. If no one has free will, then we might as a matter of pragmatism act as if we are in fact morally responsible. However, introducing a free will-enabling enhancement "forces the issue," as it were.

There are two related responses I have for this objection. One quick point regarding *Gyges' Ring*, however, is that it fails to grant importance to the other side of moral responsibility, namely praise and reward. One of the central reasons free will is worth having is particularly because we can claim full authorship over our actions and receive our due credit. My first response is that *Gyges' Ring* assumes that people without libertarian free will are not morally responsible. Second, it supposes that if only some of the population were free will-enabled, a crisis would ensue because we would have to manage a certain part of the population differently due to some having no moral responsibility. I argue that these assumptions are ill-founded.

On the first point, it is wrong to assume that anyone who is not enhanced is therefore not morally responsible. If we recall the spectrum of indeterminacy and the close-but-not-quite answer we are most plausibly and probably going to discover via neuroscience, it is extremely unlikely that any torn decision is wholly determined. We would still say of these torn decisions that they are not enough for free will, as free will is defined as consistently having *wholly undetermined* torn decision events. But, we would say of these partially determined torn decision events that they are partially free, and in the standard libertarian view, the extent to which we act freely is the extent to which we can be held morally responsible. So, unlike what *Gyges' Ring* suggests, the unenhanced would still be morally responsible to some extent, and therefore remain open to blame and punishment. On the second point, as it has been shown that responsibility is not avoided by virtue of

opting out of enhancement, it is wrong to assume that any serious problem would result from a free will-enabling enhancement. We can see how the United States legal system deals with varying levels of culpability as an example of how we might approach a society where some have free will and others do not.

Phenomenology and *Gyges' Ring* fail to object to a free will-enabling human enhancement. Maintaining our current phenomenology of choice is not more valuable than actually being able to act freely, though ideally we would want both. As with most technology, we have the ability to adjust the way it functions and fine-tune it until it works the way we want. There is also a certain expectation with new and life-changing technology that a learning curve accompanies its adoption. We have adapted to numerous technologies which fundamentally change the way we experience the world, such as the internet and smart phones, cochlear implants, or biomechanical prosthesis. It seems a small price to adapt or struggle with a new way of deciding if it would result in one consistently having authorship and control over their torn decisions. Moral responsibility is a vital component to free will, but is not totally absent in the absence of free will in this picture. So long as our decisions are not fully determined, we maintain a certain level of moral responsibility, and as a result, would not face a crisis in the face of free will-enabling technology.

Conclusion

Thinking about the future of neuroscience and the implications that would result from discoveries about indeterminacy in the brain during a particular subset of decision events may seem like an exercise in fanciful speculation. I hope to have shown, however, that if we take seriously the notion that free will is an open scientific problem, there are a number of issues worth addressing. The human body is manipulable, and our rapid advances in biotechnology may present us in the near future with an opportunity to access real libertarian agency, something core to the way we understand human life and our own forward path. As a human enhancement, free will-enabling biotechnology faces the

objections posed against all human enhancements. But, as I have shown, a free will-enabling enhancement avoids these objections in virtue of their assumptions and the nature of the enhancement itself. There are some legitimate concerns regarding a free will-enabling enhancement specifically, perhaps most significantly that we wouldn't want to change the way it feels to make decisions. However, valuing the maintenance of an illusion over actually having authorship and control is the result of confused priorities. Unless the technology we might develop completely disrupts the decision-making process, thereby not fulfilling its intended function of enabling free will, there are no good reasons not to enhance oneself in order to enable libertarian free will.

Notes

1. As far as I can tell, this conception of free will is non-controversial and advocated by most libertarians, including event-causal libertarians like Robert Kane. More importantly, this is the conception Balaguer employs for his view, which he states on page 7 of *Free Will as an Open Scientific Problem*.
2. Balaguer makes this argument in 3.3.1.1
3. Balaguer points this out on 69-70 in order to contrast his view specifically with Robert Kane's.
4. Balaguer makes a similar point in his argument from phenomenology in 3.3.1.2
5. Here, I am referring to quantum field theory and resulting conclusions based on this probabilistic interpretation of quantum mechanics.
6. Savulescu has published numerous papers on this topic, perhaps the most relevant of which is "The Moral Obligation to Create Children with the Best Chance at the Best life" (2009).
7. This objection exists across the literature, but for this paper I followed Robert Sparrow's "A Not-So-New Eugenics" (2011).
8. The origin of this objection comes from *The Future of Human Nature*, by Jürgen Habermas (2003).
9. Habermas' objection to human enhancement is discussed throughout the literature, including Giubilini and Sanyal (2015), Pugh (2015), Zylinska (2010), and Edgar (2009).

Bibliography

Balaguer, Mark. (2010) *Free Will as an Open Scientific Problem* (Cambridge: MIT Press)

Daniels, Norman. (2000) "Normal functioning and the treatment-enhancement distinction," *Cambridge Quarterly of Healthcare Ethics Committees* 9, no. 3, pp. 309

Sparrow, Robert. (2011) "A Not-So-New Eugenics," *Hastings Center Report* 41, no. 1, pp. 32-42

Modal Meinongianism and the Many Worlds of Nonexistent Objects

Cameron Takeda

I. Introduction

When we are talking about things like unicorns, Zeus, or Astro Boy, what exactly are we talking about? One thing that all three have in common is that none of them seem to exist in actual reality. However, despite that they do not exist, it does seem like there are qualities not common between them, such as unicorns being horse-like creatures that have horns while Astro Boy is a robot and has no horn at all. But what is it that makes these statements true or false? Furthermore, how could we even come to know this? So, then what these names pick out seems actually to be somewhat mysterious. Modal Meinongianism, sometimes known as Noneism, can be succinctly described as the view that some things do not exist, but we can meaningfully discuss them anyway. This seems pretty intuitive at first glance, but actually is a hugely controversial view with a rich philosophical history.

The history involved here more or less starts with Alexius Meinong and his views that have been given the name, *Meinongianism*. Meinongianism is a word often met with displeasure in much of the Western Philosophical Tradition, thanks mostly to W. V. O. Quine's 1948 paper, "On What There Is." Quine's paper is often seen to have reduced Meinongianism to something of an absurd theory. In Quine's article, the character Wyman, who is usually seen as a caricature of Meinong, is taken to "overpopulate" the universe with these nonexistent objects as somehow having being and nonexistence (Quine 1948, p. 4). I will not discuss this debate here further, but Graham Priest offers some good reasons to think that Quine has not fully defeated Meinongianism (and espe-

cially not Modal Meinongianism) in the fifth chapter of his book, *Towards Non-Being* (Priest 2016, pp. 105-115). Meinongianism saw something of a revival in the past fifty years with philosophers such as Terence Parsons and Richard Routley/Sylvan, who coined the term *Noneism*, attempting to show that Meinongianism itself is not a dead theory. These works on different, seemingly less problematic versions of Meinongianism would influence current philosophers such as Graham Priest and Francesco Berto in their formulations of Modal Meinongianism. Their views actually differ slightly, but their general accounts of Meinongianism are reasonably similar, so hereafter I will refer to their general account of Meinongianism as Modal Meinongianism. Overall, I intend this paper to elucidate features (and potential issues) of Modal Meinongianism by adjudicating between some significant objections to Modal Meinongianism and the responses by its proponents. Furthermore, I hope to show that this is a live and significant debate within the realm of philosophy and not merely discussion of a dead theory.

In this paper, I will explain the general idea of Modal Meinongianism, focusing on the Characterization Principle. Before this, I will give some intuitive reasons to think that we should take metaphysical theses about nonexistent objects seriously. Then I will address some objections by Fred Kroon toward the modal version of the Characterization Principle as well as the responses to Kroon by Graham Priest and Francesco Berto, and attempt to argue that although Kroon's objections fail to completely defeat Modal Meinongianism, his objections still highlight some issues that aren't clearly solved by Modal Meinongianism.

II. Methodology: Why Save Nonexistent Objects?

Intuitively it seems like we must be able to characterize nonexistent objects meaningfully, mostly because we do it all the time in folk-language. I think unicorns have one horn and I admire the detective skills of Sherlock Holmes, but I still accept that they

do not exist in reality. Theories like Meinongianism and Modal Meinongianism aim to save these intuitions that we really mean what we are talking about when we talk about nonexistent objects in folk-language. But here there is already an immediate trade-off in philosophical methodology. Some will accept the methodology that we ought to save our intuitions about what we mean when we speak, whereas some question the significance of intuition in latching onto metaphysical truths in reality. I will offer a somewhat short defense of this motivation, but I admit it may not be fully compelling to someone who completely denies the usefulness of intuition in metaphysics.

I believe that there is a significance found in intuition about nonexistent objects in at least two types of cases. The first type is the basic idea that people mean what they say, and can be perfectly understood regardless of the strangeness. For example, I can tell someone, "Unicorns have one horn and a horse-like body." And I would find it strange for them to respond with, "No, they don't have any horns nor a body, unicorns don't exist," despite that this is a true statement. It would be more expected to have some general agreement that this is true about what unicorns are like, namely that they are one-horned and horse-like, even though they do not actually exist. Immediately, an alternative account could be had, where 'true' meant "true-in-the-fiction-of-unicorns-existing." This is not wholly implausible, but then how could we agree that unicorns have one horn, but also agree that they do not exist under this interpretation? Certainly we do not mean, "Unicorns have one horn and don't exist" is true-in-the-fiction-of-unicorns-existing. Obviously, it would be false in that fiction. Instead perhaps it is something like, "Unicorns have one horn," is true-in-the-fiction, but "Unicorns don't exist" is true-in-actuality, and we just strangely and incorrectly conflate these two kinds of truth constantly. In fact, Priest and Berto seem to think that something like this is close to what we do, claiming that "We move seamlessly from truth in reality to truth according to a fiction and back all the time" (Berto 2014, p. 190). Now, I believe this to be right, but I also think that there still is a strangeness involved in

being able to be properly understood in these contexts that leads to thinking that our intuitions about these are, while not completely truth-seeking, significant enough to make me think further discussion is required.

The second type of case is inspired by the so-called "Paradox of Fiction." This is not a logical paradox per se, but is still a strangeness about how fiction relates to people. The "Paradox of Fiction" is basically the issue that arises out of our immediate emotional reactions to the events that occur in fictions despite the acknowledgement that these fictional objects in the stories are strictly nonexistent (Dadlez 2015, pp. A1-A2). For example, consider watching a horror movie in a movie theater, and the movie monster is hunting down the various characters in the movie. Indeed, movie-goers acknowledge that the movie monster is nonexistent, but at the same time they are scared nonetheless. One might have an increased heart rate, or be startled in their seat, or possibly even leave the theater out of fright. But, then, what were they afraid of? It seems like it was the movie monster that was frightening despite not actually existing, and this is rather strange. A further example includes reading and following along with the tragic life of the protagonist of some sad story; one might feel sadness for the character, and one might even begin to cry when the character dies within the story. So, again, why cry for what amounts to nobody dying, since the character never existed in the first place. In fact, numerous real, existing people die every day; and we usually do not cry over these people since we do not know them personally. How or why is it that we can be sadder over the "death" of a nonexistent person than of a real person?

Now, some might claim that there is some roundabout psychological mechanism happening during these bouts of immediate emotional response. Perhaps reading about "the death of X" in the story reminded you of some other person that you knew and liked in the past who had also passed away, and that memory is what caused the crying to occur. I am no psychologist or neuroscientist, so I do not want to claim that I know what this mechanism is like; however, I don't think it actually matters. We're talking

about meaning here, specifically what we mean, which seems to be most understandable by ourselves. And to bring intuition into importance again, the immediate emotional response seems to be the most brute and unthinking part of the intuition. Although these may not be utterances of language, they still have meaning and so, are about something. At the very least, the emotional responses seem to mean something about how you feel, and having an emotional response is a response to something, so that something must be what the emotional response is about. However, the significance of this can be immediately challenged. Some claim that if these were real emotions, we would not just cry or be startled in our seats; we would feel some sort of real despair or we would start running away. So, instead, we must experience some form of pseudo-emotions which has some similar features with the actual emotion, but not all (Dadlez 2015, pp. A2-A3).

I have two responses to this objection. The first response is anecdotal. When I engage with sad fictions, I sometimes feel what I think is this sort of real despair, the kind that really brings tears and emotional pains. But I must admit that I cannot claim it to be universal amongst other people and I would have no real explanation of the difference between myself and another that could be involved. However, I am not entirely sure what the difference between "real despair" and "pseudo despair" would be, and how we might know which one was which. Nevertheless, I must assume that given the same sorts of conditions, others would feel at least a similar kind of feeling, and I consider that to be a real emotion.

The second response is that it's not clear that what a proper reaction to a real emotion is. It is supposed to be clear that jumping in our seat is not a proper response to a monster appearing in front of us (on the screen/in the fiction), so it must be a pseudo-emotional response. This might seem convincing at first, but it does not really tell the whole story. From an early age, we are taught how to distinguish reality from fiction (likely even more now as technologies such as virtual reality continue to advance). Instead, consider what it might be like to be thrust from some

pre-movie time, let us say Ancient Sumer, to a modern time in the middle of a movie. I think it's quite reasonable to say that (barring that they would likely go crazy from all the cultural/technological differences) they might actually think the monster on screen is an actual physical monster coming toward them. It takes looking no further than some time-travel fiction to see that others also seem to believe that this would be the case. So, it does not seem that it is our emotion that is wrongly described, but just that we have been trained in certain cases to act a certain way; and being trained to have different responses to the same emotion in different contexts does not show that the emotion was merely a pseudo version of that emotion.

Consider the difference between being angry with a friend versus being angry with a random rude person on the street. Just because in one context you might shout expletives at one of them and in another context you do not shout the same expletives, does not seem to give any indication that in one you experienced anger and the other it was pseudo-anger. But even if it was a pseudo-emotion, it was still a pseudo-emotion about something, and that story still remains somewhat mysterious. Either way, I think these issues give us a good reason to think that there might be some significance to how our intuitions might pick up on seemingly nonexistent objects. In general, however, a rich discussion on this methodological question is beyond the scope of my paper. Nevertheless, I think these examples still give reason to think that intuition is an important factor to consider in this debate over how we discuss nonexistent objects. And this notion that intuition is significant to consider in discussion of nonexistent objects is important in the articulation of the Characterization Principle in the various forms of Meinongianism.

III. The Characterization Principle

Meinongianism itself is basically a theory of intentionality. That is to say, how our minds are "directed toward" something else. In this idea, there are some cases where our minds are directed toward things that do not exist, such as when we think about things

like Zeus or Santa Claus. From this, Meinongianism claims that, "there is indeed an object for every mental state whatsoever—if not an existent object then at least a nonexistent one" (Reicher 2016). That is, when we think about things that do not exist, we are still thinking of something, just not something that exists. However, in his theory, Meinong offers a rather strange view of how to discuss the property of existence. He brings up the notions of subsistence and absistence, which are something like lesser degrees of existence, and this is questionable as to what these are supposed to be like in reality. Overall, as Quine had noted, it is mostly a strange and rather unappealing view regarding existence. However, Modal Meinongianism does not make claims of existence quite as strange. Instead, the main relation of Meinongianism to Modal Meinongianism comes from an idea known as the Characterization Principle.

One of the important claims held by Meinong was that an object's *Sosein* is independent of its *Sein*, which can be described as the claim that, "objects have a way of being such-and-so whether or not they have any form of being" (Kroon 2012, p. 25). So, essentially, this is the separation of some thing's description and properties from its status as existing or not. To see why this claim might be acceptable, consider the properties of the golden mountain. It seems almost tautological that the golden mountain is golden and a mountain. But the golden mountain does not exist, so is the golden mountain really golden or a mountain? It seems that the answer is that the golden mountain is actually neither golden nor a mountain, and this is somewhat odd. This idea is subsumed in the articulation of the Characterization Principle, that something can have properties regardless as to whether or not it exists.

The naïve form of the Characterization Principle of Meinongianism can be described as, "Given any property A, A holds of the thing characterized as being A" (Kroon 2012, p. 25). Less formally, this is to say that, for any given property, something satisfies that property, namely the thing with that property. Obviously, this is problematic immediately, throwing in existence (or maybe something less controversial like extension if that would help)

would somehow make things pop into being by merely describing it. Furthermore, beyond that, many properties are seemingly *existence entailing;* that is to say that if it has this property of being such-and-such, then it must exist. For example, the property of being 10-feet tall. For something to be 10-feet tall, it must exist. So, of course, almost nobody (including Meinongians) continues to use this unrestricted form, instead they turn to attempting to fix the Characterization Principle by restricting it in some way or another.

Some attempted restrictions on the Characterization Principle mostly involve restricting it to only certain kinds of properties, sometimes called 'nuclear' or 'characterizing' properties to eliminate these problems. But, it is difficult and may be impossible to figure out just what properties are of this type without being arbitrarily defined to be of that type. Another approach is known as the Dual Copula method, whereby the Characterization Principle is relational, and holds that there are two modes of predication, and one of them *encodes* the properties rather than always instantiating them (Berto 2014, p. 184). These views are not without their problems, but these two forms will not be further discussed here. Here I will focus on Graham Priest's restriction of the Characterization Principle as involved in Modal Meinongianism. And here is where the modal aspects of Modal Meinongianism come in. Priest's restriction of the Characterization Principle involves holding the naïve Characterization Principle, but sometimes at worlds other than the actual. I will refer to this form of the Characterization Principle hereafter as the Modal Characterization Principle.

Before fully explaining Modal Characterization Principle, a basic note on modality and possible worlds must be first explained. Statements of modality are the statements that are about possibility, impossibility, and necessity. Furthermore, one way of understanding these statements are to describe these modal notions in reference to "possible worlds." For example, to say, "It's possible that James went to the moon," can be understood as. "There is a possible world where James went to the moon." This

view is widely received, though not completely uncontroversial, especially when it comes to what exactly these "possible worlds" are supposed to actually be, but there is a wealth of discussion written on the subject, though a full discussion on this is beyond the scope of this paper.

The Modal Characterization Principle according to Priest can be described as, "the object characterized by a representation has the characterizing properties, not necessarily in the actual world, but in the worlds (partially) described by the relevant representation" (Priest 2016, p. 84). That is to say that something like Astro Boy has the property of being a robot, just at the world(s) described and represented by the stories created by Osamu Tezuka and not necessarily the actual world. This Modal Characterization Principle differs from the independence of *Sosein* from *Sein* in Meinongianism in that only some properties of an object are independent from whether or not it exists, but some other properties are still existence-entailing at some world, which can include the actual world (Berto 2014, p. 185). But, before continuing any further, it is worth noting that the worlds other than the actual on Priest's account need not necessarily be Lewisian concrete worlds and do not entail the acceptance of some sort of Modal Realism (that possible worlds are actually concrete worlds). Furthermore, given that Priest allows for impossible worlds and even stranger open worlds (which are not relevant here), the status of what these worlds are supposed to be is left open to debate. However, even without considering the exact status of worlds in this conception, it is from the various implications of this Modal Characterization Principle that some objections are raised.

IV. Kroon's Objections

The philosopher Fred Kroon raises some objections to Priest's account of Modal Meinongianism in his paper, "Characterization and Existence in Modal Meinongianism." His objections mostly revolve around the implications of the Modal Meinongian version of the Characterization Principle. Kroon argues that accepting the Modal Characterization Principle leads to the problems that,

"[Modal Meinongianism] is not only much more unfriendly to central Meinongian intuitions than its proponents allow, but as it stands it cannot even endorse the Meinongian truism that an object like the golden mountain lacks existence" (Kroon 2012, p. 24). In this manner, Kroon raises two objections concerning the Modal Characterization Principle.

The first is that Modal Meinongianism, with its Modal Characterization Principle, still leads to statements like, "the golden mountain is not actually golden," being taken as true. Since, under this conception, the golden mountain is not actually golden, and this is because it is not existent at the actual world. Instead the golden mountain is only golden at some worlds other than the actual that realizes the way that it would be if it did exist. But Meinongianism overall seems like it aims to save the intuitive notion that the golden mountain has to actually be golden for it be a golden mountain. So, if Meinongianism generally aims to save these intuitive notions, it is strange that the supposed "better" form of Modal Meinongianism does not even save the intuitive notion that it aims to save. Of course, this objection would not show Modal Meinongianism to be false per se, but then we ask what impetus do Meinongians (and others) actually have to endorse Modal Meinongianism over other views that try to account for these intuitive notions?

The second, more important, objection that Kroon raises is that there is an issue involving just what we know (or can know) about nonexistent objects given the Modal Characterization Principle. He notes that Modal Meinongianism accepts that some nonexistent objects have properties still at the actual world such as being thought about by someone who exists at the actual world. But then this leads to a further question; he then asks, "How do we know that nonexistence is among [the object's] properties?" (Kroon 2012, p. 28).

Kroon gives two arguments to show this problem; the first is about *de re* imagining. By *de re* imagining, that is to say, imagining that is about the specific object picked out by a definite description. This argument says that we can direct our imagination

towards actual existing things, and this can lead to some counter-intuitive implications. For example, consider playing an imagination game; say you are looking at an actual existing wooden stick and imagining that it is a magical mythril blade. Additionally, one can further imagine more scenarios while continuing to imagine this wooden stick as a magical mythril blade. So, this object exists at the actual world as a wooden stick, but at the same time exists as a magical mythril blade at the worlds which realize the sort of world you imagine when you do imagine it. Therefore, in some respects at some other worlds, at least some wooden sticks are magical mythril blades that exist somehow (Kroon 2012, p. 29). This notion is odd, how can a wooden stick be mythril (so not wooden) and a blade (so not a stick)?

The second argument is a modal argument that concerns fictional characters. When it comes to questions about fiction, Modal Meinongianism claims that these fictional objects/characters have properties, just only at the world(s) that realize the way that they are represented as being when one engages with these fictions. But then consider the making of arbitrary counterfactuals when engaging with fiction. For example, perhaps when reading Catch-22, the world I represent includes the claim, "Had Yossarian not been in the military, he might have become an opera singer." Appealing to Modal Meinongianism says that this claim is true at the worlds that realize the way I represent it when I engage with this fiction. Then Kroon adds in the note about modality that, "the 'might have' is the 'might have' of logical possibility" (Kroon 2012, p. 30). From this note, under Modal Meinongianism, there must be worlds where Yossarian exists, but is an opera singer and not in the military. And he continues with this argument, adding that, "There is no a priori reason, however (at least no a priori reason officially sanctioned by Priest's theory), for thinking that the actual world is not one of these worlds" (Kroon 2012, p. 30). Given this, we additionally have no a priori reason to conclude that the claim that Yossarian was an opera singer at the actual world is false. This is a very strange implication, since it seems to be a given (or even seemingly analytic) that Yossarian, as a

fictional character, does not exist in the actual world. But to put this another way, it is to say that we have no a priori reason to think that Yossarian, a fictional character, is actually nonexistent. But perhaps it could be said that I represent all fictional characters as nonexistent, so then I can know that they're nonexistent. However, the Modal Characterization Principle would read that statement as just saying that at some world or another, the characters do not exist, so I still cannot know whether the actual world is one of those worlds or not.

This further leads to the issue of how one can actually intend a nonexistent object. In order to intend a nonexistent object, it seems like one needs to actually know that it is nonexistent first. Given that we have no a priori knowledge about whether or not some intended object exists at the actual world, Kroon reminds us that, "We can make mistakes when intending an existent object . . . So why not when intending a nonexistent object?" (Kroon 2012, p. 32). For example, when we see an oasis in a desert, we can think about this oasis and how cool and refreshing the waters will be. However, this oasis is a mirage, so does not actually exist. And so, there's the question as to whether there is a reversed analogue where one intends a nonexistent object, but it turns out that it actually exists. And again, since under this conception, we cannot know a priori whether the object is nonexistent, it seems like this mistake can occur. So, if these objections are correct, not only does Modal Meinongianism not actually save the intuitions that it is out to save, but also cannot even account for what is actually nonexistent.

V. Modal Meinongian Replies

In reply to Kroon's objections, Graham Priest defends Modal Meinongianism in the second edition of *Towards Non-Being* (2016), as well as in "Modal Meinongianism and Characterization Reply to Kroon" with Francesco Berto.

In response to Kroon's first objection that Modal Meinongianism does not account for the intuitive notions that they are supposedly out to save, they note that what is claimed to be an

important intuitive notion shared by most forms of Meinongianism is actually the acceptance of a related theory called Literalism. Literalism is the view that fictional/nonexistent objects, "literally and really have the (nuclear) properties they are characterized as having (in the relevant fictions)" (Berto 2014, p. 187). Not only do they note that other forms of Meinongianism, such as the Dual Copula version, do not accept literalist intuitions, but also argue that the seemingly intuitive notions for Literalism are actually counterintuitive.

To elucidate this counterintuitive notion, it is helpful to consider reference to non-fictional actual existing people within fiction. For example, I could write some fan-fiction where I meet all my favorite fictional characters and they each give me a memento to remember them by. But this does not make it literally the case that I have met my favorite fictional characters, nor do I literally have these mementos. Not only are literalist intuitions just straight up false at the actual world, but not actually intuitive at all. It would be similar to asking for directions to take you to Metropolis or Gotham City. Just because they are in the U.S. in the worlds described by DC-Comics writers, it does not mean that we can literally find them in the U.S. And asking any U.S. citizen for these directions, they would quickly say something to effect of, "Those aren't real places." If this is not enough to get rid of these literalist intuitions, there is a further argument that literalist intuitions are counterintuitive because they are committed to numerous significant contradictions. There are at least three ways that literalists are committed to contradiction. The first is that there can be internally inconsistent fictions that contain contradictions as part of the story. For example, in Grant Morrison's *Animal Man*, the titular character recognizes that he received his powers at age 30 and has no children, but he also remembers that he received his powers at age 19 and has two children. However, this is an intentional contradiction in the character to describe the strange continuity of comic-book characters and fits into the larger metafictional plot. The second way is that some fictions can contradict other fictions, such as the case in Superman comics. In Superman:

Red Son, Superman lands in the Soviet Union after Krypton is destroyed, whereas the normal Superman-fictions has Superman having landed in Smallville. The third is the most commonly noted; some fictions contradict reality (Sainsbury 2010, p. 31). This is just as was mentioned previously; a fiction might say "Sherlock Holmes lived at 221b Baker Street," but in reality, something else occupied the space of 221b Baker Street. So, given this large set of contradictions to accept in Literalism, these "intuitive" notions about Literalism are actually wildly counterintuitive.

However, although the literalist intuitions actually seem both false and counterintuitive, it also seems like there is still some area of explanation to do here. One worry here is that if all statements about nonexistent objects are prefixed with "according to the (relevant) story," then we can end up with some strange implications. Here is a question: who is physically stronger, me or Superman? I suppose that should be read as, "according to the world described by DC-comics writers, who is stronger, me or Superman?" According to that world, Superman is inhumanly strong, whereas I do not exist in that world, so that question does not have an answer, or is meaningless. Instead maybe it is, "According to the actual world, who is stronger, me or Superman?" But the same problem arises, I am not very strong, but Superman does not exist, so there is no meaningful answer. Yet at the same time, we do want to say that it even seems somewhat obvious that Superman is stronger than I am. So, while the literalist intuitions are counterintuitive (and wrong), rehashing them as "according to the story" does not manage to solve all problems either. I will not press this further, but I think there must be some more complete way to account for this issue.

To the second of Kroon's objections, of how we can know that nonexistent objects have nonexistence as a property, Priest and Berto's response involves distinguishing between speaker referent and semantic referent (Berto 2014, p. 191). The speaker referent is what the speaker intends to refer to when they perform an utterance. The semantic referent is what is actually being referred to. This distinction can be highlighted with an example

involving being slightly incorrect in describing a referent. For example, perhaps I say, "That woman over there is beautiful," but unknown to me, the person to whom I'm referring is actually a man. The semantic referent is still this man; that is to say the meaning of the utterance did not wholly depend on my intention. Continuing, they note that additionally, the semantic referent of a definite description is context dependent. Furthermore, what gets picked out by this definite description may not be what actually satisfies the description due to differences in time and/or place. For example, if I say, "John Wilkes Booth is infamous for having killed the President." Even though I am in the US and it is 2017, I was discussing U.S. history, so the referent of 'the President' is not Donald Trump, but Abraham Lincoln. From this, Priest and Berto add that there is a modal version of this same context-dependence phenomenon. For example, consider, "The strongest person living on Earth." This certainly picks out someone on Earth, likely a muscular person who competes in weight lifting competitions. However, if I was discussing the world described by the various writers of DC-Comics, "The strongest person living on Earth" likely points to someone else, such as Superman. So, these definite descriptions can semantically refer to things that do not actually have these properties since it can refer to objects that do not exist.

With this in mind, when it comes to Kroon's argument from *de re* imagining, Priest and Berto argue that Kroon is mistaken in thinking that "what was referred to by the description was, context-independently, a unique thing" (Berto 2014, p. 193). That is to say that there are numerous answers based on different contexts and/or different speaker-intentions. So how we come to know that nonexistent objects, such as the golden mountain do not exist, can be figured out differently in many different contexts. Sometimes it is just empirically discovered, such as in the case of my nonexistent thirteenth finger; all it takes is counting my ten fingers to discover that my thirteenth finger does not exist. So, again consider the imagination game involving the wooden stick. We make up a story *de re* about it, and we refer to it as a magical

mythril blade. But it is understood that this wooden stick is not a magical mythril blade at the actual world. So, we can know that "magical mythril blade" just refers to this wooden stick at the actual world, but also can refer to a magical mythril blade at the worlds which realize the way my imagination represents it to be. But this does not mean that the wooden stick *is* mythril and blade-like, just that the object referred to at the actual is a wooden stick and that the object referred to at some other worlds is a magical mythril blade.

But still the problem remains of how we can intend a nonexistent object. Since Kroon's modal argument showed that, since we cannot know the nonexistence a priori, intending a nonexistent object is actually quite problematic. They acknowledge that, "this cannot in general be settled by [Modal Meinongianism] as such: it depends on how things turn out in the world" (Berto 2014, p. 197). One can intend to refer to all sorts of things, and if what they intend to refer to does not obtain at the actual world, then they have intended a nonexistent object. It can just happen that we can think about things like your thirteenth finger. I really think that your thirteenth finger is a nonexistent object, though it could exist (not to discriminate against any thirteen-fingered beings reading this), but if it is not the case that you have thirteen (or more) fingers, then I have intended a nonexistent object. It seems to be the case that we can do this, as when we are mistaken about existing objects, or not-fully-knowing whether something exists or not, but having it be the case in the actual world that it does not exist. But can we intend something about objects that we seemingly know do not exist? How could we pick them out amongst all the nonexistent objects? Priest argues this is just an act of "primitive intentionality," that is to say it is just the nature of intentionality and phenomenology (and maybe imagining), that we can do this (Priest 2016, pp. 208-210). And it seems that we can; for example, I can imagine a woman named Holmes* with a deerstalker cap, who is a detective (so now this nonexistent object is phenomenologically present to me); I can admire Holmes* as well. I do not think I'll ever get this Holmes* confused with the

Sherlock Holmes of Doyle's stories. When people talk of Holmes, the detective, I will never be confused whether they are talking about the Holmes* I imagined. However, it still must be stated that this primitive phenomenological/mental pointing at nonexistent objects is not wholly uncontroversial, so it may be outright denied by those who take it to be the case that we cannot do that. But, it really seems to be the case that we can do that, so I take it that we can intend something about nonexistent objects. So, though these objections do not actually show any damning problems with Modal Meinongianism, the problems they do raise show that there is still further discussion to be had on the topic, and that Modal Meinongianism is far from being a complete view.

VI. Conclusion

Claims about nonexistent objects seem to be much less mysterious than they actually are when we discuss them. And intuition appears to be a great motivator in thinking that we actually can properly find a theory to account for the strangeness involved with characterizing (and knowing about) nonexistent objects. It seems that, intuitively, we need to be able to account for how we can properly discuss (or even engage with) nonexistent objects. Various forms of Meinongianism attempt to account for these intuitions. The Characterization Principle, and the modal version of it, are central to whole idea of Meinongianism. And though the Modal Characterization Principle is what separates Modal Meinongianism from other forms of Meinongianism, it carries some implications that some may find problematic. But though Modal Meinongianism has some answers to the issues that Fred Kroon raises, it was noted in the response by Priest and Berto that Kroon's criticisms were, "the most perceptive criticism of [Modal Meinongianism] to date" (Berto 2014, p. 199). But equally, they acknowledge that there is a further discussion to be had on the subject. So, although the objections raised by Kroon have some reasonable responses, they still highlight some issues with all forms of Modal Meinongianism. Some questions that remain are how exactly are we able to engage with these nonexistent objects?

And how justifiable is appealing to primitive intentionality to answer that question? Also, there still lies the question of how much the intuition should be considered as justification, given that the literalist intuitions were shown quite rightly to be problematic and incorrect. But Modal Meinongianism is a relatively recent theory, and seems to actually get a lot correct about how we talk about nonexistent objects. So, Modal Meinongianism still remains a reasonably acceptable theory to hold, but there is still more explanation to be done in order for it to be complete.

Bibliography

Berto, Francesco & Priest, Graham. (2014) “Modal Meinongianism and Characterization Reply to Kroon,” *Grazer Philosophische Studien* vol. 90, pp. 183-200

Dadlez, E. M. & Haramia, C. M. (2015) "Fictional Objects, Future Objectives: Why Existence Matters Less Than You Think," *Philosophy and Literature*, vol. 39, no. 1, pp. A1-A15

Kroon, Fred. (2012) “Characterization and Existence in Modal Meinongianism,” *Grazer Philosophische Studien* vol. 86, pp. 23-34

Priest, Graham. (2016) *Towards Non-Being: Second Edition* (Oxford University Press)

Quine, W. V. O. (1948) “On What There Is,” *Review of Metaphysics* vol. 48, pp. 21-38 (Cambridge, Mass.: Harvard University Press)

Reicher, Maria. (2016) "Nonexistent Objects", *The Stanford Encyclopedia of Philosophy* Retrieved on January 21st, 2017 from https://plato.stanford.edu/archives/win2016/entries/nonexistent-objects/

Sainsbury, R. M. (2010) *Fiction and Fictionalism* (New York, NY: Routledge)

Names, Rigidified Descriptions, and Persistency

David Fonth

Introduction

This essay aims to examine the discussion surrounding the plausibility of Kripke's rigid designation thesis. In particular, I will examine the debate between those who try to appeal to Kripke's modal argument as a motivating force for rigid designation, and those who try to undermine the argument by appealing to descriptivism. One well-known descriptivist stance, as noted by Chen Bo, against rigid designation is the idea that any description can be rigidified and thus turned into a rigid designator, which would seem to undermine Kripke's enterprise on rigid designation, since rigid designation was meant to explain the semantic difference between proper names and descriptions. However, even if we accept Chen Bo's thesis that any description can be successfully rigidified, this does not entail a dissipation of the seemingly intuitive difference between proper names and descriptions, since the distinction between obstinately rigid and persistently rigid designators may offer another option to consider. Ultimately, I aim to show that, insofar as we grant the strain of rigidified descriptivism described in this essay, either we should accept the consequences of obstinacy, or we should accept the failure of Kripke's modal argument. Either way, however, a difference between proper names and rigidified descriptivism will still remain.

At the base of all of this lies the question, "What do names mean?" Most of us believe that we could provide a brief answer to this question, since we use names on a daily basis. You may think, for instance, that a name is just shorthand for a description that you associate with an individual. However, does this mean that, if we are false in our description, the name fails to refer to

any unique individual? If this is not true, then it does appear to aid the position of those that do not take names to be shorthand for descriptions. But then how else would names get their meaning? Are names completely devoid of descriptional substance? Or are names actually descriptions masquerading as singularly referring terms? This essay examines one possible way that one could answer these questions—through the adoption of Saul Kripke's rigid designation thesis—and assesses one facet of its defensibility against those who seek to undermine it.

In Section I, I will set up the context of the discussion at hand by providing a general summary of the descriptivist atmosphere that permeated much of the 20th century discourse on philosophy of language leading up to Kripke's theory of rigid designation. In Section II, I will briefly introduce the semantics of modality, especially those pertaining to possible world semantics and counterfactual conditionals, in order to make clear the distinction between modal and non-modal propositions. Section II will culminate with Kripke's theory of rigid designation, which is the main focus of the debate in question. In Section III I will introduce Chen Bo's argument against rigid designation, which focuses on the idea of rigidifying any type of description in order to equivocate, in general, rigidified descriptions with rigid designators. Section IV will examine the distinction between *obstinately rigid* and *persistently rigid* designators, and will provide the motivation behind adopting obstinacy in order to save Kripke's modal argument. Section V, I will introduce Millianism and demonstrate its applicability to the obstinacy thesis. Section VI will then evaluate the compatibility between Millianism and persistency, and discuss its consequences for Kripke's modal argument. Lastly, Section VII will include my final remarks on the status of the debate entertained in this essay.

I. Background

In order to first understand the philosophical climate that surrounded discussion on names and definite descriptions, it is be beneficial to examine the precursors to Kripke's theory of rigid

designation. The philosopher who arguably left the earliest impact on modern discussion of philosophy of language was Gottlob Frege, who tried to formulate an extensional theory of language; that is, a theory which bases an expression's meaning on its extension, or reference, within the external world. In order to solve the problems associated with 'a=b'-type statements and propositions with intensional contexts, he introduced the notion of "sense," which communicated the cognitive significance—the thought that the expression introduces in the individual's mind—of an expression (Frege 1892, pp. 37-40). A sense, or a "mode of presentation", therefore, could be understood as identifying some property of the object being referred to, which oftentimes took the form of a description of the object (Frege 1892, p. 37). Frege's conception of meaning was thus a descriptive theory of names, since descriptions played a role in the meaning of a name.

This descriptivist stance on proper names was also held by Bertrand Russell, who believed that proper names were actually definite descriptions in disguise. Although Russell rejected Frege's notion of "sense", since it seemed to require too much of an ontological commitment, it is obvious that his attitude towards proper names was descriptivist in nature (Russell 1905, pp.481-493). This philosophical perspective on proper names was further echoed with John Searle's cluster theory, which, despite also rejecting the notion of "sense" put forth by Frege, associated a name with a cluster of descriptions that were true of the object that was being referred to (Searle 1958, pp.171-172). And even though Searle's conception of the relationship between proper names and descriptions was not as strong as that of his predecessors, descriptions were still needed in order to fix the referent of a name to a unique object in the external world. Saul Kripke, however, strayed from his predecessors by conceiving of a theory of proper names that did not rely on definite descriptions, while simultaneously solving the flaws of the earlier descriptivist theories.

One of the main flaws in question was how one could successfully refer to someone, using a name, while being completely wrong in the only description that she could attribute

to the individual referred to. For instance, if I were speaking to a philosophically-illiterate colleague of mine, who believed Plato to be the teacher of Socrates, about the works of Plato, then it seems obvious to say that both of us are referring to the same individual, even though her only description of Plato was false. What ought to be said in this situation is not that she failed to refer, but that she failed to accurately attribute a description to the individual that she was referring to. This seems to suggest that descriptions, contrary to Russell and Frege, are not necessary in order for someone to refer successfully through the usage of a name, since my colleague was able to refer without the mediation of a description. Moreover, suppose that the description "the teacher of Aristotle," a description typically attributed to Plato, had been true of Socrates. Any sentence that I utter which has "the teacher of Aristotle" as its subject term would be referring to Socrates, since that is what the expression denotes in this counterfactual scenario. However, I would not be under the impression that Plato was not Socrates; I would simply be under the impression that Plato was not the teacher of Aristotle. Here is where Kripke's theory of rigid designation excels: it offers an explanation to the way in which we manage to successfully refer to an individual without the mediation of a description, and it provides one with the ability to differentiate the modal status of definite descriptions and proper names.

II. The Semantics of Modality

Before explicating Kripke's notion of a rigid designator, however, it is important to clarify a couple of terms relating to modality, since his theory has modal involvement and implications. It is first important to note that his definition of rigid and non-rigid designators depends on the concept of "possible worlds." A "possible world" can basically be understood as a way in which the world could be, or could have been. For instance, since it could have been the case that John Adams was the first President of the United States, not George Washington, then that is tantamount to saying that, in some possible world, John Adams is the first President of the United States. It should be made clear, however, that

talk of possible worlds does not ontologically commit oneself to believing in these worlds as real, since all that this talk of possible worlds is contributing to the discussion is providing a semantics on which we can rest our modal talk. Moreover, since modality deals with what is possible or necessary, a proposition of the form "*it is possible that x*" or "*it is necessary that x*" would be considered *modal* propositions. These are not the only kinds of modal propositions, however, since there are also counterfactual conditionals, which are similar to conditional propositions in form, yet differ in the sense that they discuss what is contrary to fact. An example of this would be the statement, "If Donald Trump had not won the 2016 United States presidential election, then Hillary Clinton would have," since, despite having the form of a conditional statement, it is contrary to the actual state of affairs.

Actuality is also another term which has some modal import. If we abide by Lewis' interpretation of actuality—that it is an indexical—then the actual world is just the world that we are a part of (Lewis 1973, p. 85-86). This allows us to analyze, for instance, the statement "George Washington was the first President of the United States" as true in the *actual* world, but false in some *possible* world in which John Adams was the first President of the United States. Actuality plays a substantial role in Chen Bo's criticism of Kripke's rigid designation thesis, which I will introduce later in this essay, since he attempts to undermine the modal difference between proper names and definite descriptions through qualifying descriptions as being about objects in the *actual* world, and as such always referring to these objects across all possible worlds.

Having laid out the semantics and philosophical import of the theory in question, it would now be appropriate to define Kripke's theory of rigid designation. For Kripke, a rigid designator is a term that refers to the same object in all possible worlds in which that object exists, and which refers to no object in all possible worlds in which that object does not exist.[1] A non-rigid designator, on the other hand, is a term that does not refer to the same object in all possible worlds. For instance, the expression

"the teacher of Aristotle" would be considered a non-rigid designator, since, although it refers to Plato in the actual world, it is possible that it could have referred to Cicero, if the state of affairs had been different. In other words, the expression does not refer to the same object in all possible worlds, since it refers to Plato in the actual world, but to Cicero in some other possible world. Definite descriptions, in general, are considered by Kripke to be non-rigid designators, since they could have easily been true of another individual in some counterfactual scenario.[2]

The singular term "Plato," however, does refer to the same object across all possible worlds in which the object exists because, whenever I entertain counterfactual situations in which "Plato" is the subject term, I hold the referent of "Plato" fixed. This referent-fixing facet is what characterizes proper names as rigid designators, since, without it, one would have to construct a "transworld" identity criterion for every proper name, which would clearly be an arduous task. Kripke describes this referent-fixing process by appealing to his causal theory of reference, which states that a proper name becomes rigid—has its reference fixed to a specific object—the moment someone initially baptizes the object with that name. Moreover, reference involving that proper name is successful whenever there is a causal chain that exists between the utterance of the proper name to the initial baptism of the object. Besides the potential metaphysical consequences that emerge from this theory, Kripke manages to create a theory of proper names that does not equivocate proper names with definite descriptions, while simultaneously solving the issue that descriptivist theories seemed to have trouble with.

III. Rigidified Descriptivism

Philosopher Chen Bo, however, does not take Kripke's modal argument to be a strong enough reason for adopting his theory of rigid designation. Roughly, Kripke's modal argument revolves around the difference in the modal profile between proper names and definite descriptions. To reiterate, Kripke understood there to be a semantic difference between proper names and definite

descriptions, in regards to modal propositions: proper names are rigid designators while definite descriptions are non-rigid designators. This goes against descriptivist theories of names, however, since descriptivists maintain that proper names and definite descriptions do not differ in regards to their modal profile. After all, if they did differ, then it would seem to undermine the connection between proper names and descriptions that descriptivism aims to demonstrate. Kripke clearly argues against this, however, by holding that there is a difference. Namely, that one has the property of rigidity while the other does not. So, if one assumes that descriptivist theories of proper names are true, then, by simply applying a reductio-style argument from the premises just laid out, the original assumption turns out to be false.[3]

Chen Bo criticizes Kripke's argument by maintaining, among other things, that descriptions can, "with an implicit qualification," become rigidified, and thus have the same modal profile as proper names (Bo 2012, p. 235). Bo evaluates his criticism by examining statements of the following variety:

(1) Plato might not have been the teacher of Aristotle

(2) The teacher of Aristotle might not have been the teacher of Aristotle

It is clear that (1) is true, since there is nothing stopping us from creating a counterfactual scenario in which Cicero was the teacher of Aristotle. However, (2) appears false, since it seems to be the case that the only thing that the teacher of Aristotle could not have not been was the referent of "the teacher of Aristotle"—whoever that may be. As a result, (1) and (2) do not have the same truth value. And since their truth values differ in a modal context, then their modal profiles are not the same. This lends itself to Kripke's modal argument for adopting rigid designation (Kripke, 1981, pg. 281-282).

Bo understands Kripke as being partially correct here, since he maintains that (1) and (2) do differ in truth value, but that they only differ in truth value if 'Plato' in (1) and 'The teacher of Aris-

totle' in (2) both take narrow scopes relative to the implicit modal operator—possibility—that is apparent in both statements.[4] If, instead, they both take wide scopes relative to the modal operator of the proposition that they are within, then (1) becomes:

(3) There is exactly one person who is Plato such that it is possible that *he* was not the teacher of Aristotle.

And (2) becomes:

(4) There is exactly one person who is the teacher of Aristotle such that it is possible that *he* was not the teacher of Aristotle.

Under these wide-scope readings of (1) and (2), then they do both turn out to be true. It would thus seem to be the case that Bo has, on some level, undermined Kripke's modal argument for rigid designation, since, on one reading of the modal operator, there is no difference in the modal profile between proper names and definite descriptions. However, since I am not primarily concerned with this aspect of Chen Bo's criticism of Kripke's thesis on rigid designation, I will not bother myself with providing an argument against his assumption that simply having a reading on which (1) and (2) both have the same truth values—a wide-scope reading—may cause one to doubt Kripke's modal argument. Chen Bo himself even recognizes that this assumption of his needs substantial support, since he gives a considerable amount of attention to defending the idea that propositions containing proper names, like definite descriptions, are also susceptible to scope ambiguities (Bo 2012, p. 239). Rather, it is what he continues to say after he puts forth this initial criticism that I want to examine.

According to Bo, one reason why someone may hold (1) to be true and (2) to be false is because, when we use names, we intend to refer to some unique individual, while, when we use definite descriptions, we intend to refer to whoever has the quality or property that the description is attributing to the individual. If this were the case, then it is clear that (1) and (2) do not have the same truth value. However, Bo goes on to claim that, contrary

to this line of reasoning, we do not use descriptions in this way. He claims that, when we use descriptions, we use them intending to refer to a particular individual whom we have already picked out as the referent of the description. For instance, when we use "the President of the United States" we intend to refer to Donald Trump, not to whoever might have fulfilled the quality of being the current president of the United States (Bo 2012, p. 235). Bo continues this line of reasoning by stating:

> From the referential perspective, descriptions can also be rigidified, and become "rigid designators", as long as they are supplemented with an implicit qualification: "in the actual world (which I shall rephrase by 'in @')". The reason why the qualification is not made explicitly is that we *always* talk in @, so there is no need to refer to @ all the time. When we talk counterfactually, we go outside the actual world and into some other possible worlds, and then "in @" should be added: "the teacher of [Aristotle] in @". In this way the description is no less rigid than the name 'Aristotle'; they both refer to the man *Aristotle* fixedly, invariably, rigidly. (Bo 2012, pp. 235-236)

Before I mention my thoughts on Bo's criticism of the difference between statements like (1) and (2), I would like to clarify the two claims that he has made thus far. The first claim seems to be a claim on the way that people use language; that is, it seems to be an empirical claim about what people mean when they use descriptions to refer to someone or something. The second claim, however, is not entirely empirical, since it is also making a claim on the modal quality of descriptions. Since I want to dedicate the rest of this essay in examining his second claim, I would like to quickly address his first claim and hopefully show that it at least does not assist in the veracity of his second claim.

To reiterate, Bo's first claim centers around the idea that, when people use definite descriptions as the subject of their statements, they do so intending to refer to some particular, unique individual. This is in contrast to the view that someone who

understood (1) as being true and (2) as being false would hold; namely, that, when people use definite descriptions as the subject of their statements, they do so intending to refer to whoever fulfills the quality that the description is attempting to attribute. I do not completely reject Bo's claim, since there are certainly moments in which someone uses a definite description intending to refer to some particular individual whom they already have in their mind. However, I do not agree that this is always the case, since there are certainly scenarios in which someone would use a description in order to refer to whoever happens to fulfill the quality or property that the description is attempting to attribute. For instance, let us say that, for some reason, I was unaware of who Thomas Edison was. If, one day, I was to marvel at the complexity of the light bulb that lights my desk, and if I were to then utter "the inventor of the light bulb was a genius", I would not have had any particular individual in mind, since I am merely referring to whomever was the inventor of the light bulb. Mid-20th century philosopher Keith Donnellan understood this facet of descriptions by conceiving of a distinction between a description's *referential* use and its *attributive* use (Donnellan 1966, p. 284-286). The example I just provided would have been an instance of the attributive use of the description "the inventor of the light bulb," while Bo's intuitions about descriptions align with a description's referential use. Moreover, since the use of a description is determined by the intention of the speaker, and the issue over how people intend to use descriptions is an empirical one in nature, Bo's first claim does not amount to a strong criticism of Kripke's thesis, since not all uses of descriptions are referential.

Bo's second claim, however, is more substantial than the first, since it is a claim that seems to attack the root of the problem; namely, the difference in the modal profile between proper names and definite descriptions. Bo claims that, by adding the implicit qualification "in the actual world," a definite description can be rigidified and thus have the same referential power that a proper name does. A rigidified description that refers to some object in the actual world, therefore, is no different, in regards to the object

that it refers to in other possible worlds, from a proper name that refers to the same object in the actual world. To better illustrate this claim, let's rewrite (2) as:

(5) The teacher of Aristotle *in the actual world* might not have been the teacher of Aristotle.

With the implicit qualification now added into the definite description, (5) now appears to be true; the teacher of Aristotle in the actual world is Plato, and it is clear that Plato, as evidenced by the truth of (2), might not have been the teacher of Aristotle. Moreover, since the truth of (2) is captured by (5), and since 'the teacher of Aristotle in the actual world' and 'Plato' both appear to refer to the same object in all possible worlds in which the object exists, then it seems that Chen Bo has succeeded in undermining Kripke's modal argument for the adoption of rigid designation. After all, if, in regards to their modal import, a rigidified description is no different from a proper name, then there is no reason to favor names as the sole, exclusive bearers of rigidity.

What avenues, if any, are available to the Kripkean that hopes to combat Bo's rigidified descriptivist stance and save the semantic difference between proper names and definite descriptions? There are three possible ways in which I believe one could respond. The first way one could respond would be to attack the addition of the "implicit qualification" of 'in the actual world' to definite descriptions. The argument could be made that, similar to how it is not implicit that every use of a description by someone is not implicitly referential, not every usage of a definite description is implicitly qualified to be about the actual world. The fact that the qualification is implicit could be criticized, since, if all descriptions were always implicitly qualified to be about the actual world, then counterfactual scenarios, with definite descriptions as their subject term, would never have been a problem for metaphysicians dealing with modality. Perhaps there are usages, probably more arcane than that of attributive uses of definite descriptions, in which a description, uttered in the actual world, is not implicitly qualified to be about the actual world. For the sake of argument,

however, I would like to grant Bo this point, and instead focus on what, if anything, could be done to maintain some difference between rigidified descriptions and proper names, if both are to be thought of as rigid designators.

IV. Obstinacy and Persistency

The second way in which rigidified descriptions and proper names can still be thought of as different, in regards to their modal consequences, is through appealing to the distinction between an obstinately rigid and a persistently rigid designator. Before offering the definition of obstinately rigid and persistently rigid designators, which is often credited to Nathan Salmon, let us reiterate the definition of a rigid designator and a non-rigid designator, as conceived by Kripke:

> Rigid Designator: a term *d* is a rigid designator for an object *i* iff it designates *i* with respect to all possible worlds in which *i* exists, and to no object in all possible worlds in which *i* does not exist.
>
> Non-rigid Designator: a term *d* is a non-rigid designator for an object *i* iff it does not designate *i* with respect to all possible worlds.

With these definitions clear in our minds, let us now introduce the definitions of an obstinately rigid designator and a persistently rigid designator, as conceived by Nathan Salmon:

> Obstinately Rigid Designator: A term *d* is an obstinately rigid designator for an object *i* iff it designates *i* with respect to every possible world, whether that thing exists there or not.
>
> Persistently Rigid Designator: A term *d* is a persistently rigid designator for an object *i* iff it designates *i* with respect to every possible world in which that thing exists, and which designates nothing with respect to possible worlds in which that thing does not exist. (Salmon 1982, p.

33-34)

It should be clear now how Chen Bo's rigidified descriptivist argument against Kripke's modal argument for the adoption of his rigid designation thesis could be amended: a rigidified description is a persistently rigid designator, while a proper name is an obstinately rigid designator. This would preserve a semantic difference between rigidified descriptions and proper names, which would support Kripke's modal argument and thus weaken descriptivism. The issue, however, now becomes whether these terms are even terms worth entertaining in the first place. After all, an obstinately rigid designator seems to entail the necessarily-existing-object consequence that Kripke, with his added clarification on the definition of a rigid designator, wanted to avoid, since he did not want his thesis on rigid designation to entail the necessary existence of rigidly designated objects. As a result, it appears to be that, in order to save the difference between rigidified descriptions and proper names, one would need to demonstrate that the distinction between an obstinately rigid designator and a persistently rigid designator is not superfluous, and that names are obstinately rigid designators while rigidified descriptions are persistently rigid designators.

First, it would be useful to examine exactly when the distinction between these two terms would even be pertinent to our current discussion. It appears to be that the only time an obstinately rigid-designating term would differ from a persistently rigid-designating term would be when we consider possible worlds in which the object designated by the term does not exist; if the term is persistently rigid, then it fails to refer, while if the term is obstinately rigid, then it manages to successfully refer. If we look to Kripke's definition of a rigid designator, it most likely resembles the definition of a persistently rigid designator, since they both make the clarification of guaranteeing a failure of reference at possible worlds in which the designated object does not exist. As a result, the defendant of Kripke's modal argument could begin her defense by first making the claim that rigidified descriptions and

proper names are—at least—persistently rigid designators. Once she has made this claim, she will then need to show how one could go about elevating proper names to the status of obstinately rigid designators. This second claim, however, will clearly be in more need of justification. I would now like to examine three possible ways in which she could argue this second claim.

It should quickly be mentioned that each of the three ways involves making use of David Kaplan's bi-dimensional framework of "context of use" and "world of evaluation" (Kaplan 1978, p. 494). Kaplan believes that an expression's content was a product of the expression along with the context at which it was uttered within. Once we had an expression's content, we can then evaluate it at different possible worlds in order to determine its extension at those worlds. For example, let's take the following sentence:

(6) Donald Trump is the President of the United States.

Under Kaplan's bi-dimensional framework, if we were to check the veracity of this statement at some possible world, then the world at which the content of 'Donald Trump' was generated at will be different from the world at which its extension is to be determined. This is because the term's context-of-use is the actual world, while the term's extension is some other possible world.

Having clarified Kaplan's bi-dimensional framework, I would like to examine the first possible route. This first route criticizes the idea that obstinately rigid designators refer to nonexistent objects. This preconception about obstinately rigid designators is, after all, one of the reasons why a defender of Kripkean rigid designation might be skeptical about believing in obstinacy. In order to defend this, Joao Branquinho, a proponent of the obstinacy thesis, capitalizes on the difference between "the notions of ¬*reference with respect to a world*¬ and *reference in a world"*, noting that "the former notion concerns the reference of words as used by us in describing certain counterfactual situations; the latter concerns the reference of words as used in those counterfactual situations" (Branquinho 2003, p. 5). Branquinho is clearly

trying to completely dissociate the two cases of reference from one another, since he takes the former instance of reference as being an instance in which the world of evaluation is the world at which the context is generated; that is, the actual world. He then takes the latter instance of reference as being an instance in which the world of evaluation is some other possible world. He then claims that the terms which we would consider obstinate, such as proper names, will have a reference in the actual world; since that is the world at which they were generated at, the objects reside in that world. Furthermore, even if the world at which the term is evaluated at does not have the object as part of its domain, it does not entail that the obstinately rigid-designating term will completely fail to refer. This is because, in situations in which the object does not exist at the target possible world, the obstinately rigid designator's reference can still be guaranteed through its existence at the world of generation. In other words, although the term does not have a reference *in* the possible world, the term has reference *with respect to* the possible world.

If we accept this line of reasoning, then we manage to reinforce our belief in Kripke's rigid designation thesis by reaffirming his modal argument, and we manage to do so without having to believe in the reference of nonexistent objects. However, there are two counter-arguments that I think can be levied against the proponent of this thesis. The first involves the distinction made between reference at a world of generation and reference at a world of evaluation. Although I have no issue with making the distinction between the worlds themselves, I do not subscribe to the idea that, when we take our rigidly designating term and entertain counterfactual situations in which the referent of the term does not exist at some possible world, we are to believe that its extension at the possible world—the world of evaluation—is not the sole determinate of its reference. If I want to know whether or not an obstinately rigid designator's reference is successful at a possible world, then I am interested in whether or not the object that the obstinately rigid designator is supposed to designate is a member of the set of objects in the possible world's domain.

Under the analysis thus far presented, however, the object's nonexistence at a possible world would not entail a failure of reference, since reference, as was hitherto mentioned, can still be guaranteed, albeit through a different form of reference. But this is simply counterintuitive, since it goes against what we take to be true, in regards to an object's existence and a term's reference to that object; namely, that if an object does not exist, then reference to that object cannot be successful. It should be noted that, simply because a theory is counterintuitive does not entail that it is false or should be rejected immediately, since it is oftentimes the case that our intuitions can lead us astray from the truth of the matter. However, given that this entailment has been taken as true by the vast number of individuals, descriptivists and anti-descriptivists alike, I would not accept this response to rigidified descriptivism until this counterintuitive notion is argued more thoroughly.

However, for the sake of argument, let us accept that this split between an object's existence and a term's reference to said object is demonstrated to be a tenable position; that is, let's accept that an object's nonexistence at some possible world does not entail a failure of reference for an obstinately rigid-designating term. There would still need to be a further argument that would demonstrate why proper names are the only privileged class of expressions to receive the treatment of obstinacy. This is because, insofar as Kaplan's bi-dimensional framework is concerned, there does not seem to be any reason as to why I should not consider rigidified descriptions as being obstinately rigid. Whenever I utter an expression of the form '*The actual F is G*' in the actual world, a content is generated at that world. I could then take that content, which was generated at the actual world, and evaluate it at some possible world in order to determine the expression's extension at that world. If the evaluation world is a possible world in which the object that the expression referred to in the actual world does not exist, then reference *in* the possible world will fail. However, the expression can still have reference *with respect to* the possible world, since reference is still guaranteed through the object's existence at the world of generation. So, unless it can be shown

that rigidified descriptions do not follow this same process, then alluding to obstinacy and persistency will not be of use to saving Kripke's modal argument, since both proper names and rigidified descriptions can be classified as the former.

V. Millianism and Obstinacy

There is a way, however, in which one can adopt the sole, exclusive obstinacy of proper name-rigid designators, which involves a more indirect route dealing with Millianism and the concept of direct reference. Millianism, named after 19th century philosopher John Stuart Mill, is a theory about the *content* of an expression. Moreover, it is a theory that states that the content of a proper name is just the individual that the name is referring to.[5] For instance, someone who adheres to a Millian interpretation of a content's expression will understand the content of, for instance, the name 'Donald Trump' to just be the individual that the name refers to; namely, Donald Trump. This is in contrast to an individual who adopts a descriptivist stance on proper names, since they would most likely understand the content of the name 'Donald Trump' to be some sort of description that is meant to uniquely refer to him.

Given this understanding of Millianism, how does it enter into the issue of whether or not a proper name is an obstinately rigid designator? As Brendan Murday notes in "Names and Obstinate Rigidity", one may invoke Kaplan's interpretation here on what it means to say that an expression is directly referential, "I intend to use 'directly referential' for an expression whose referent, once determined, is taken as fixed for all possible circumstances, i.e., is taken as being the propositional component" (Kaplan 1978, p. 493). This interpretation of direct reference is highly suggestive of obstinate rigidity, since, once the content of an expression is created and set, then it will remain as such in all other possible evaluations of the name. So, in regards to our discussion of proper names, because the content of a proper name just is the individual being referred to, then, in any counterfactual situation involving the name, the aforementioned object will be available for reference, due to the object itself being a propositional component.

It is still vague, however, how this process is to be understood, because, although it would provide a response to my second criticism, it does not appear to provide an adequate response to my first criticism. In other words, it manages to provide a coherent understanding of obstinacy, but it does not do much in regards to formulating a clear way in which we can think of reference occurring in possible worlds in which the object does not exist.

Kaplan, however, does at least provide a way in which we could think about this issue that is not as vague. Kaplan, in "Afterthoughts", states:

> If the individual is loaded into the proposition (to serve as the propositional component) before the proposition begins its round-the-worlds journey, it is hardly surprising that the proposition manages to find that same individual at all of its stops, even those in which the individual had no prior native presence. The proposition conducted no search for a native who meets propositional specifications; it simply "discovered" what it had carried in. (Kaplan 1989, p. 513)

Kaplan is aware of the need to demystify the way in which obstinacy is to be coherently understood, and does so through his idea that a proposition, upon being evaluated at a certain world or circumstance, merely discovers the propositional component that it had carried in. And since, given a Millian understanding of proper names, the propositional component of a name just is the individual that the name is to refer to, the individual will be available for successful reference. Thus, not only does this analysis make the obstinacy thesis for proper names plausible, it also, through the inclusion of Millianism, makes a case as to why obstinacy would be a property of proper names and not of rigidified descriptions.

Despite the coherence, at least at face value, of this theory, it is still far from criticism. The most peculiar feature of this interpretation is that of the "carrying" function of propositions. Reference is successful due to the individual being carried into the world of evaluation, as a result of being the propositional compo-

nent of the name being evaluated. But what does it mean to say that an individual is "carried into" a world of evaluation? If this "carrying into" function is to be granted, then it seems to suggest that the individual that was "loaded into" the proposition becomes, as a result of simply being the content of the name, a member of the objects at a given possible world's domain (Murday 2013, p. 231). And since it becomes an eligible candidate for reference within the world of evaluation, reference to that object within the possible world will be guaranteed, and the obstinacy thesis is seen as somewhat plausible.

But this entailment, as Murday notes, appears to turn on a misunderstanding of Kaplan's own bi-dimensional framework. If we review Kaplan's notions of content and extension, then it becomes clear that the former cannot entail the existence—or nonexistence—of the latter in any way (Murday 2013, p.232). To reiterate, a content of an expression is created from the expression itself and the context of the expression's use. Once the content of an expression has been established, then one can figure out its extension at a possible world by evaluating that expression at that possible world. However, just because the individual was available in the world of generation to serve as the content of, say, a name, does not entail that it will be available to serve as the extension of any counterfactual situation involving the name. This is because, if we acknowledge that the only objects that are eligible to serve as the extension of a name in a particular possible world are just the objects that exist at that possible world, then it should come across as highly suspicious that the individual referred to in the world of generation becomes an eligible object for reference at any possible world just because it is the propositional component of the expression. If it sounds incorrect to assert that, by simply being the content of the name, the individual referred to in the world of generation consequently becomes an eligible member of any possible world's domain, then this understanding of the "carrying in" function ought to be met with an incredulous attitude.

Given this peculiar facet of obstinacy, the only other way

that I can imagine someone approaching and adopting this issue is by somehow reifying the "carrying in" metaphor and asserting that it just is a property of expressions that have Millian contents. Although this reification may be more appealing to those individuals who would like to adopt the obstinacy of proper names, since it is at least a coherent stance that one could take, all it does is shift the discussion to instead be about whether or not we should believe in the "carrying in" function of these types of expressions. If we are fine with accepting this, then we at least provide a response to Chen Bo's criticism of Kripke's modal argument, and thus manage to save his rigid designation thesis. However, if we do not feel comfortable with this solution, then there is still one more alternative that I would like to consider.

VI. Millianism and Persistency

If the reasons thus considered do not seem sufficient for establishing the obstinacy thesis for proper name-rigid designators, then what stance should one adopt on this issue? Murday argues that we should accept proper names as being persistently rigid designators, similar to rigidified descriptions, but that we should take a Millian understanding of their content. Since the question of whether or not a name is obstinately or persistently rigid pertains to a name's extension, there is nothing contradictory about formulating a Millian interpretation towards a name's content because, as has been previously demonstrated, the content and extension of an expression are two distinct concepts.

What, however, are the motivations for accepting such a position? Well, most of the motivations are the rejections of the reasons for obstinacy that have been discussed thus far. After all, we started off the discussion of obstinacy and persistency by comfortably granting that the defender of obstinacy would have to show that proper names, not rigidified descriptions, are obstinately rigid designators, since we were comfortable with assimilating Kripke's notion of a rigid designator with Salmon's notion of a persistently rigid designator. And since we were comfortable with treating both proper names and rigidified descriptions as persis-

tently rigid designators, then the idea of holding proper names as being persistently rigid designators ought to appear reasonable. The idea of names being persistently rigid while having a Millian content, however, may not appear as plausible. How should we respond to Kaplan's claims, after all, about the apparent entailment of the obstinacy of proper names through the acceptance of the theory of direct reference?

Murday suggests that, all that the directly referential aspect of the content of proper names does is tell us how we are to pick out the referent of a name in a particular world of evaluation, not that the referent of the name somehow becomes a part of the domain of the world of evaluation (Murday 2013, p.233). To clarify, let us once again examine the proper name 'Donald Trump.' The content of the name, as has been mentioned before, just is the individual that it refers to; namely, Donald Trump. Now, this means that, when we evaluate this name at other possible worlds of evaluation, all we are doing is looking for the individual that just is Donald Trump. In other words, we are not committed to saying that the individual, simply by being the content of the name, gets carried into the world of evaluation. All that we *are* committed to, by our acceptance of Millianism, is that, when we evaluate a name at a particular possible world, what determines the extension of the name is the individual that just is Donald Trump at that world.

Although this demonstrates that one can hold proper names as having Millian contents without being obligated to believe in the obstinacy of proper names, how should we then understand counterfactual situations involving the negated existence of an individual? After all, the distinction between obstinacy and persistency only enters the discussion when we are examining a possible world in which the designated object does not exist. For instance, take the following propositions:

(7) Donald Trump does not exist.

(8) Donald Trump exists.

The defender of obstinacy could say that, (7) is true at a world in

which Donald Trump does not exist because the extension of the name, which in this case just is Donald Trump, has the property of nonexistence at that world. Moreover, the truth of (7) entails that (8) is false, since Donald Trump would have the property of nonexistence, not of existence. Salmon himself would agree with this analysis, since he would take this to be the best way that we could understand the truth of (7) at some possible world, given, after all, that names are obstinately rigid (Murday 2013, p. 236). Despite its being clear now how the obstinacy of proper names allows the individual to enter into the domain of any possible world, it is unclear just how an object can exemplify the property of nonexistence. Even though we want to be able to give an explanation for the truth of negative existentials, we probably do not want to be committed to the idea that an object can exemplify nonexistence, since that seems counterintuitive (Murday 2013, p. 236).

The defender of persistency, however, can affirm the truth of (7) without having to take recourse to this odd property of objects. First, she would say that (8) would be false at a possible world in which Donald Trump does not exist because there would simply be no extension at that world. Furthermore, because the falsity of (8) means that (7) is true, the defender of persistency manages to demonstrate the truth of (7) without taking recourse to the property of nonexistence. So it seems that, even in these instances, one can hold to the persistency of proper name-rigid designators and respond with the same explanatory power of obstinacy.

What, then, does this mean for Kripke's modal argument? After all, the use of the distinction of obstinately rigid and persistently rigid designators were introduced in order to support the modal argument and to rebuke descriptivism. There are three comments I would like to make here. First, if someone were to subscribe to one of the versions of obstinacy previously described, then Kripke's modal argument is not in danger. Obstinacy, especially together with Millianism, does provide one with the semantic difference needed in order to answer Chen Bo's criticism of the argument, albeit with some peculiar consequences. Secondly, we

granted Bo's premise that, when we use definite descriptions, we implicitly qualify them to be about the actual world. This, as I had pointed out earlier, is not an easily defensible premise, and, as such, can be targeted by anti-descriptivists in order to weaken rigidified descriptivism. And lastly, even if we assert that proper names are rigid in the same way as rigidified descriptions, we do not evaluate rigidified descriptions in other possible worlds in the same way that we evaluate proper names. Given the difference in content between a rigidified description and a proper name, we can still hold onto a difference between way in which their extensions are determined in other possible worlds. So, although this may not demonstrate a difference in the extensions between proper names and rigidified descriptions, it does demonstrate a difference in their contents.

VII. Conclusion

It should now be clear where we stand in regards to the possibility of finding a difference between rigidified descriptions and proper names. I began this essay by first examining Chen Bo's criticisms against Kripke's rigid designation thesis, and mentioned the possible avenues one could take in response to his criticisms. One of the routes—that of persistently rigid and obstinately rigid designators—seemed to show promise, since it provided a way in which one could neatly amend the modal argument in order to rebuke rigidified descriptivism. Alternative understandings, in regards to obstinacy, were discussed in order to entertain the plausibility of adopting the obstinacy/persistency distinction to save Kripke's modal argument. Although they at least managed to provide a coherent understanding of obstinately rigid designators, each had a peculiar facet to them that only seemed to push the issue onto a more troubling problem instead. Finally, we arrived at a more plausible difference between proper names and rigidified descriptions, albeit one that was different from the one we had originally tried to save. Although this difference was not the kind of difference that we had hoped to gain by infusing the obstinacy/persistency distinction into Kripke's modal argument, it is

still a form of difference nevertheless, and as such should be taken in stride by proponents of rigid designation, and should be taken seriously by proponents of descriptivism.

Notes

1. Simply maintaining that a rigid designator is a term that refers to the same object in all possible worlds would be to suggest that the object being referred to is a necessarily existing object. Since Kripke does not want to have this be a consequence of his theory, he makes the noted clarification.
2. I say "in general" since Kripke does recognize some definite descriptions as being rigid. The definite description "the sum of 2 and 4" will always refer to 6 in all possible worlds, for instance, since mathematical propositions are necessarily true.
3. A *reductio*-style argument involves making an assumption, and then demonstrating how that assumption leads to a contradiction. If the assumption leads to a contradiction, then the assumption cannot be true.
4. Different "scope" readings typically occur whenever it is ambiguous what the scope, or domain, of an operator or expression is within a sentence. An example of this would be the popular sentence "Everyone loves someone", since it is ambiguous as to what the scope of 'everyone' and 'someone' is within the sentence.
5. Mill, in his writings, used the terms 'connotation' and 'denotation', as opposed to 'content' and 'extension'. Throughout this essay, however, it is presupposed that what Mill meant by 'connotation' is captured with our usage here of 'content', and that what Mill meant by 'denotation' is also captured with our usage here of 'extension'. If it is somehow argued that that is not what Mill meant when he used the terms 'connotation' and 'denotation' in his writings, then that will not undermine the integrity of the analysis presented in the essay, since Millianism is merely being used here to characterize a directly referential attitude towards proper names.

Bibliography

Bo, Chen. (2012) "A Descriptivist Refutation of Kripke's Modal Argument and of Soames's Defence" *Theoria: A Swedish Journal of Philosophy*, vol. 78, no. 3, pp. 225-260

Donnellan, Keith. (1966) "Reference and Definite Descriptions" *The Philosophical Review*, vol. 75, no. 3, pp. 281-304

Frege, Gottlob. (1892) "On Sense and Reference" *Translations from the Philosophical Writings of Gottlob Frege*, pp. 36-56 (Oxford: Blackwell)

Joao Branquinho. (2003) "In Defense of Obstinacy" *Nous-Supplement:Philosophical Perspectives*, vol. 17, pp. 1-23

Kaplan, David. (1989) “Afterthoughts” *Themes from Kaplan*, pp. 565-614 (New York, NY: Oxford University Press)

Kaplan, David. (1978) “Demonstratives” *Themes from Kaplan*, pp. 481-564 (New York, NY: Oxford University Press)

Kripke, Saul. (1980) “Naming and Necessity”, pp. 279-294 (Cambridge: Harvard University Press)

Lewis, David. (1973) *Counterfactuals* (Oxford: Basil Blackwell)

Murday, Brendan. (2013) “Naming and Obstinate Rigidity” *Southern Journal of Philosophy*, vol. 5, no.2, pp. 224-242

Russell, Bertrand. (1905) “On Denoting” *Mind*, vol. 14, no. 56, pp. 479-493

Salmon, Nathan. (1982) *Reference and Essence* (Oxford: Basil Blackwell)

Searle, John. (1958) “Proper Names” *Mind*, vol. 67, No. 266, pp. 166-173

A Practical Approach to Free Will and Moral Responsibility

Johnathan Poh

Introduction

The free will debate as it currently stands is far from settled. However, recent advances in the fields of neuroscience suggest that our world is deterministic. If our world is deterministic, how does this affect free will? For incompatibilists such as Derk Pereboom, determinism would entail that we do not possess free will. Here lies the conundrum. Most of us, if not all of us, live as though we have ability to freely choose during certain situations. However, if we do not have free will, would we still be able to justify moral responsibility? Moreover, how would a deterministic worldview affect our psychology and first-person experience? Pereboom believes that there can be no moral responsibility. He suggests that we accept a worldview that does away with moral responsibility. Interestingly, Pereboom believes that most of what we are accustomed to can be salvaged even though accommodations must be made. Without moral responsibility, we will be able to eliminate moral retribution from society. As such, accepting a deterministic worldview will be a positive for society. In his article, "The Moral Psychology of Determinism", Jeremy Evans examines the consequences of accepting a deterministic worldview. Evans deviates slightly from Pereboom in that Evans believes a deterministic worldview will have both positive and negative implications on our moral psychology.

In response to Pereboom and Evans, I argue that determinism is not yet a foregone conclusion. Since we are epistemologically limited beings, we may never find a definitive answer to the free will debate. Instead, we should take a more pragmatic approach to the free will debate. Section I will outline the free will debate and

the resulting implications. In Section II, I will outline Pereboom's argument for why we should adopt a worldview that precludes moral responsibility. Section III will cover the weaknesses in Pereboom's argument. Section IV will introduce Evans' view on the consequences of a deterministic worldview. Section V will relate Evans' views with Pereboom's views. Some of Evans' views reinforce Pereboom's claims while others highlight problems Pereboom does not adequately address. Section VI will articulate a more practical approach to the free will debate. The approach I suggest aims to preserve our status quo regarding our conception of free will as much as possible. However, my approach does not completely dismiss the consequences of determinism for some consequences of determinism can be rather helpful.

I. The Free Will Debate

A deterministic conception of the world is one on which all events are caused by prior events and conditions together with the laws of nature. Furthermore, no other sequence of events could have occurred (Hoefer 2016). This also applies to human psychology including our thoughts, actions, and desires. The opposite of determinism is indeterminism. Indeterminism allows for scenarios where an action is not wholly determined by prior events or causes. Determinism has a profound effect on how we think about free will. When we think about free will, in the way the common person thinks about it, we think of being able to make free decisions between options. When I talk about making a decision, I am referring to conscious decision making. But if determinism is true, there can be only one actualized sequence of events and no other possibilities. So, our belief that some of our conscious decisions are free might be no more than a mere illusion. Those who believe determinism and free will cannot both be true are labeled as incompatibilists. Interestingly, however, most philosophers subscribe to the view that determinism and free will can co-exist with one another. These philosophers are often called compatibilists or soft determinists. Among the incompatibilists, there exists another divide. There are those who believe determinism is true

and free will cannot exist. These people are who we shall call hard determinists. On the other end of the incompatibilist spectrum are the libertarians. Libertarians believe free will exists and determinism is false. All in all, the debate regarding how determinism and free will interact with one another is strongly tied to moral responsibility. Responsibility entails the ability to choose. Not every decision needs to be free, but we need to be able to make choices in order to be held morally responsible or accountable for an action. Otherwise, there would be no justification for a person to be praised or blamed for some action he or she commits. Many ethical systems depend upon the existence of alternate possibilities. In particular, our conception of deontological or deontic ethics is a prime example of an ethical system that depends on alternate possibilities. Deontic moral theories revolve around rules and duties required of a person or agent (Hoefer 2016). Often times, an agent receives praise or blame for his or her action or inaction based on what they "ought" to do. If an agent "ought" to do something, then this implies the agent "can" do something. This is the ought-implies-can principle made famous by Immanuel Kant (Kant 1785). Under deontic moral reasoning, if the agent receives blame, then it is because the agent ought to have done something different. But the agent only could do so if there was an alternate possibility that allowed the agent to do otherwise. Thus, determinism seemingly undermines the very core of this widely supported ethical orientation.

II. Pereboom's Argument for a World Without Moral Responsibility

In Derk Pereboom's book, *Living Without Free Will*, he comes to the conclusion that we do not possess free will. Thus, there is no justification for moral responsibility. Furthermore, we should accept the notion that there exists no moral responsibility and adjust our worldview accordingly. We cannot and should not attribute moral responsibility to human agents (Pereboom 2001). Pereboom refers to himself as a hard incompatibilist but for the

most part is a good representative of the hard determinist viewpoint. The difference between a hard incompatibilist and hard determinist is that a hard incompatibilist (i.e. Pereboom) believes an indeterministic world aside from one of libertarian, agent-causation would also entail no moral responsibility. For a hard determinist, free will and determinism are incompatible with one another. Hard determinists believe determinism is true. Thus, free will does not exist. Furthermore, moral responsibility requires free will, and thus there is no moral responsibility.

In his book, Pereboom reaches his conclusion by eliminating other competing theories in the free will debate. In theory, Pereboom believes a libertarian, agent-causation view of free will could provide the free will required for moral responsibility. But Pereboom dismisses the agent-causation view on the basis of a lack of evidence for its existence. He believes it to be highly unlikely for agent-causation to exist based on empirical evidence and our best physical theories. Pereboom does not accept compatibilist views. Compatibilist views have two weaknesses that are often criticized. One is that determinism entails that there are no alternate possibilities. Leeway incompatibilism rejects compatibilist views on the grounds that moral responsibility requires alternate possibilities. Pereboom believes that moral responsibility does not require alternate possibilities and thus does not reject compatibilist views on these grounds. The second weakness is that compatibilist views do not provide the control necessary for human agents to be held morally responsible. Source incompatibilism rejects compatibilist views on the grounds that the agents in compatibilist theories do not have enough control over their actions to warrant the free will required for moral responsibility. Pereboom rejects compatibilist theories based on source incompatibilism as opposed to leeway incompatibilism. Though Pereboom believes moral responsibility is unjustified, he does not believe accepting this notion would undermine morality or meaning in our lives. He prescribes that we live our lives as if there is no free will and relinquish our notion of moral responsibility. In some ways, Pereboom believes the absence of moral responsibility could be a benefit to society. Pere-

boom maintains that the only significant thing that would change is that the absence of moral responsibility would preclude retributive punishment. Retributive punishment would be unjustified. He believes this would be a positive consequence. There would be less anger directed towards those who commit actions outside their control such as drug addicts and a push for theories advocating correction of behavior, such as the moral education theory of punishment. Moral education theory would involve rehabilitating a criminal to refrain from committing similar actions in the future. Criminals would be quarantined in a way akin to the quarantine of biologically diseased individuals with dangerous contagious diseases. A deterministic outlook on life, however, would leave intact rational deliberation, morality, social attitudes concerning interpersonal interactions, and approbation of virtues.

III. Weaknesses in Pereboom's Argument To Eliminate Moral Responsibiity

Pereboom's prescription for us to accept a worldview without free will and eliminate moral responsibility is far too quick. Though Pereboom may be correct in that we do not possess free will, his argument is not strong enough to warrant such a drastic change from our status quo. Our current status quo is that we do possess free will in certain situations. We have instances where we are able to freely choose. Our concept of free will is heavily interwoven into our lives on both a personal level as well as a societal level.

Let us borrow a concept from statistical hypothesis testing. A null hypothesis is statement that is reflective of the status quo. A null hypothesis claims that there is no relationship between two phenomena. Then, there is the alternative hypothesis, which claims that there is a relationship between the two phenomena. In inferential statistics, evidence is gathered to disprove the null hypothesis. The idea is that we are to give ourselves enough reason, usually via empirical evidence, to accept that there is a relationship between two phenomena. Otherwise, we remain with our status quo that there is no relationship. In most scientific

experiments, the conclusion that is hoped to be true is the alternative hypothesis. If there is enough evidence, the null hypothesis is rejected and we conclude that the alternative hypothesis is true. In hypothesis testing, there are also cases of statistical error: type I error and type II error. Type I error is when the null hypothesis is rejected when it is in fact true. Type I error is a false positive. Type II error is when the null hypothesis is not rejected when it is in fact false. Type II error is a false negative. With all this in mind, let us use this statistical framework and combine it with Pereboom's suggestion that we accept a worldview that is absent of free will. Bear in mind, Pereboom's argument does not fit exactly into a hypothesis test. I am using the framework simply as a way to better illustrate a point. For the purposes of this example, I will adjust what qualifies as a null hypothesis as well as an alternative hypothesis. Let us adjust our null hypothesis to simply be what our status quo is: we possess free will. Next, let us have our alternative hypothesis be Pereboom's suggestion: we do not possess free will and there is no moral responsibility. First, the onus is on Pereboom to offer enough evidence for us to reject our null hypothesis and conclude the alternative hypothesis. Furthermore, let us discuss type I error vs type II error. In type I error, we would reject our notion of free will and accept a worldview that precludes moral responsibility. However, the null hypothesis is actually true. This would then be a false positive. We would accept a worldview in which free will and moral responsibility are absent when in actuality free will exists. In type II error, we would not reject our notion of free will. But in this case the null hypothesis is actually false. This would then be a false negative. We would accept a worldview in which free will and moral responsibility exist when they in fact do not. Since we cannot reliably rule out agent-causation completely, there is a possibility that we could make a type I error, a false positive. So, we would live as though there is no free will and no moral responsibility even though we do have free will. This scenario, in my opinion, is much worse than believing in free will and moral responsibility when in fact they do not exist. We should do as much as we can to prevent

making a type I error. An argument to change our status quo and accept free will needs to be very strong. The consequences of the drastic attitude and cultural shift are too great to be for naught. This type of theoretical safeguard is used in other fields as well. For example, our court system hedges against convicting an innocent person. That is why it requires that every single member of a jury find a defendant guilty instead of half or eleven out of twelve jurors. It increases the chance that a guilty person will go free, but our justice system believes that the trade-off is worth preventing an innocent person being wrongly sent to jail. With regard to free will, especially our acceptance of free will, I advocate the same attitude. Our epistemic access to the metaphysical truth of free will is limited. Determinism could be true but we may never be sure that it is. In these cases, it is better to refrain from accepting free will in part to avoid a situation where we mistakenly reject free will when it actually does exist.

Pereboom's dismissal of agent causation as a genuine possibility is too quick. His argument is that agent-causation could allow for free will and moral responsibility, but agent-causation is highly unlikely based on our current physical theories. The agent-causal view involves the agent being the origin or source of his acts. He can be influenced by his desires and beliefs but there must be an influence originating from the agent that is not determined by prior events. Pereboom makes the case that the agent-causal view has not been supported by evidence and other evidence supports the contrary (Pereboom 2001). However, there are weaknesses in empirical data that must be considered.

Science is epistemologically limited. However, we treat certain scientific conclusions as certainties for pragmatic purposes, such as the notion that gravity exists throughout the universe. Science is a form of inductive reasoning. It relies on generalizing about a class of properties based on observations of a small subset of objects within the class. Moreover, the method presupposes that sequences of events will always occur in the future as they did in the past across all space and time (i.e. Hume's uniformity of nature principle). Science also relies on falsifiability and test-

ability. Furthermore, not all sciences are created equal. In general, the field of physics is much more reliable in accuracy and predictive capability than a field such as biology. Neuroscience, a subdivision of biology, is relatively nascent and still fraught with uncertainty. In science, it is assumed that the universe is materialistic and that there are laws of nature that govern it. In physics, we have been able to make empirical observations, extract laws from the data, and test the laws through experimentation. Another indicator of scientific prowess is predictive capability. For example, Albert Einstein's Theory of Special Relativity was widely accepted before it was confirmed during the total solar eclipse of May 29th, 1919. Afterwards, he became famous worldwide and his theory displaced the Newtonian Laws of Gravity. Neuroscience is not nearly at the level of predictive capability as physics is. For many experiments, there are still issues of whether the data correspond enough to indicate correlations and whether the correlations are causal. It is a common assumption among neuroscientists that the mind and mental experiences can be completely explained by neural states, structures, and their functions. The issue is that there could be something that allows for agent causation but is difficult or impossible to empirically evaluate. This, according to Gardar Arnason, is a big part of the epistemological-reductionist challenge (Arnason 2011). There is a possibility that neuroscientists are erroneously reducing our mental capabilities to only those that can be empirically tested in a lab. Furthermore, most cognitive studies are done under the assumption that that our minds operate in a deterministic way. Studies are done with the intent of discovering the underlying laws governing brain activity. This makes it even more unlikely that an empirical study will show us to have free will or give evidence for agent-causation. Most studies revolve around explanations. Empirical data are collected, and the interpretation of the data is key. There is a distinction between correlation and assumed causation that often hinges on how the data is interpreted. There is also the empirical-cognitive challenge. The challenge is that all decisions are made unconsciously and forwarded to the conscious mind in sort of a rationalizing manner (Arnason 2011).

This would be fatal for agent-causation. However, this event is extremely difficult to prove. Studies suggesting this tend to gravitate towards correlation and give weak causal claims. The field of neuroscience is not nearly advanced enough to predict outcomes with the accuracy and scope of physics. The experimental designs and the explanations of the empirical data can have a great effect on how much we should trust certain scientific conclusions. In general, the strength of causal connections made in neuroscientific studies is much weaker than those of physics. Also, we should be aware of confusing correlation with causation. That said, Pereboom is a bit quick to dismiss agent-causation based on a lack of empirical support. Agent-causation may seem implausible in virtue of recent research, but agent-causation is not necessarily defeated by these insights.

Pereboom's view would undermine one of our most prominent and useful systems of ethics—deontological ethics. Pereboom believes that our conception of moral worth and ethics would not be diminished as some fear. Pereboom suggests we would still have consequentialist based ethics (Pereboom 2001). Under Pereboom's suggested worldview that precludes moral responsibility, we can still have aspects of morality similar to those of virtue-based ethics. For example, people can receive praise or blame based for their actions. However, I disagree with Pereboom. Deontological moral reasoning is very useful because it relies on obligations and duties. It also happens to be one of our most prominent prescriptive moral theories. In comparison to other ethical systems, deontology is better at accounting for motivations and intentions. Another benefit of deontological theories is that they promote contemplation and deliberation before making choices. By accepting Pereboom's view, we would be sacrificing a lot more than what Pereboom suggests.

Pereboom believes that the conditions for moral responsibility are not satisfied by compatibilist theories. He rejects compatibilism on the grounds of "source incompatibility". For Pereboom, "An action is free in the sense required for moral responsibility only if the decision to perform it is not an alien-

deterministic event, nor a truly random event, nor a partially random event" (Pereboom 2001, pp. 54). Since all of our actions under determinism would be influenced in some degree by prior events, there is no moral responsibility. However, the issue with Pereboom's conception of moral responsibility is that the only thing that can garner blame is the original cause of the universe (e.g. the big bang). Conceptually, moral responsibility is an infinite regress to something prior. It seems to me that this standard is far too high and may be a bit absurd. A compatibilist, David Hume, simply says this:

> The whole dispute, in this respect also, has been hitherto merely verbal. For what is meant by liberty, when applied to voluntary actions?... By liberty, then, we can only mean a power of acting or not acting, according to the determinations of the will; that is, if we choose to remain at rest, we may; if we choose to move, we also may. Now this hypothetical liberty is universally allowed to belong to every one, who is not a prisoner and in chains. Here then is no subject of dispute. (Hume 1748, pp. 63)

Hume believes that free will is essentially the ability to choose in accordance with our desires and to act in accordance with our choices. From my standpoint, Hume's definition of free will is enough to justify moral responsibility. Pereboom's definition is much too stringent. Simply because we cannot satisfy Pereboom's definition of the free will that is required for moral responsibility does not mean we should rid ourselves of moral responsibility. Moreover, moral responsibility is more or less a tool we use as justification for ethical purposes. By using this Humean version of free will, a notion of controlling one's desires, we can demarcate a cause that avoids an infinite regress. Pereboom would likely say that the infinite regress is reason to dispose of retributive punishment altogether, which is what he advocates. The Humean version of free will is a bit different than the free will we believe we have and want to have. But it is enough to justify our current actions. I would say it would be more pragmatic to keep believing in free

will even if our notion of free will is slightly different from what we are used to envisioning.

IV. Evans' Deterministic Conception of Human Action

According to Jeremy Evans in "The Moral Psychology of Determinism", "the central question of the reactive attitudes debate is whether the DCA will exacerbate or mollify … the reactive attitudes central to social life" (Evans 2013, pp. 644). Recent research in cognitive science has suggested our decisions are heavily if not completely influenced by deterministic factors outside of our control. But if we accept that our world is deterministic, does it mean the end of morality and pro-social attitudes? As Strawson puts it, the free will debate is mostly a reactive attitudes debate (Strawson 1962). Strawson accepts the metaphysical reality of determinism. However, he believes the truth of determinism would not stop people from expressing their inter-personal attitudes (i.e. reactive attitudes) such as feeling moral responsibility, blame and approbation, guilt, pride, resentment, gratitude, etc. These attitudes are not merely ways of regulating behavior, but an expression of our moral attitudes. The replacement of what Strawson termed "participant" attitudes with "objective" attitudes would deny or rationalize all human behavior in virtue of the behavior being determined. This would undermine ordinary human relationships. For Evans, the DCA accepts determinism as true and rejects ultimate control. However, incompatibilist views can still be accepted as well as hard determinist views. Under the DCA, none of our actions are agent-caused (i.e. none of our actions originate from our minds without any prior causes outside our control). Although the DCA has negative implications for our reactive attitudes towards things such as morals and meanings, it helps give a causal explanation of those afflicted with mental diseases such as addiction and schizophrenia. Moreover, we may be less retributive with respect to certain deviant behaviors if we understand that the behaviors are outside of a person's control.

This might lead to less retribution and more rehabilitation. A deterministic worldview weighs heavily on our outlook on life as well as on our motivation. The DCA might lead us towards more emotional apathy. Evans suggests the DCA undermines our illusion of ultimate control. We may have "an unconscious trigger for rationalizing self-serving behavior and avoiding manifesting empathic concern" (Evans 2013, pp. 11). Self-control requires exertion. In moral psychology, there is an ever-growing belief that there is a human tendency to search out ways to rationalize and excuse behavior in order to pursue our self-interests. Self-restraint takes energy. Evans suggests there is a "tendency for individuals to engage in strategic ignorance and self-deception in order to actively avoid situations where altruistic behavior might be required" (Evans 2013, pp. 653). Evans cautions that we be aware of the metaphysical and epistemological distinction of a deterministic worldview. Fixed futures do not mean fixed natures. "A common mistake is to think that determinism implies not only that our futures are fixed, but that our characters are fixed in their current state" (Evans 2013, pp. 657). Events from the past do not necessarily have to repeat themselves in the future even in a world that is deterministic. Moreover, a deterministic worldview does not mean we are omniscient. We are epistemologically limited beings and thus we may never know what the exact sequence of events are before they happen. Overall, Evans concludes that the realization of a deterministic world has and will continue to have an impact on our psychological perception of the world. For many of us, we hold on to the illusion that we possess ultimate control. Evans believes the DCA will obviate this illusion and as a result have a profound effect on our moral psychology.

V. Evans and Pereboom

Evans' paper highlights many aspects of a deterministic worldview that Pereboom neglects. Similar to Pereboom, Evans believes the acceptance of a deterministic world will benefit society in a positive manner. It will diminish moral retribution. However, Evans highlights negative aspects in a graver manner than Pereboom

does. Both agree that prima facie (i.e. at first glance) misconceptions of free will should be avoided. We should avoid fatalistic attitudes towards our futures simply because of a deterministic worldview. Evans differs from Pereboom in that he suggests this fatalistic attitude could become a big problem. Moreover, this attitude is a part of the larger psychological impact of a deterministic worldview. The DCA, advocated by Evans, suggests an increase in moral apathy and a decrease in motivation to help others. These potentially psychologically damaging consequences are good reason to reconsider accepting Pereboom's suggestion to accept a deterministic worldview that precludes moral responsibility.

VI.

I suggest we take a more pragmatic approach to the free will debate. As I have discussed throughout the paper, the hard determinist view, as articulated by Pereboom, has its weaknesses. Moreover, the consequences of accepting a deterministic worldview are quite severe. The hard determinists have not made strong enough arguments to warrant eliminating moral responsibility from our collective first-person worldview and experience. We are epistemologically limited beings. So, we may never have a definitive answer to the free will debate. Instead of focusing on whether determinism, free will, compatibilism, etc. are true, it may be better to focus on what we do know and can know. This approach is more pragmatic in the sense that it will allow for a better opportunity to enhance our everyday lives and worldviews.

In theory, the significant benefit of a deterministic world is its ability to predict the future. In principle, we would have full knowledge of every single future event based on the laws of nature and prior events. However, we are epistemologically limited beings. Even if determinism is true, it does not mean we will possess omniscience. We are currently unable and very unlikely in the near future to acquire omniscience. In the case of the DCA, I propose we take a more nuanced approach that is inspired by Evans' position. It would be unwise to wholly reject determinism on the sole basis of our wishing to maintain a belief of ultimate

control. This would deny us the opportunity to use the predictive aspects of determinism available to us. For example, the concept of determinism has allowed us to discover laws, make predictions, and progress technologically. At the same time, fully accepting the notion of determinism is an extreme psychological shift that risks a great deal of negative psychological impact. A misconception of determinism can lead to a fatalistic attitude. Moreover, even a non-misconstrued view such as Pereboom's would lead to consequences such as the rejection of moral responsibility. A rejection of moral responsibility would undermine some of our best ethical theories such as deontic moral reasoning as well as decrease the efficacy of our deliberations. Justification in our day-to-day lives would be unnecessarily harder without using concepts such as possibilities and choice. My position is that we accept deterministic outcomes when absolutely necessary but in all other cases, we should act as though we have free will. For if we cannot predict what will happen, it leaves us in the same position regardless of whether the world is deterministic or indeterministic. If we knew the exact sequence of events, then we would be unable to change anything outside of that sequence. However, we do not always know the exact sequence of events, so that particular condition is not invoked. Knowing our futures does affect our phenomenology, but it is dependent on the scope and clarity of the predictions. When we can accurately predict that a particular outcome will happen within a certain standard of accuracy (e.g. the predictive accuracy found in physics), only then should we adopt a deterministic view for that particular circumstance. Otherwise, we should assume we have free will.

Our first-person experience of free will may simply be how we perceive the world. Take green grass for example. We see the color green. But scientifically, the grass is anything but green. The grass itself is absorbing all light that is not green. Then the grass reflects green light into our retinas. We are aware of this but for general practicality, we say the grass in green. This is analogous to my suggestion of believing we have free will. It could be the case that our world is deterministic. However, we can believe and

use the notion of free will because it is simply more practical to do so. As long as we are aware of determinism, it should be fine. Our world runs much more smoothly as it does than it would if we took the approach of the hard determinists and eliminated things such as moral responsibility.

A better question moving forward would be, for what circumstances we have predictive capabilities and in what instances we can continue to believe we have free will. My approach errs on the side of caution while also not eliminating the benefits of a deterministic world. Not only do we avoid committing the error of rejecting our standard notion of free will in the case that free will does exist, but also this approach allows for a gradual acceptance of a deterministic worldview in the case that free will does not exist. However, it seems unlikely that we will ever achieve the omniscience necessary to achieve a completely deterministic worldview. Thus, pragmatically speaking, free will should continue to exist for the foreseeable future.

Bibliography

Arnason, Gardar. (2011) "Neuroscience, Free Will and Moral Responsibility" *Trames* vol. 15, no. 2, pp. 147-155

Evans, Jeremy. (2013) "The Moral Psychology of Determinism" *Philosophical Psychology* vol. 26, no. 5, pp. 639-661

Hoefer, Carl. (2016) "Causal Determinism" *Stanford Encyclopedia of Philosophy* Retrieved on April 21st, 2017 from https://plato.stanford.edu/entries/determinism-causal/

Hume, David. (1748) *An Enquiry Concerning Human Understanding* (Indianapolis: Hackett Publishing)

Kant, Immanuel. (1785) *Groundwork of the Metaphysics of Morals* (New York: Cambridge University Press)

Pereboom, Derk. (2001) *Living Without Free Will* (Cambridge: Cambridge University Press)

Strawson, P.F. (1962) "Freedom and resentment" *Proceedings of the British Academy* vol. 48, pp. 1-25

Down the Rabbit Hole of Consciousness: All the Way Down, All the Way Up, or Panpsychism, Anthropomorphism, Aporia, and Quietude

Eduardo Salazar

Introduction

The notion of *psyche* has been a salient philosophical concept since the ancient Greeks. It generally referred to mind, breath, or soul (Skrbina 2009, p. 3). Various disciplines now explore the notion of consciousness and set out scientific means of inquiry. Nevertheless, two millenniums removed from the Greeks, we still seem to lack an adequate theory of consciousness. The reemergence of panpsychism in philosophy (Skrbina 2009, pp. xii-xiv) further problematizes contemporary theories of consciousness, beyond just mere novelty (Skrbina 2009, pp. xii-xiv). Succinctly stated, panpsychism is the position or worldview which claims that things, both animate and inanimate, have consciousness, mentality, or mind (Chalmers 2006, pp. 432-433). However, the literature on panpsychism reflects variations regarding the specifics of what mentality means and how mentality is a property or quality of things (Skrbina 2009, pp. 1-29). For our purposes, we will discuss the work of Galen Strawson and Graham Harman to make sense of panpsychism; reference Jacques Derrida to further problematize the issue; and conclude with the notion of the *quietude* of consciousness.

First, in "Real Monism: why physicalism entails panpsychism," Strawson argues that the physicalist's (or materialist's) negation of the experiential or phenomenal aspect of reality as not being physical is not only incorrect but denies the fundamental

reality of consciousness. Strawson concludes that any coherent real physicalist theory will not only entertain panpsychism, but ultimately weave it into physicalist theory. Second, we will discuss Harman's "Zero-person and the psyche" (2009) article on panpsychism and his critique of reductive-descriptive panpsychism, which reflects Strawson's position. Harman presents three arguments that build on one another that lead us to positing a panpsychist worldview. Although I argue that Harman presents a more nuanced view of consciousness in terms of panpsychism than Strawson, Harman's endorsement of *endopsychism* denies full blown panpsychism, stating that while "...all entities *contain* experience, not all entities *have* experience" (Harman 2009, pp. 282). I ultimately conclude that the merits of both positions lie in their taking us beyond *physicsalism*, but critique their anthropomorphism and idealism about consciousness. I use Derrida's notion of *aporia* to help illustrate their shortcomings and, consequently, develop the notion of the quietude of consciousness. In Derridian fashion, we are left with a paradox.

Background

Before fully engaging in panpsychist notions of consciousness, we will briefly explore why panpsychism has reemerged in philosophy. Traditional positions in the philosophy of mind, such as behaviorism, emergentism, and physicalism/materialism, have all fallen short of providing a substantive theory of mind (Alter and Nagasawa 2015, p. 1). Physicalism still holds significant merit, however, due to its close relationship with science, physics, and neurophysiology. In fact, physicalism seems to reflect our general assumptions about reality and consciousness: that the physical is the ground of reality and that our physical brain houses consciousness. Though we may hold immaterial religious beliefs close to heart, Nietzsche's diagnosis of the death of god (Nietzsche 1974, pp. 181-182) seems also to have resulted in our physicalist assumptions. So, what is physicalism exactly?

Briefly, physicalism is the metaphysical thesis that everything in the world is physical (Stoljar 2016). The world is composed of

physical objects working under natural laws and scientists help us understand the nature of objects through observation, experimentation, theories, and application/technology. There is little room for the non-physical. Even psychological immaterial things reduce to the physical or are dependent on the physical. Again, this seemingly neutral matter of fact position seems to pervade our assumptions about the world, even if we hold immaterial beliefs.

I refer to Alyssa Nye's physicalism as a paradigmatic position on contemporary physicalism because she presents a compelling argument to accept the view that physics can in fact adequately describe the nature of objects without recourse to speculative metaphysics. Also, her position values the law of parsimony by not introducing new entities in explaining physical objects and reality. Interestingly, her position says little about consciousness itself.

Responding to Bertrand Russell's critique of physics' inadequacy to explain the intrinsic nature of objects, and thus leaving an explanatory gap, Nye contends that the universe can, in fact, be solely explained by physics. That is, physics' power to describe, analyze and predict events is sufficient to understand the nature of the universe (Nye 2006, p. 362). Nye's philosophical insight is in positing that knowledge of the intrinsic nature of objects does not require experience, and that physics can provide a viable account of intrinsic properties in the form of theoretical models, namely mathematical formulas (Nye 2015, p. 355). She states, "We understand what the individual masses and charges are like in themselves in terms of what are the proper sort of mathematical objects that may be used to represent them" (p. 362). Mass, for example, is unique by the type of property it is as a mathematical representation. Consequently, we do not require a full-blown theory of consciousness beyond our theoretical formulations and representation of objects to explain the intrinsic nature of objects.

Strawson's Realist Physicalism and Panpsychism

Galen Strawson offers us a radicalized version of physicalism. Indeed, he calls his version *realistic physicalism* because it fully engages with the problem of existing in a physical world as a physical (human) object, yet also having supposed immaterial properties, namely consciousness or non-physical mental phenomena (Strawson 2009, pp. 33-34).[1] He offers a position that incorporates consciousness into its metaphysics. He writes, "I take physicalism to be the view that every real, concrete phenomenon in the universe is physical" (Ibid, p. 33). More specifically:

> What is it to be a realistic physicalist, or, more simply, a real physicalist? … You're certainly not a realistic physicalist, you're not a real physicalist, if you deny the existence of the phenomenon whose existence is more certain than the existence of anything else: experience, 'consciousness', … explicit conscious thought as we have it and know it at almost every waking moment. (Ibid, p. 33)

Here, we see Strawson challenging the orthodoxy of physicalism by touting the notion of experience at the heart of not only theoretical means to knowledge, but experience as a more fundamental starting point of explaining how things exist in the physical world. The notion of experience, then, becomes tantamount in Strawson's realistic physicalism, which reorients how we describe objects or phenomena in the physical world. So, how does incorporating the reality of experience in the physical world present to us a better understanding of the world?

Below I outline Strawson's panpsychist realistic physicalism followed by justificatory commentary:

1. Monism (one physical substance) and ontological plurality of objects is true.
2. Any real concrete phenomena in the universe is physical.
3. All mental goings-on are concrete mental/experiential

phenomena.

4. Consciousness is a necessary occurant (non-dispositional) phenomena.
5. If (1) – (4) are true, then all forms of physicalism must accept mental/experiential phenomena as physical phenomena.
6. Physics and neurophysiology do not exhaust the physical.
7. Either consciousness is present from the start or it is not present.
8. Emergentism is false.
9. If emergentism is false and (5) – (8) are true, then realistic physicalism is true.
10. Thus, realistic physicalism entails panpsychism.

We will now discuss each premise that leads us to conclude that realistic physicalism entails panpsychism.

Premise (1) serves as a starting assumption: monism is true. Strawson writes, "It [physicalism] is a view about the actual [monistic] universe, and I am going to assume that it is true" (p. 33). Whatever the world is and manifests itself as, it is of one substance or kind, physical. As philosophers, it is anathema to make gross assumptions and/or propose self-evident truths, but I think here our inclinations and intuitions may allow for such laxity. What there is is the physical, and we will learn that the physical encompasses much more than what standard physicalism posits.[2]

Premise (2) takes monism as manifesting into ontological pluralism. That is, although there is only one substance, there are many objects made of this substance. But let us be more specific. First, concrete refers to an object being spatio-temporally located; second, phenomenon refers to any particular existent, both a rock and an idea, for example. I think most forms of physicalism would likely see nothing controversial about either premise (1) or (2).

Premise (3) is where we first note Strawson's argument

drawing away from other forms of physicalism. He redefines and expands the notion of *phenomenon* to characterize what exists, namely, the physical. Humans have mental/experiential phenomena. Mental/experiential phenomena, then, must be necessarily real and true qua physical phenomena. He writes:

> Full recognition of the reality of experience, then, is the obligatory starting point for any remotely realistic (indeed any non-self-defeating) theory of what there is. It is the obligatory starting point for any theory that can legitimately claim to be 'naturalistic' because experience is itself the fundamental given natural fact; it is a very old point that there is nothing more certain than the existence of experience. (Strawson 2009, p.33)

Thus, we must take serious the claim that mental/experiential phenomena is just physical phenomena, which standard physicalism rejects. For to exclude the experiential would mean to negate the reality of existing. For Strawson, there seems to be no instance in the physical world where there is physical existence that is devoid of experiential phenomenon. Perhaps some liberal standard physicalists would not find premise (3) extremely controversial. The real controversy comes in following Strawson's logic about the ontological nature of things (elemental, ultimates) all the way to a panpsychist conclusion.

We should clarify that what Strawson means by 'mental goings on' refers to the process or state of mental experience as can be discerned by science, physics, and neurophysiology which we typically deem as consciousness or conscious experience. In premise (4), Strawson further radicalizes physicalism by positing that the fundamental fact of experience is a necessary component in understanding the nature of consciousness. That is, we must understand consciousness in terms of the myriad nature of experience, which includes concrete, mental phenomena. Consciousness, then, is weaved into the physical world in terms of experience. In a very real sense, this weaving of consciousness into the physical world demystifies our inclinations and intuitions towards

a more transcendent view of consciousness. In understanding consciousness, we may assume and even desire a transcendence from the physical because we want to attribute consciousness a higher, special status. We may desire a transcendent consciousness to account for a life beyond just the inert, dead physical stuff that we encounter daily. But Strawson physicalizes our desire for transcendence into immanent experience of consciousness.

Premise (5) conjoins premises (1) – (4) to assert that the scope of any form of physicalism must account for the reality of consciousness as experience and as physical phenomena. If (5) is true, then we must further say that consciousness goes *all the way down*. This is a significant component of Strawson's position, which we will detail later.

Premise (6) serves to critique theories that help justify standard versions of physicalism, namely physics and neurophysiology theories. He writes:

> That experience is 'really just neurons firing,' at least in the case of biological organisms like ourselves. But when I say these words I mean something completely different from what many physicalists have apparently meant by them. I certainly don't mean that all characteristics of what is going on in the case of experience can be described by physics and neurophysiology or any non-revolutionary extensions of them… It is that experiential phenomena 'just are' physical, so that there is a lot more to neurons than physics and neurophysiology record (or can record). (Strawson 2009, p. 36)

It seems to be the case that if premises (1) – (5) are correct, then premise (6) must be true. At this point of his argument, if we accept it, standard physicalism is reduced to a false position about the physical world. Perhaps Strawson's re-appropriation of the notion of the physical may not be as unorthodox to contemporary physics due to discoveries and developments in quantum mechanics. Counterintuitive discoveries that radicalize physics, for example, can be seen in the findings of quantum entangle-

ment theory.[3] And the further down we go into subatomic structure of fundamental elements, the more we seem to find counter intuitive events. But here Strawson is not being complacent about whether standard physics will someday be able to account for all physical phenomena. Perhaps his claim is stronger. Perhaps the claim is that standard physics or even realistic physics can never fully define and exhaust the nature of the physical. It is the mental/experiential that in the end keeps eluding us. However, we do not need to interpret this stronger version of premise (6) to validate it. All that premise (6) needs to show is that standard (non-revolutionary) physics cannot capture the full nature of the physical, of what there is.

Premises (7) and (8) describe Strawson's critique of emergentism. He discusses how some see emergentism as a plausible defense of the standard notion of physicalism, while strangely also positing realistic physicalism. He ultimately renders the emergentist position untenable and absurd. For Strawson, two hypotheses characterize emergentism:

a. Non-experiential thesis (NE) of standard physicalism — physical stuff is, in itself, in its fundamental nature, something wholly and utterly non-experiential;

b. Realistic physicalism (RP) — experience is a real concrete phenomenon and every real concrete phenomenon is physical. (Strawson 2009, p. 41)

Here is Strawson's interpretation of emergentism:

> Experiential phenomena are emergent phenomena. Consciousness properties, experience properties, are emergent properties of wholly and utterly non-conscious, non-experiential phenomena. Physical stuff in itself, in its basic nature, is indeed a wholly non-conscious, non-experiential phenomenon. Nevertheless, when parts of it combine in certain ways, experiential phenomena 'emerge'. Ultimates in themselves are whole non-conscious, non-experiential phenomena. Nevertheless, when they combine in certain

> ways, experiential phenomena 'emerge'. (Strawson 2009, p. 44)

Emergentism attempts to explain the mystery of consciousness and even goes beyond standard physicalism by allowing experiential phenomena in the physical world to emerge or rise from non-experiential properties, ultimates (quarks, etc.).

The instance of liquidity is taken to be one of the best examples of the reality of emergentism (Ibid, pp. 42-44). We are apt to understand that, given the combination of certain physical ultimates that make up water molecules, the formation of water fluidity *emerges* from such a combination. Liquidity, it is argued, is one instance of an emergent property that arises from non-liquidity parts, water molecules. Thus, we can explain consciousness in like fashion as being an emergent property of certain non-conscious characteristics coming together to give rise to it, so emergentism is correct. Strawson also formulates emergentism in terms of spatial/non-spatial, extension/non-extension, proto-experiential, and neutral monism, but shows us how each proposed form of emergentism fails for the same reason. He writes, "If it really is true that Y is emergent from X then it must be the case that Y is in some sense wholly dependent on X and X alone, so that all features of Y trace intelligibly back to X" (p. 46). Although emergentism can explain liquidity as an A from non-A properties, but only *in virtue* of the fact that both water molecules and the liquidity of water ultimately are made up of the same physical stuff. It is only in virtue of the fact that both are physical that there is an emergent property. It is misleading, then, to claim that consciousness can be understood as an emergent property because we are to accept that non-physical consciousness emerges from physical elements.

The NE thesis, which partly defines emergentism, above says that ultimate physical stuff lacks experiential properties. But we now know that consciousness is to be understood as experiential. We thus cannot establish an *in virtue of relation* between non-experiential ultimate physical properties to non-physical experi-

ential consciousness. Emergentism in any form makes sense if and only if the emergence of a property Y, say, liquidity, is already contained in X, water molecules, but this is not plausible in the case of emergentist consciousness. At best, emergentism gives us magic and miracles; at worst absurdity and rejection of scientific laws: consciousness just pops into physical existence.

In rejecting emergentism, Strawson's view ends up with that mental/experiential reality must be *all the way down*. That is, if we are to avoid the problems of emergentism of positing experiential properties from non-experiential properties resulting in the miracle instance of consciousness, we must accept that physical ultimates (being all the way down the ontological pipeline) must have an experiential characteristic to give rise to experiential consciousness. This fills the explanatory abyss left by emergentism. We are left, then, to accept premise (9). And if premise (9) is true, then we must accept panpsychism in our realistic physicalism, premise (10).

Strawson does leave us with some interesting last remarks. First, we still must account for the combination problem as noted by William James (Strawson 2009, p. 54). Strawson answers by saying that "One needs a vivid sense of the respect in which every object is a process; one needs to abandon the idea that there is any sharp or categorical distinction between an object and its propertiedness" (Ibid, p. 55). Here, we see Strawson not only reject dualism but perhaps call for a new framework of doing physics in terms of process and non-relational, intrinsic properties. But is Strawson's view of physicalist panpsychism and consciousness fully adequate?

Harman's Critique of Reductive Ontology and Endopsychism

We have now seen how we can make sense of panpsychism via Galen Strawson. Indeed, he has done much of the heavy intellectual work needed to just position it along other theories of what there is in the world, which is why I focused half of this paper on

his work. No longer marginalized, we are apt to simultaneously engage with panpsychism and physicalism coherently, as their coupling present a very real and strong candidate for explaining the physical world. Nevertheless, we will now explore some potential problems facing Strawson's physicalist panpsychism.

Graham Harman approaches panpsychism in a manner different from Strawson, which ends up highlighting the limitations of Strawson's panpsychist ontology. For Harman, understanding panpsychism and *what there is* is intimately related with two other factors: the mind-body problem and the relation between first-person and third-person descriptions. All three factors, however, each have an underlying problem that hinders us from understanding their greater depth and relatedness. Thus, Harman presents three different arguments that lead to a metaphysical conclusion of a variation of panpsychism he terms *endopsychism*.

First, Harman argues that the modern mind-body problem is derived from the body-body problem (Harman 2009, pp. 255-261). He notes that the latter problem dates to the occasionalist tradition in the medieval period of philosophy. Philosophers of that period came up with occasionalism as an explanation of causal powers and how things interact. It was ultimately God who *occasioned*, was the causal power, in the interactions between objects. Harman further tells us that the moderns, especially Descartes, narrowed the initial global body-body interaction problem into the more singular mind-body problem we still currently debate. Descartes framed the human mind opposite and radically different from the physical world. It resulted in the problem of how two different substances can interact.

The mind-body problem, for Harman, arose as a case of falsely assuming to understand object-object interaction and thus privileging the human mind with a god-like status. He writes:

> While quantum theory may add certain complications to the mechanistic view of nature, it does not alter the basic model of physical entities slamming together in space or interacting with fields… But the basic features of causation are

> taken for granted, and have assumed an air of self-evidence that makes materialism [physicalism] the default intellectual position of our time. (Harman 2009, p. 257)

We could see here how both Strawson and Harman critique standard physicalism using similar reasons. For Harman, the problem of standard physicalism is that it assumes that its models and calculations adequately understand atom-atom interaction, for example, and thus enables and forces philosophy to accept such models and move on to try to resolve the problem of how mind and bodies interact. Instead of giving the power to God for explaining causation, as the occasionalists did, moderns (Hume, Kant) gave such power to human experience for explaining causation; but, Harman states, "it [philosophy] cannot continue to leave the vast majority of relations outside its mandate. We need to reawaken a body-body problem ignored by the sciences" (Ibid, p. 256).

To explain body-body interaction, Harman posits the idea of *vicarious causation*. He writes, "Any two entities must interact vicariously, by way of a third. And just as importantly, any entity can serve as such an intermediary—not just God or the human mind" (Ibid, p. 259). Here, we start seeing Harman's departure from Strawson's ontology. Harman pushes Heidegger's ontology to highlight the body-body problem. Using Heideggerian tool-analysis, Harman first notes that Heidegger revealed something about Dasein and objects in his tool-analysis, "The reason the hammer can sometimes malfunction is because it is not reducible to Dasein's current use of it" (Harman 2009, p. 258). The point of the tool-analysis is that the hammer itself is more dynamic than both practice and theory. Dasein's relation with the hammer does not exhaust all the hammer's properties. We must then say that human experience with objects is limited to its own human-object experiences, which allows us to discuss the feasibility of object-object interaction and even experiences.

Harman's version of tool-analysis leads us to see that objects withdraw themselves not only from humans, but from other things. In no one instance, then, does one thing ever fully encounter

another. That is, bodies encounter other bodies only in fragmented, incomplete, and distorted ways. A body-body encounter reveals a third entity or manifestation. Eduardo encounters rock A as only the fragmented entity of rock A, namely rock A*, a different entity from rock A. Likewise, rock A encounters Eduardo as Eduardo*, and Eduardo* is only a fragment, distorted instance of Eduardo not exhausted by any and multiple encounters. Vicarious causation, then, pushes the counterintuitive idea that "relations per se are always a *translating* force, always giving us something different from that to which they relate" (Ibid, p. 259). If vicarious causation is accepted as a framework for understanding causation in the world of object-object interactions, then we can also accept that the mind-body problem is only one instance of the more global body-body problem of interaction and communication. We can further conclude that all objects have independent substance, as was once only granted to humans. Harman's radical position argues for object-substance that is prior to any relation or interaction.

Vicarious causation further radicalizes Strawson's physicalism not by granting objects intrinsic properties, but by acknowledging objects substance as "hidden things-in-themselves." That is, physical objects as described by Strawson limited them to descriptions of their intrinsic properties, assuming that such descriptions exhausted objects or ultimates. But for Harman, no description of ultimates or relations/encounters can ever fully capture the substance being of an object. Strawson's ontological rabbit hole, *all the way down*, may now be tunneling upwards to reanalyze the nature of objects. The reality of objects may be more complex than what Strawson's physicalism holds. We shall explore this below.

Developing the vicarious causation argument to reawaken the legitimacy of the body-body problem, Harman adds the zero-person argument to replace the first-person/third-person dualist view describing experience (Harman 2009, pp. 261-272). First-person reports of experience give us descriptions of *qualia*, the what-it's-like of taste or smell that only first person direct expe-

rience presents. Third-person reports of experience, in contrast, give us objective data or information of events. For Harman, however, the first and third person reporting of experience results in mere descriptions, but mere descriptions are radically different from objects-in-themselves.[4] He writes, "For this reason I will coin the adjective 'zero-person' [object] to refer to the reality of any entity apart from its interactions with other entities of any kind" (Ibid, p. 261).[5] That is, a body or mind is not equivalent or the same as what can be described of them. The description-reality dilemma had focused on bridging the gap between first/third person descriptions of reality, but we can now see how that project seems inconsequential. For Harman, the zero-person object position more aptly grounds the problem by stating that we need to bridge mind/body descriptions and reality.

Harman specifically refers to David Chalmer's work *The Conscious Mind* to show how the zero-person object description-reality gap problem is misunderstood, but can still be shown to underlie prominent works on consciousness (Harman 2009, pp. 272-278). For Chalmers, everything in the universe is logically supervenient, dependent, on the physical, and reduced to structural, functional, or relational properties (Ibid, p. 264). But consciousness (and causal laws and fundamental elements) is a special case where there is no logical supervenience (Ibid, p. 262). They are explained by natural supervenience on the physical (Ibid, p. 264). Ultimately, Chalmers is wary of allowing for consciousness to be understood as an epiphenomenon, which has no causal force in the physical world (Ibid, p. 271). He is also wary of the combination problem, but resolves it by stating that "everything in the world comes down to what is phenomenally intrinsic," as information or description (Ibid, p. 271).

Harman notes that Chalmers does not address his biggest problem, "The major problem he fails to acknowledge is this strangely asymmetrical treatment of body and mind, which grants no macroscopic-sized entities in the physical case but is plagued with an odd tension between tiny and large sized minds" (Ibid, p. 272). As noted above, Chalmers holds a logical-natural super-

venience dualism in his naturalistic position. Chalmer's logical supervenience position leads him to conclude that only microscopic physical particles are real, while macroscopic objects, tables, chairs, etc., are not autonomous objects but functional representations of microscopic structure. For consciousness, however, Chalmers allows for both logical supervenience of physical ultimates, but also natural supervenience, consciousness irreducible to the physical (Ibid, p. 266). But Harman collapses Chalmer's logical-natural supervenience (mind-body) dualism because physical objects and phenomenal experience can be understood merely as information or abstracted descriptions. If this is the case, then:

> He [Chalmers] should drop the idea that there are two basic classes called bodies and minds, and replace it with a dualism of *intrinsic* [object] *realities and the information transmitted about them*. Objects would be zero-person intrinsic realities, prior to any informational abstraction [in encountering, meeting] by other entities. But for objects to become accessible to other objects means that they must be reduced to abstractions, translated into informational holograms. And this is all the dualism we need. (Harman 2009, p. 272)

Mind and body cannot fully be revealed or disclosed through information/descriptions, but such information is the means of accessing objects. Thus, we see how Chalmers indirectly discusses the first/third person description-reality gap problem via supervenience, but we also see Harman pushing supervenience, as he did with Heidegger's tool analysis, to give us the 'zero-person intrinsic object reality and informational model,' which bridges the gap between description and reality, albeit a fragmented reality.

Harman's arguments create an interplay between his notions of vicarious causality and zero-person object to further make sense of the plausibility of panpsychism, but in the form of *endo-psychism* (Ibid, pp. 278-282). He spends significant amount of time detailing how the object-object interaction of informational descriptions model plays out. Thus, we can now make sense of

how, for example, an electron may interact with another electron as zero-person objects, or how two birds interact vicariously. But even if we accept his model, we need to ask where this interaction occurs, if humans do not in fact exhaust the domain of zero-person experience. Harman first reminds us that objects (humans, particles, amoebas) do not in fact interact, "They only interact vicariously in some shared medium where they are somehow able to meet," and the medium is purely informational (Harman 2009, p. 274). Here, Harman says, phenomenology will help decipher the interaction.

In the last section of his article, Harman tries to make sense of an intuition:

> My suspicion is that intentional objects [of consciousness] are a primitive phenomenon found in all experience. If this is so, then even the most rudimentary inanimate experience is torn by a rift between unified intentional objects and their shifting accidental profiles. (Ibid, p. 281)

By ascribing the experience of intentionality to all objects, including microscopic entities, Harman concludes that inanimate objects must also experience consciousness in the intentional realm (Ibid, p. 281). If we accept vicarious causation and zero-person interaction, then we must also accept that even elemental things have an interior and they are "in the interior of a larger entity," as intentional objects (Ibid, p. 254). The move that makes Harman reject full blown panpsychism occurs when he tells us that the relational object, *a real object*, that contains the experience created by zero-person object-object interaction itself is not contained in another entity or object. It itself lacks intentionality and thus is not able to be an intentional object. He writes, "It need not be the case that such a larger object enters into relation with anything else. It certainly has an interior, because that is where my experience occurs right now… It is nearly certain that there are many objects that have a genuine reality, but which still enter into no further relations... Hence, they would be real, but without experience" (Ibid, p. 282). This is where full-blown panpsychism

unravels for Harman. This third object or entity that the zero-person object-object interaction creates undermines pan-psychism because although all real objects contain experience, contain the relational reality between object-object interaction, yet, nevertheless, *the object created by such an interaction itself lacks experience*, lacks relationality and intentionality.

The Anthropomorphic Problem of Consciousness

We have now explored two different positions that lead us to at least postulate the reality of panpsychism, or at least something like panpsychism. Strawson's work pushed physicalism *all the way down* to show ultimates as experiencing entities. Harman pushed all the way up to show how the relationality of zero-person object-object interaction ends in consciousness as experience contained within a larger object. One tentative conclusion we can draw about consciousness is the *if all the way down, then all the way up* perspective. Consciousness, then, permeates all forms, from the microscopic to the macroscopic without one being reduced to or by the other. Consciousness is not inserted or dropped in somewhere along the line of evolutionary experience. As Strawson tells us, it is there from the start. The exploration of consciousness here makes a full circle in its explanation. However, as much as we push towards understanding the reality of consciousness, even as manifest in the immanent objectivity of the zero-person object, there seems to be an inherent problem. It is the very theorizing that helps us understand consciousness that also betrays its reality. The theorizing problem of consciousness could be understood in terms of the anthropomorphic problem.

Simply defined, to anthropomorphize refers to the process of attributing human traits, mental states, and/or features to non-human objects (Andrews 2016). The Judeo-Christian conception of god, for example, usually takes on an anthropomorphic character, as god being benevolent, forgiving, fatherly, etc. This process seems to be a human psychological tendency. For our

purposes, we will simply assume that it refers to some (perhaps archaic) psychological process that allows us to make sense of the world. It could play a significant component in creativity or empathy, but such questions are beyond the scope of this project.

In terms of consciousness, to theorize about it is to filter it through anthropomorphic machinery and down into the realm of the human. But there is no way around this anthropomorphic machinery. It seems necessary and inevitable. Whether you go all the way down or up, it results in understanding it in theory. Strawson's physicalism ends up positing hypothetical ultimates as the building blocks of reality and thus consciousness. Likewise, Harman's zero-person object intentionality theory not only denies us full disclosure of objects' consciousness, but it results in consciousness as a type of relational object that lacks intentionality. We must take notice of how both end up in some speculative metaphysics about consciousness. We do not seem to see or experience these ultimates or relational objects beyond their hypothetical being. Perhaps it is a problem with our propositional language which pushes us to theorize and anthropomorphize.

This may seem like a radical form of skepticism or even mysticism regarding consciousness. But we cannot seem to fully escape anthropomorphism. It denies entry into consciousness, regardless of whether you go all the way down or up. We ultimately fall into the process of human machinery. One day, wander into a desert, forest, or downtown area, and ask yourself, exactly where consciousness is? In my mind as theory? In my bodily experience? Both seem inadequate, leaving us more baffled and searching for a different explanation of consciousness. However, perhaps the very nature of consciousness should lead us in a different direction, away from wanting to theorize, away from wanting to understand, away from a finite resolution. Here, I reference Jacques Derrida's notion of *aporia* not to solve the problems of consciousness, but to help us cope with our anthropomorphic machinery.

Consciousness and Aporia

The notion of *aporia* has been around since the time of the ancient Greeks (Woodruff 2016). It generally refers to a state of being perplexed or a state of impasse without a foreseeable resolution. Aporia is due to holding two equally plausible conclusions, which are inconsistent with one another. "It is raining. It is not raining. It is the case that it is both raining and not raining." This is not aporia exactly, as we can easily verify whether it is raining or not. True aporia leaves us in an irresolute state and impasse. However, I simply wish to present the logical structure of aporia to later make sense of its irresoluteness. The logic and meaning of it is further developed by Derrida in his later writings, which will now help us peer a little more into consciousness.

In his book *Giving Time I: Counterfeit Money* (1992), Derrida uses the notion of *the gift* as presenting us with an aporia. In short, the possibility of gift giving inherently denies the possibility of genuine gift giving. When we give a gift, we typically intend for it to reward or celebrate another. For Derrida, the structure of genuine gift giving is devoid of the element of giving and taking, self-interest, or equivalent actions. The simple response of receiving a gift prompts us to thank the giver. However, in acknowledging the gift to the giver, it nullifies gift giving. 'Thank you' starts the cycle of giving-taking as it assumes that it can exchange equivalent gestures, but a genuine gift does not expect any equivalent gesture, no 'thank you's.' But then how can we understand the gift giving gesture? We cannot acknowledge it as a gift, as to avoid the problems noted above, yet it is a gift. We cannot really say anything about receiving the gift, even to the giver, as it would nullify it, but it is only a gift if we acknowledge it as such. Are we to keep silent and maintain the genuine gesture of a gift? We typically are polite and acknowledge it though, which undermines it. In analyzing the aporia of gift giving, I noticed a distinct and even strange characteristic about aporia in general, which helped illuminate consciousness. Specifically, analysis of aporia results in discovering the *quiet* nature of gift giving, which seems to also be present in consciousness.

Regarding consciousness, aporia develops due to being able to meaningfully theorize and experience the reality of consciousness, yet also holding the view that consciousness cannot be reduced to theory or even describe its full experience. The language that we use to describe experience is not consciousness itself. So here is the insight: the aporia of consciousness shows us that *consciousness as experience—consciousness is just experience all the way down and up—is just experience without language, or description*. We can even be brash and posit that it is language, not consciousness, that turns out to be something like an epiphenomenon, but such brashness requires its own analysis not considered here. Objects may experience consciousness all the way down and up, but we cannot fully reason or talk about it. In trying to understand it, we hypothesize it in terms of experience, but theorizing about experience in any form is the furthest edge we can thread that allows us to peer into consciousness, before we go into the quietude of consciousness itself. This is not skepticism, but it may seem to teeter on mysticism.

The aporia of consciousness takes the form and state of quietude. Juxtaposing the works of Strawson, Harman, and Derrida pushes us to ask and re-ask a seemingly simple question: what does it mean to have and experience consciousness? We know it; we experience it; it is. Yet we cannot further assume that we ever truly and honestly make sense of it. We can only describe it, whether scientifically, logically, aesthetically, etc. But language itself ends at the dark chasm of consciousness.

Conclusion

We have gone down the rabbit hole *all the way down and all the way up*, and theoretically constructed a circular enclosure that captures consciousness. It permeates throughout the enclosed circle, in the vast enclosure of the cosmos and universe, from amoebas to black holes. Yet there seems to be something missing here, and not even qualia itself can fill in the abyss of not being able to further peer into consciousness. Some simple consider-

ations: the physicist splits atoms in her lab and records observations. Does she observe consciousness? The philosopher theorizes about consciousness in her study. Does she experience consciousness? The lay person swims in the ocean and realizes her being conscious in the thrusts of waves upon reflection. Is her reflection consciousness?

In the section focused on Strawson, I wrote that his work demystifies both panpsychism and our inclination to characterize consciousness as something transcendent of the everyday. Although the notion of the quietude nature of consciousness may seem to again mystify panpsychism and consciousness and, even worse, present a sort of mysticism, I need only to remind myself that we, objects, humans, quarks, are physical all the way down and up, but we are physical consciousness in a radical (quiet) sense, not dependent or limited by language.

Notes

1. Nye presents physicalism as a position that offers us a theoretical explanation of being in terms of intrinsic and extrinsic properties present fundamental properties of reality (atoms, etc.), but leaves us much to question about how consciousness fits into theoretical, reductive view of reality. And my assumption is that her physicalism has an explanatory gap regarding consciousness.
2. By 'standard physicalism' Strawson means any type of physicalism that rejects or does not account for the experiential or conscious element of reality in their descriptions and theories. He uses the term 'physicSalism' to refer to standard physicalism, "It follows that real physicalism can have nothing to do with physicSalism, the view—the faith—that the nature or essence of all concrete reality can in principle be fully captured in terms of physics" (p. 34).
3. For a general yet comprehensive introduction to quantum entanglement, please see: Bub, Jeffrey, "Quantum Entanglement and Information", *The Stanford Encyclopedia of Philosophy* (Spring 2017 Edition), Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/spr2017/entries/qt-entangle/>.
4. We could be tempted here to frame Harman's objects-in-themselves in terms of Kant's distinction between phenomena and noumena, but this need not be the case. Whereas for Kant noumena is beyond any experience, for Harman we do have limited access to objects-in-themselves, albeit in a vicarious way.

5. This reflects Husserlian phenomenology where human intentionality regarding objects undergo *epoche*, but Harman later tells us that objects too have intentionality which leads us to assume that they also undergo epoche of humans and objects, but he does not elaborate on this matter.

Bibliography

Andrews, Kristin. (2016) "Animal Cognition" *The Stanford Encyclopedia of Philosophy* Retrieved on February 1, 2017, from https://plato.stanford.edu/archives/sum2016/entries/cognition-animal/

Derrida, Jacques. (1992) *Given Time I: Counterfeit Money* (University of Chicago Press)

Harman, Graham. (2009) "Zero-person and the Psyche" *Mind That Abides: Panpsychism in the New Millennium*, pp. 254-282 (John Benjamins Publishing Company)

Nietzsche, Friedrich. (1974) *The Gay Science* (Trans. Walter Kaufmann) (Vintage Books)

Nye, Alyssa. (2015) "Critique of Russellian Monism" *Consciousness in the Physical World*, pp. 346-369 (Oxford University Press)

Skrbina, David. (2009) "Panpsychism in History: An Overview" *Mind That Abides: Panpsychism in the New Millennium* (John Benjamins Publishing Company)

Stoljar, Daniel. (2016) "Physicalism" *The Stanford Encyclopedia of Philosophy*, Retrieved on January 15, 2017, from https://plato.stanford.edu/archives/spr2016/entries/physicalism/

Strawson, Galen. (2009) "Realistic physicalism" *Mind that Abides: Panpsychism in the New Millennium*, pp. 1-30 (John Benjamins Publishing Company)

Woodruff, Paul. (2016) "Plato's Shorter Ethical Works" *The Stanford Encyclopedia of Philosophy* Retrieved on January 10, 2017, from https://plato.stanford.edu/archives/win2016/entries/plato-ethics-shorter/

A Kantian Approach to the Problem of Shortage of Kidneys

Van Doan

Abstract: In this paper, I aim to explore the negative moral implications of the problem of shortage of kidneys. I will then explain two current attempts around the problem of the shortage of kidneys, namely, transplant tourism and having a legalized, and regulated market for kidneys. I will then explore moral problems—using a Kantian approach—with both transplant tourism and having a legalized, regulated market for kidneys. Next, I will present my own suggestion in an attempt to solve the problem of the shortage of kidneys, namely proposing the harvesting of kidneys of deceased, unclaimed persons. I will argue that if a Kantian approach—which takes into account humanity as an end-in-itself and the concept of rationality—can be used to argue for abortion, then, through logical consistency, the same argument must also apply to the harvesting of kidneys from unclaimed deceased persons. I will also use the concept of consent and argue that deceased persons have no capacity to give consent because they do not have rationality. Finally, I will address a possible objection and offer a response.

Introduction

With the advancement of medical technology, a person who has kidney failure or any other kidney diseases, and so are in need of a kidney, can undergo a kidney transplant surgery in about three hours. Provided that the patient's body does not reject the new kidney, they have a 96% survival rate one year after the transplant and an 86% survival rate five years after the transplant, according to statistics given by Mayo Clinic. In many cases, a kidney transplant is the option that has the highest survival rate. Furthermore,

in many cases, a patient's quality of life is often better than that of patients who undergo dialysis and, considering the fact that many diseases which would cause a patient to undergo dialysis can be diminished significantly, even cured in some cases, it is natural that patients are motivated in considering a kidney transplant as their best option for treatment. This consideration causes a large demand for kidney transplants, and thus kidneys. The immense demand for kidneys, in turn, creates a black market for kidneys—in the form of transplant tourism. Transplant tourism (TT) refers to a patient in need of an organ traveling to another country in order to purchase the transplant of the needed organ. In this article, the focus will be on TT for kidneys. In order for TT to work, there needs to be a demand for kidneys. In addition to the high demand for kidneys, there has to be people willing to give up a kidney in exchange for money. The demand for kidneys and the number of people who are willing to give up a kidney for money combines to fuel the practice of TT. In many cases, the people who are willing to sell their kidneys are very poor and, thus, are in a position where they are coerced by organ brokers for their organs. Kidneys in the black market, which transplant tourism uses, cost between $70,000 to $160,000 per kidney (Shimazano), but the kidney sellers in Pakistan only receive an average of $1,311.40 (Cohen 270).

One way to attempt to 'solve' the shortage of kidneys in the U.S. is TT, another is legalizing and regulating the organ market. Unfortunately, as I shall present, this also has negative moral impacts. Thus, we cannot consider the concept of TT as a morally plausible way to solve the problem of kidney shortages, nor can we consider legalizing and regulating the kidney market.

Section 1 – Background Information on the Problem of Shortage of Kidneys

The problem of kidney shortages is a medical problem with moral implications and consequences. This problem arises due to the fact that there are more patients who need a kidney transplant in order to cure kidney-related diseases than there are available kidneys

for transplantation. When the demand for kidney transplants is far greater than the supply of kidneys for transplantation, it is very difficult for a patient to undergo a kidney transplant, given the current procedures that pertain to kidney transplants. Furthermore, the average wait time for a patient to receive a kidney transplant once they are on a waiting list for it is 3 to 5 years. For many patients however having a kidney transplant is the only option for survival. As of November, 2016, there are 100,791 patients awaiting life-saving kidney transplants, while, according to statistics from 2014, only 17,107 kidney transplants took place in the US. There were also 4,761 people who died while waiting for a kidney transplant, according to statistics given by the National Kidney Foundation. In addition, there are around 3,000 patients added to the kidney transplant waiting list every day; thus, we see that the chances of receiving a kidney transplant gets smaller if the number of available kidneys remain constant. Considering the extreme difficulties and amount of time a patient needs to endure in order to be able to receive a kidney transplant, many patients end up not utilizing the current procedures to get a kidney transplant. Instead, many opt for alternative methods in procuring a kidney. Currently, there are two popular alternative methods to the current kidney transplant procedure in the United States, namely having a regulated kidney market and transplant tourism (TT).

Section 2 – Two Alternatives to the Current Kidney Transplant Procedures in the U.S.

Many people who do not want to go through with the current procedures in order to receive a kidney, due to long waiting times, would end up engaging in what is termed transplant tourism (TT). The patient would go to a developing country—where organ transplant laws are lax, such as India and the Philippines, in order to purchase a kidney. In order to procure a kidney transplant by engaging in TT, the patient pays a certain amount of money and receives a new kidney, surgery included. TT is a black market for kidneys which was created to supply the high demand for kidney

transplantations.

Another type of market in which kidneys are available for transplantation is the regulated organ market where people would be compensated with such things as money, health, or life insurance in exchange for their kidneys. Proponents of regulated organ markets—which includes kidneys—argue that having a regulated market, where people are able to sell their organs, would decrease the gap between the supply and demands of kidneys, while at the same time increasing the well-being of the kidney sellers in the form of compensations.

Section 2.1 – Transplant Tourism, a Closer Look

The constituents of TT consists of kidney sellers, kidney transplant recipients and brokers. Notice that the term 'donor' is not used to indicate the seller of the kidneys because in TT the kidneys are not donated, they are sold. According to a study in 2006 by Syed Naqvi, in which 239 kidney sellers (186 male, 53 female) from Eastern Pakistan were interviewed, all the sellers were very poor (Cohen 2013, p. 270). Of the 192 sellers who agreed to answer the interview questions on monthly incomes, 62% earned between $10 to $30 a month, with the mean income being $15.40 USD, while 32% earned less than $10 USD a month (Cohen 2013, p.270). 77% of the 176 respondents to the question about debt reported that they owed between $1000 to $2500 USD, with a mean of $1311.40 USD (Cohen 2013 p.270). The sellers were promised between $1,146 to $2,950 USD (mean $1,737 USD) per kidney, but no seller in the sample was actually paid that price; instead the range of the amount received by the sellers was $819 to $1,803 (mean $1,377.40 USD) (Cohen 2013, p. 270). The big difference in how much a kidney seller receives and how much they were promised shows that most of the kidney sellers were exploited and lied to. The deductions are largely due to the nephrectomy, hospital stay, and travel expenses (Cohen 2013, p. 270). The kidney brokers did not seem to clearly let the kidney

sellers know that they would actually have to pay for expenses related to kidney surgery. All sellers reported having good health before the transplant, but only 1.2% reported that their health was good after the transplant, 62% indicated that they were physically weak and were unable to work the long hours they did before the transplant (Cohen 2013, p. 270). Essentially, TT is an unregulated activity that exists in the black market where people who need a kidney transplant, and want to skip the legitimate kidney waiting list, will procure a kidney transplant at the expense of the people of a less developed country. In section 3.1 I will evaluate the moral problems associated with this activity.

Section 2.2 – A Closer Look at the Arguments for a Regulated Kidney Market Model

In order to reduce the activity of TT some argue for a regulated kidney market where a person is compensated for selling a kidney. An economist who wanted to maximize utility would argue that a ban in the sale of organs is inefficient for a capitalist economic system, while a libertarian who believes in minimal governmental regulations would argue that it is illegitimate to have a ban on organ sales. The economic argument for a regulated kidney market contends that the problem of the wide gap between the demand and supply for kidneys would be somewhat reduced when there is a regulatory market for kidneys, while at the same time allowing people to have the liberty to sell a kidney for some form of compensation, in addition to the attempt to decrease the activity of TT. While it may seem like a good way to address the problem of kidney shortages and the problem of TT, there are some moral problems to consider, which will be presented in section 3.2

Section 3.1 – Moral Problems with Transplant Tourism

There are many moral problems associated with TT. In this section, I will first explain a Kantian analysis of TT, namely that activities such as TT would violate Kant's conception of a person as an end-

in-itself. I will then use a study done by medical anthropologist and bioethicist Monir Moniruzzaman. He gives further concrete evidence of the social injustices that loom large over the working class people who are trying to get out of debt in Bangladesh.

According to the second formulation of Kant's Categorical Imperative, one must act such that the humanity—whether of your own person or of others—is always an end, and never as merely a means to an end. In other words, a person should be treated as an end in herself, and not as a means to someone else's end. Further, it should not be the case, according to this formulation of the Categorical Imperative, that she uses her person—or body—as a means to some other end.

Thus, according to this formulation of the Categorical Imperative, it is morally wrong to engage in TT, both as the recipient of a kidney and as a seller of a kidney. As the recipient of a kidney—also the broker in the transaction—one is treating the seller of the kidney as a means to an end, where the means is the seller and the end is a kidney for the recipient. The end for the broker in these transactions would be profits, and the means is also the seller of the kidneys. This is a clear violation of the Categorical Imperative. Further, the dignity of the seller is lost in these transactions because the sellers are treated as mere objects by the recipients and the brokers and, by objectifying the sellers, the recipients and brokers of these kidney transplant transactions end up exploiting the kidney sellers. Instead of respecting the kidney sellers as humans with dignity, the brokers treat the kidney sellers as a mere means to an end. As studies revealed, most—if not all—of the sellers of kidneys are very poor or are trying to get out of debt, and to make money as a broker or to buy a kidney as a recipient exploits the fact that the kidney sellers are poor and are trying to get out of debt. Hence, the activity of TT is immoral on many levels. The violation of the Categorical Imperative not only applies to recipients of kidneys and brokers, but also applies to the sellers as well.

By selling a part of your body, i.e. a kidney, in order to receive monetary compensation, the seller of a kidney is violating

the Categorical Imperative because she is treating her body as a means to an end and not as an end-in-herself. Furthermore, according to Kant: "To deprive oneself of an integral part or organ (to maim oneself)—for example, to give away or sell a tooth to be transplanted into another's mouth, or to have oneself castrated in order to get an easier livelihood as a singer, and so forth—are ways of partially murdering oneself" (Kant 1996, p. 177). For Kant, one's humanity is included in one's body, and thus to commodify a part of one's body is to treat it merely as a means to an end, which is a violation of the Humanity Formulation of the Categorical Imperative. These are the theoretical moral arguments against TT based on Kant's Categorical Imperative, there are also some other moral issues with TT raised by a study done by Monir Moniruzzman.

According to information from a study by Monir Moniruzzaman, The People's Republic of Bangladesh is an emerging organ bazaar that has been in existence for decades. It is endorsed by national media that openly publish newspaper classifieds seeking kidneys among other transplantable organs (Moniruzzaman 2009, p. 70); and, every day, these organ classifieds reach millions of poor people in the country, some of whom will eventually sell a part of their body in hopes of getting out of poverty. The recipients, according to Moniruzzaman, are either local or overseas residents (Moniruzzaman 2009, p. 70). Seeing the expansion of this practice, many organ brokers have expanded their networks and run the business for an enormous fee (Moniruzzaman 2009, p. 70); thus, we currently see transplant tourism as a lucrative and attractive practice for those who want to make money at the expense of others.

Moniruzzaman uses the term bioviolence to describe the activity of procuring "fresh" organs, and considers bioviolence as a blend of physical, structural, and symbolic violence, all of which are carried out to extract organs from the oppressed bodies of the poor (Moniruzzaman 2009, p. 72). Moniruzzaman's ethnographic study reveals that as kidney sellers' health deteriorated, their economic conditions actually *worsened*, and their social standing

declined after they sold their kidneys (Moniruzzaman 2009, p. 79). Moreover, Bangladeshi kidney sellers reveal that selling a kidney has profound psychological and psychosocial impacts on them, especially in relation to their selfhood (Moniruzzaman 2009, p. 72)

Moniruzzaman's study reveals horrible unethical facts associated with TT. The physical, psychological, social, and psychosocial damage to sellers of kidneys are very real. The main reason why a person would sell a kidney was because they hoped that it would get them out of poverty, but according to Moniruzzaman's study, not only did they not get out of poverty, most of them actually ended up financially worse.

Transplant tourism is an exploitive activity with bad consequences for kidney sellers. It is a form of social injustice. From a Kantian perspective, it violates the humanity formulation of the Categorical Imperative, and studies actually reveal evidence that people are exploited and used as a means to an end where they suffer tremendous damages to their bodily health, psychology, and even finances.

Section 3.2 – Moral Problems with a Regulated Kidney Market

The argument that in selling one's kidney, one is violating the humanity formulation of the Categorical Imperative still holds in a regulated kidney market. One is still selling one's body part, which, according to Kant, is part of one's humanity and, thus, one is using one's humanity, or personhood, as a means to an end. When one sells one's kidney, one is, in essence, disregarding the respect of one's dignity. Thus, it is still immoral for one to sell one's kidney, even in a regulated market.

In a regulated market, the sale of a kidney would cause kidneys to be regarded as a commodity in a market, not part of one's personhood or humanity. Further, if kidneys are viewed as a commodity, it can potentially be viewed as collateral for money lenders. In a regulated kidney market, it can very well be the case

that in addition to being viewed as a commodity, a kidney could be used as collateral.

Imagine a situation in which a person is in debt in a society where a regulated kidney market exists, it can very well be the case that the lenders of the person in debt might have a lawsuit against the person in debt. If the person in debt has no way of paying any part of the debt owed, it could be the case that the lender would require the person in debt to sell her kidney in order to generate some money to take care of a part of the debt. Thus, a regulated market, legalizing the selling of kidneys, can potentially force a person to sell a kidney in order to pay for a part of their debt. Even if there were no trial in this situation, money lenders would be inclined to place pressure on debtors to sell their kidneys in order to repay some of the owed amount. This is a form of coercion, and thus immoral. Furthermore, if selling kidneys became legal, lenders may create contracts where a borrower would be contractually obligated to sell a kidney if they have no other means of repaying the loans, which is another form of coercion.

Thus far, I have presented two ways that are employed to address the problem of shortage of kidneys. One way is transplant tourism, which is illegal, and the other way is having a regulated organ market. Both ways are immoral because both violate Kant's humanity formulation of the Categorical Imperative. In addition to violating Kant's Categorical Imperative, both TT and regulated organ markets each have other ethical problems. I will now shift to my own suggestion for how to address the problem of kidney shortage.

Section 4 – Harvesting Kidneys from Unclaimed Deceased Persons and the Concept of Consent

Every year, there are thousands of people in cities across the United States who are deceased but are unclaimed as a family member of anyone. These bodies end up being cremated and their remains end up in mass graves. Recently, 1,400 remains

were cremated and buried in Los Angeles county. Can it be said that these persons, after they are dead have the capacity to give consent?

Currently the laws regarding next-of-kin are unclear, but according to a study published by the journal of Nephrology Dialysis Transplantation "the next-of-kin are involved in the organ procurement process in most nations regardless of the consent principle and whether the wishes of the deceased to be a donor were expressed or unknown." Thus, in cases where the deceased have a next-of-kin, they have more power in determining if the deceased is going to be a donor or not. In many cases, regardless of whether consent was given by the deceased,-consent is assumed to be given or not before the deceased has actually died.

When a person is deceased, they cannot give consent on any matter. Current organ donation laws seems to acknowledge this fact because they give more power to the next-of-kin when deciding whether the deceased is going to donate his organs or not. Thus, if we consider this fact and extend it to deceased persons who are unclaimed, it would not be far-fetched to suggest that it is morally permissible to harvest kidneys from unclaimed deceased persons.

Section 4.1 – A Kantian Argument for Harvesting Kidneys from Unclaimed Deceased Persons

According to Kant's humanity as an end-in-itself formulation of the *categorical imperative*, we should treat others as ends in themselves, and not treat them merely as means. This moral principle applies to rational beings, i.e., human beings capable of rational thought. So, in essence, we should treat rational beings as ends in themselves and not merely as means to an end, i.e. a person should be treated with respect as an end in herself, and never to be used as a means to some other end. This formulation of the *categorical imperative* states that rational beings are intrinsically valuable. According to Kant, it is our duty to preserve our own rational

nature and develop our own natural abilities. Thus, considering Kant's humanity as an end-in-itself formulation, we see that rationality is a very important end, an end that ought to be preserved. When we are talking about rationality, we are talking about an attribute that a living human being has, thus, a fetus cannot be said to have rationality.

Section 4.2 – Applying a Kantian Argument for Abortion to Harvesting the Kidneys of Unclaimed Deceased Persons

A Kantian argument can be used to justify abortion, more specifically, the argument that a fetus is not an end-in-itself precisely because a fetus does not have rationality. Abortion, under this conception is not at all like murder, as some pro-life advocates would claim. The end-in-itself does not exist for a fetus because a fetus does not have rational agency. Thus, using this argument, abortion is morally permissible.

According to Kant's humanity as an end-in-itself formulation of the *categorical imperative*, we should treat others as ends in themselves and not treat them merely as means. This moral principle applies to rational beings. So in essence, we should treat rational beings as ends in themselves and not merely as means. This formulation of the *categorical imperative* states that rational beings are intrinsically valuable. According to Kant, it is our duty to preserve our own rational nature and develop our own natural abilities. Thus, considering Kant's humanity as an end-in-itself formulation, we see that rationality is a very important end, an end that ought to be preserved. When we are talking about rationality, we are talking about an attribute that a living human being has, thus, a fetus cannot be said to have rationality.

In the case of abortion, a living organism is actually killed-a fetus is considered a living organism as it is composed of living cells. In contrast, the harvesting of organs from deceased people does not involve actively killing a living organism, thus, one cannot argue that abortion is morally permissible while at the same

time argue that it is not morally permissible to harvest organs from unclaimed deceased persons because this would be inconsistent.

To hold the view that abortion is morally permissible on the grounds that a fetus does not have rationality, and thus not an end-in-itself entails that one must also hold the view that harvesting organs from an unclaimed deceased person is morally permissible. In the abortion case, the fetus can be terminated because the fetus does not have rationality. Thus, the same should hold for the harvesting of kidneys of unclaimed bodies: the deceased person is not capable of giving consent, while, at the same time, the deceased person is no longer an end-in-itself because she lacks rationality, which is an attribute that is important in Kant's formulation of humanity as an end-in-itself.

Kant argues that it is our duty to preserve our own life, rationality, and to develop our natural abilities. This duty only exists for people who are alive and does not exist for a deceased person. Thus, when a living person needs a kidney it can be said that it is their duty to acquire a kidney in order to preserve their own life and rationality. According to Kant, it is not morally permissible to treat others as means to an end, but, as argued previously, a deceased person cannot be said to be an end-in-herself because she is no longer a rational being. If we take Kant's claim that we ought to preserve our own lives and rationality, and considering that a deceased person is not an end-in-herself, then it is actually morally required that the kidneys of the deceased be harvested in order to preserve the life and rationality of living rational agents.

Section 4.3 – Harvesting Kidneys from Unclaimed Deceased Persons to Help with the Problelm of Shortage of Kidneys

The problem of the shortage of kidneys is prevalent in modern societies. It creates immoral situations in the form of transplant tourism, where people living in poverty are preyed upon and coerced into selling their kidneys. Their hopes in selling a kidney is to get out of poverty, however the contrary is true. Most kidney sellers, according

to the study done by Moniruzzaman, end up being economically worse off than if they had not sold a kidney. Further, the studies reveal that most kidney sellers go through physical, emotional, psychological and psychosocial damages after the removal of their kidneys. Having a regulated kidney market also has moral problems, e.g., commodifying kidneysand kidneys potentially being used as collateral by money lenders. On the theoretical level, both methods of harvesting organs violate the humanity as an end-in-itself formulation of the *categorical imperative*.

I maintain the position that harvesting kidneys from unclaimed deceased persons is not only moral according to Kant's Categorical Imperative but also our duty, tied to the preservation of life and rationality. Harvesting kidneys from unclaimed bodies is not only morally permissible, it seems to be a moral requirement. In harvesting kidneys from unclaimed bodies we will diminish the problems of transplant tourism and, also, prevent the need for regulated kidney markets, both of which present moral problems in themselves. Thus, with harvesting the kidneys from unclaimed bodies, we are not only saving and preserving the life of a sick person, we are limiting the unethical practice of transplant tourism by reducing the demand for kidneys. Therefore, the injustices of transplant tourism would be diminished.

Section 5 – Possible Objection and a Response

A possible objection to this view is that perhaps the unclaimed deceased person did not give consent to donate her kidneys before she died, that it is her right to not want to donate her kidneys to those in need. A response to this is that it is true that her consent is necessary but her will and desire to not donate her organs exists only when she is alive. After she is deceased she does not have the ability to give-or not give-consent. Furthermore, she cannot be said to have a will or desire to donate or not donate. Thus, under the view that it is morally permissible—perhaps even morally required—to harvest organs from an unclaimed deceased person, but at the same time, while a person is alive, we ought to treat them as ends-in-themselves and give them all the dignity

and respect that ought to be given to rational beings. Insofar as a person is living, they are entitled to be treated as an end and not as a means and their humanity ought to be respected.

Conclusion

In this paper I addressed the ethical problems of transplant tourism and having a regulated kidney market by using Kant's humanity as an end-in-itself formulation of the Categorical Imperative. I also gave evidence from various studies in order to demonstrate the negative effects endured by kidney sellers in developing countries. I then suggested another way to address the shortage of kidney problem, namely harvesting kidneys from unclaimed deceased persons, and I used Kant's humanity formulation of the Categorical Imperative to argue that harvesting kidneys from unclaimed deceased persons is morally permissible.

Bibliography

Cohen, Glenn. (2013) "Transplant Tourism: The Ethics and Regulation of International Markets for Organs" *Journal of Law, Medicine and Ethics*, pp. 269-285

Kant, Immanuel. (1996) *The Metaphysics of Morals* (Trans. Mary Gregor) (Cambridge: Cambridge University Press)

Kidney transplant. (2016) Retrieved on April 20, 2017, from http://www.mayoclinic.org/tests-procedures/kidney-transplant/details/results/rsc-20203643

Moniruzzaman, Monir. (2012) "Living Cadavers' in Bangladesh: Bioviolence in the Human Organ Bazaar" *Medical Anthropology Quarterly*, pp. 69-91

Organ Donation and Transplantation Statistics (2016) Retrieved on April 20, 2017, from https://www.kidney.org/news/newsroom/factsheets/Organ-Donation-and-Transplantation-Stats

Rosenblum, A. M. Horvat, L. D., Siminoff, L. A., Prakash, V., Beitel, J., & Garg, A. X. (2012) "The Authority of Next-of-Kin in Explicit and Presumed Consent Systems for Deceased Organ Donation: An Analysis of 54 Nations" Retrieved on April 20, 2017, from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3363979/

Shimazano, Youske. (2007) "The state of the international organ trade: a provisional picture based on integration of available information" *Bulletin of the World Health Organization*. vol. 85, no. 12, pp. 901-980

Utopia and Fabulation: Erasures and Affirmations

Craig Laubach

Introduction

In his paper "Deleuze and Guattari and the Future of Politics: Science Fiction, Protocols and the People to Come," Ronald Bogue suggests that the concept of *fabulation* in the work of Deleuze and Guattari is to be thought of as an improvement upon their concept of *utopia*. My purpose here is to partially mirror the sentiment that fabulation indeed represents a culmination of the work undertaken by Deleuze and Guattari to develop a concept of utopia; however, I will advance a claim that thinking fabulation *merely* as an improvement of utopia may motivate a forgetfulness or dismissal of the important nuances apparent throughout Deleuze and Guattari's development of these ideas. I will begin by offering a rendering of Deleuze and Guattari's concept of utopia in conjunction with an analysis of its relevant usage in the history of Marxism and critical theory. I will then discuss the concept of fabulation with primary reference to Deleuze's *Cinema 2* alongside Bogue's explication of the concept. The exegetical difference I will propose, albeit minor in some respects, raises the question of why, despite their reservations about the term itself, did Deleuze and Guattari nevertheless work towards a positive rendering of utopia as a concept? Notwithstanding transparent comments they have made about utopia being a "bad concept," I will challenge Bogue's hasty relegation of utopia and instead advise a reading of it which affirms its commensurate importance within Deleuze and Guattari's conceptual topography. Despite this minor objection, I will nonetheless argue in line with Bogue that the concepts of utopia, fabulation, and *a people to come* all support a conceptual framework suggestive of a unique revolutionary praxis which is

(1) capable of facilitating the development of provisional political aims while remaining inoculated to effete notions of utopian thought in its conventional sense (2) prepared to mitigate the potential implosion of resistance movements which face the unpredictable ways of evolving, such as encountering the possibility of an authoritarian turn. Furthermore, I follow Bogue in opposing Phillipe Mengue's claim that the *micropolitics* of Deleuze and Guattari merely espouses a misguided antipathy towards modern forms of Western democracy, a mode of governance which Mengue points to as most suitable for potentially realizing the political ideals envisaged by their work. In response to Mengue's claim, this essay will conclude with notes on the work of Etienne Balibar, whose own analysis of the concepts addressed here serves to support the claim that the work of Deleuze and Guattari sets forth a viable politics of resistance in stark contrast democracy as it has been traditionally practiced in the West.

UTOPIA

Casual employment of the term 'utopia' quickly evokes its caricature: the quixotically impossible, the capriciously idealistic, the politically unfeasible. However, the consideration of utopian thought as a philosophical concept in Deleuze and Guattari (and more generally in the tradition) has entailed both affirmations of and departures from its conventional associations. Deleuze and Guattari's treatment of the term utopia in *What Is Philosophy?* is unique in its attempt to absorb something of its affirmative usage in the history of political philosophy in that it attempts to formulate a new concept with which to challenge a familiar notion of possibility. They claim that utopia "designates that *conjunction of philosophy, or of the concept, with the present milieu*—political philosophy (however, in view of the mutilated meaning public opinion has given to it, perhaps utopia is not the best word)" (Deleuze 1994, pp.99-100, emphasis in the original). Despite their suggested reservations, their concept of utopia is nonetheless robustly utilized in fleshing out what is at stake in *"[summoning] forth a new people, a new earth"* and a vision for philosophy itself:

> [U]topia is what links philosophy with its own epoch, with European capitalism, but also already with the Greek city. In each case it is with utopia that philosophy becomes political and takes the criticism of its own time to its highest point. Utopia does not split off from infinite movement: etymologically it stands for absolute deterritorialization but always at the critical point at which it is connected with the present relative milieu, and especially with the forces stifled by this milieu. (Deleuze and Guattari 1994, pp. 99-100)

Deleuze and Guattari cast the conceiving of utopia as an epoch's most radical form of critique. As a political concept, they formulate it as the production of a collective fantasy angled against the political crises of the present. They construe utopia as a fantasy whose attempt at realization creates new exigencies which factor back into an ongoing shift of the horizon of crisis. In its linkage with philosophy, utopia designates the conditions for the possibility of overcoming the involution of *thought itself* at its critical limit. Whereas utopia has been commonly conceived of or attempted in the form of detached spatial localities, Deleuze and Guattari identify the site of utopia's *becomings* as temporal events at the edge of the present. Deleuze and Guattari work towards a definition in which to become utopian means not to be uprooted from the present or remain fugitive from its antagonisms; instead they cleave to a sense in which utopia represents a kind of political creativity set against its own horizon of contingencies. They furthermore emphasize that the site which connects utopia with the present is constituted by a suppression of "forces stifled by [the present] milieu." These forces comprise rival political and philosophical tendencies whose actualization operates critically towards or in full opposition to institutional hegemonies. Deleuze and Guattari also indicate that the edge between present and future is marked by a complex emotional tonality. The disintegration of the territorial boundaries of dominant paradigms accelerates a sense of limit often experienced as intolerability by individual bodies, collectivities, factions and so on which have

been repressed. The surge of an antagonistic political climate motivates "breaks" or disinvestments from the political field. The most decisive of disinvestments would be achieving utopia in its conventional sense, an absolute dislocation from the political topography. Ruptures in the political field are continually ensued by a plurality of forces in a process of constant reterritorialization; modes of political dissent which entertain fantasies of total escape, however, disavow the occurrence of their own creativity. Utopia in its conventional sense gives up taking on a valuable affirmative character by abandoning the very site constitutive of its own act of imagining itself. However, by committing utopia to an embeddedness in the present, Deleuze and Guattari motivate a sublimation of the impulse towards a total flight from the political field into the force of absolute philosophical critique. The form of radical critique implied by Deleuze and Guattari calls upon the forces suppressed in the present to self-liberate in the face of dominant systems of power. Philosophy appears in this dynamic as emulative of the deterritorializing power of dominant forces in the form of philosophical critique. What's more, this critique emerges in conjunction with a *summoning forth of a new people, a new earth*—new political figurations which induce flights towards the new, but not from the present itself.

Deleuze and Guattari's attempt at setting forth a positive conception of utopia calls into question the extent of overlap with and departure from Marxist and critical theory standpoints. In consideration of their commitment to an identity as Marxists and their affinities towards Frankfurt School thinkers, we should be curious about their intention to recover a term whose treatment in the mainstream leftist thinking from the 19th to 20th centuries had ranged from its ambivalent appropriation to outright dismissal. Moreover, we should be curious why Deleuze and Guattari strived at recasting utopia as a concept definitive of philosophy's grandest of gestures: its ceaseless rejuvenation of itself as a discipline. I would like to propose a method of reading utopia in Deleuze and Guattari which (1) situates it positively in their conceptual topography, not only as a scaffold for the concept of fabulation but one

which uniquely connotes a sense of full critique and maintains a character of political obstinacy, (2) remains invocative of specific philosophical forebears who have also developed utopia as philosophical concept, and (3) celebrates the audacity of conceiving of a future not necessarily bound to a present facticity. I also believe that it is precisely the navigating of anxieties which attend to some critical approaches towards utopia which have brought to bear positively for philosophy, as is the case with the figures I will soon present.

Deleuze and Guattari's hesitancy at employing the term utopia may not only be derived from its mangling in public discourse, but is also informed by its identity as a pejorative term in philosophy as well. In *The Communist Manifesto*, Marx and Engels famously excoriate historical attempts to found independent utopian settlements in the absence of the socioeconomic conditions necessary to sustain them. Not only did they reproach such communities for their indifference towards historical contingencies, such societies often incubated a tendency towards recreating a host of reactionary social relations (feudal, religious, etc.) which Marx and Engels felt exacerbated class antagonisms (Marx and Engels 1848, p.33). However, like their successors, Marx and Engels did not fail to notice the important ways in which utopian thinking functioned to aggregate the negative collective sentiments towards capitalism in its earliest phases. They concede that some utopian movements were even founded under revolutionary pretenses, but all succumbed to the naivety that they could exist independently of the ravages of capital. Contrast this observation with Marx's fascination with the rise of the bourgeoisie in late feudal society in which a collective extrication from feudal dominance occurred in accordance with a shift in the economic mode of production. Tracking with the timely technological advances of the era, the bourgeois revolution began to emerge at the margins and through the interstices of a degrading feudal order long before it proclaimed victory. Its nascent expressions were experimental in character: localities reconfigured methods of production and exchange by establishing small factory operations as the commu-

nities around them budded with a new political imagination. Marx and Engels attempt to invoke this dynamic, suggesting a similar systematic development of the proletariat to bring about the historical contingencies needed to achieve a communist society. In one sense, we can undertake a reading of *The Communist Manifesto* as a critique of modes of political invention, one which contrasts both the productive effects and challenged sustainability of untimely political fantasies.

Marx and Engels' critique of utopian socialism is carried over into the discipline of critical theory; however, critical theory ascribes a refurbished importance to political fantasy which strives to envision that which is not the case in a way criticized by Marx and Engels. It is the Frankfurt School to whom Deleuze and Guattari credit some measure of influence in their provisional acceptance of the term 'utopia'; through a cursory analysis of critical theory's treatment of utopia, we can better navigate the coordinates plotted by Deleuze and Guattari. We begin with Horkheimer, who articulates a sentiment which parallels the *Manifesto*'s critique of utopia:

> But Utopia is no longer the proper philosophic form for dealing with the problem of society. It has been recognized that the contradictions in thought cannot be resolved by purely theoretical reflection. That requires an historical development beyond which we cannot leap in thought. Knowledge is bound up not only with psychological and moral conditions, but also with social conditions. The enunciation and description of perfect political and social forms out of pure ideas is neither meaningful nor adequate. (Horkheimer 2002, p.269)

Here, Horkheimer reiterates the ineffectual nature of utopian thinking mentioned by Marx and Engels. Conversely, Horkheimer identifies an important role of critical theory, which is to produce an alternative to the notion of utopia, an alternative that employs a mode of thought akin to fantasy to produce an image of the future divergent from an extant state of socioeconomic affairs:

> One thing which [critical theory] has in common with fantasy is that *an image of the future which springs indeed from a deep understanding of the present* determines men's thoughts and actions even in periods when the course of events seems to be leading far away from such a future and seems to justify every reaction except belief in fulfillment. It is not the arbitrariness and supposed independence of fantasy that is the common bond here, but its *obstinacy*. (Horkheimer 2002, p.220, all emphases added)

Horkheimer's comparison illuminates some contours suggestive of critical theory's broader methodology. Much of the scope of critical theory's concern is an analysis of present historical development, however it is not entirely limited to it. Moreover, it is suggested that the facticity of the present socio-economic reality and its imagined trajectory should not be established as a limited criterion for determining what is possible in the political realm. The role of *obstinacy* in the development of critical theory's methodology is also picked up by Marcuse, who not only relates it to a concept of utopia but also deems it to be an important feature of philosophy's legacy of articulating ideality:

> Like philosophy, [critical theory] opposes making reality into a criterion in the manner of complacent positivism. But unlike philosophy, it always derives its goals only from present tendencies of the social process. Therefore it has no fear of the utopia that the new order is denounced as being. When truth cannot be realized within the established social order, it always appears to the latter as mere utopia…The utopian element was long the only progressive element in philosophy, as in the constructions of the best state and the highest pleasure, of perfect happiness and perpetual peace…Critical theory preserves obstinacy as a genuine quality of philosophical thought. (Marcuse 2009, pp.105-106)

The specter of utopia haunts Marcuse's valorization of ideality

and his bid to capture something of its essence for a critical methodology. The flourish of technological progress, in Marcuse's view, has brought with it a new way of thinking of possibility in the face of normativity. Utopia (conceived of as investments in unrealizable dreams), for Marcuse, is now behind us; however, the possibilities utopian thought had once envisioned have now been delivered over to modernity. The once ridiculed hopes to create plentiful surpluses have reemerged with the growth of industrial productivity; the once unfeasible projects lambasted as utopian have now become possible. Under capitalism, new contingencies and crises continually merge in a historical confluence in a way that is called 'progress': the once untimely becomes timely, the impossible dissolves in the flourish of technological advance—despite that social ills remain. Marcuse believes the tension between the intrinsic possibilities of advanced economies and their indifference towards a complete redress of systemic abuses will likely eventuate a break in the continuity of the current political modality such that new standards will emerge. In the same vein, Deleuze and Guattari predict the relaunching of past political struggles to win a future once disparaged as wishful thinking (Deleuze and Guattari 1994, p.100).

The tentative character of Deleuze and Guattari's choice to employ the concept of utopia as definitive of the conjunction of philosophy with the present becomes easier to understand once we outline what has been at stake in their project: (1) constructing an alternative to the notion of possibility by partially liquidating prior definitions of utopia and (2) emphasizing the significance of the thought's event in the present as that which creates the new. However, what they have sought to articulate in utopia finds more suitable expression in fabulation.

Fabulation

Deleuze and Guattari adopt the term fabulation from Bergson, who employed it to describe the "instinctive tendency of humans to anthropomorphize and attribute intentionality to natural phenomenon" (Bogue 2014, p.99). In his new formulation, Deleuze inverts

Bergson's otherwise negative characterization of this tendency to articulate the positive artistic capability of humans to render collective narratives which are ultimately transformative. Fabulation takes place when someone is 'caught in the act of legending', says Deleuze; in other words, it occurs through participation in a type of collective storytelling (Deleuze 1995, pp. 125-126). Its purpose is "capturing the affects and percepts of sensation", the creation of a pervasive aesthetic sensibility which overcomes normative modes of experiencing the world (Bogue 2014, p.100). Like utopia, fabulation points to a juncture between the world in its material sense and the articulation of the images which shape it. Fabulation, too, is a rethinking of possibility which brings to bear productive effects in the real world through an "[activation of] the 'powers of the false', to falsify orthodox truths in the process of generating emergent truths" (Bogue 2007, p. 81). Unlike conventional utopian thought, fabulation does not presume a pre-ordained notion of itself. It has no *a priori* determined political coordinates. Hence, the process of fabulation remains open-ended and always incomplete.

Most importantly, fabulation is the invention or summoning forth of *a people to come*. Deleuze best elaborates this concept in his description aesthetic creation. He views art not only in terms of the object created by the artist, but it also involves the envisioning of a new people receptive to newly developed aesthetic sensibilities. The artist is naturally limited in their ability to do so, only being able to invoke a people and not create them. Fabulation can also be construed as the invocation of a suppressed people, those who have been relegated politically or whose presence has been diminished by dominant political tendencies, such as imperialism, nationalism, and so on. The enactment of fabulation in its specifically political sense is marked by experiments in collective determination, the creation of a "scaffolding for resistance" against the repressive power by majoritarian forces (Bogue 2007, p. 82). Deleuze's *Cinema 2* details important conceptual peculiarities of fabulation which avail us here. He points to the various film cultures of the third-world and those subjugated by colonial

imperialism as social groups who have used cinema as a medium to generate self-affirming representations in the face of their oppressors. Their struggles play out in defiance of portrayals by masters and colonizers who have either abandoned them to silence or altogether purported their non-existence as a people. Deleuze observes how films, like ethnographies, offer patronizing depictions of marginalized people and how national propaganda pieces often aim at either diluting or resituating minority presences within totalizing narratives. In short, the suppression of minority voices in the cinematic venue acts as a continued provocation towards them to tell a different story about themselves.

Fabulation, however, is not merely a recovery of the lost images of a people; it also infuses old representations with new ones. Fabulating entities may be triangulated between traditional mythic figures, subjected to the creolizing forces of imperialism and global capitalism, and otherwise challenged to shed all kinds of enervating images of itself. Deleuze highlights the films of Rocha as evacuations of myth which motivate such a transformation. Fabulation entails "raising misery to a strange positivity", reframing adversity in many possible "creative simulations": caricature renditions, accentuations of the grotesque, empowering speculative fiction, and so on. Amplifications and contortions of a mythic memory are undertaken to induce a cinematic *trance*:

> "[I]t is the trance which makes the speech-act possible, through the ideology of the colonizer, the myths of the colonized and the discourse of the intellectual. The author puts the parties in trances in order to contribute to the invention of his people who, alone, can constitute the whole [ensemble]." (Deleuze 2009, p.223)

The trance state creates a sort of suspension of dominant perceptual tendencies. It occurs as an aesthetic intervention which catalyzes the development of new emotional modalities in the form of percepts and affects. In doing so, it precipitates the conditions for both the invocation of the new and the reproach of one's masters—it prepares the cinematic space for the inauguration of

a new discourse, a collective utterance. As Deleuze notes, the speech-act uttered within the trance "create[s] itself as a foreign language in a dominant language, precisely in order to express an impossibility of living under domination" (Deleuze 2009, p. 237). The coming into presence of a missing people thus occurs through a stylistic aberration of cinematic themes and familiar discourses. The theme of radical semiotic reconfigurations undertaken as an emboldening of resistance recurs throughout the work of Deleuze and Guattari. In *A Thousand Plateaus*, Deleuze and Guattari claim that to become "a foreigner in one's own tongue" and "making a language stammer" as a mode of subversive expression is capable of enacting transformative difference. (Deleuze and Guattari 2007, p. 98) The aesthetic moves suggested here are intended to actuate the new as a kind of qualitative shift through a rupture of dominant semiotics.

Denying that Deleuze and Guattari thought fabulation was a superior concept to utopia is a difficult, if not impossible, claim to make. In a late interview with Toni Negri (which Bogue also cites), Deleuze says, "Utopia isn't the right concept: it's more a question of "fabulation" in which a people and art both share. We ought to take up Bergson's notion of fabulation and give it a political meaning" (Deleuze 1995, p.174). Bogue takes this comment and others like it to quickly move past the significance of utopia and towards fabulation. Admittedly, fabulation is more strongly cast as having a provisional character which maintains a connectedness to a present state of affairs, a quality which Deleuze and Guattari seemingly worked towards in their development of utopia. Fabulation's ability to think that which is not the case eschews the kind of temporal disjunction associated with utopian thought, proceeding more carefully on the narrow footing from present to future. Moreover, its enactment does not entail the creation of totalizing microcosms that utopia does.

What seems understated in Bogue's analysis is how utopia was utilized as a scaffold in the development of fabulation. Its elevation to the level of a concept which represents something important about the nature of philosophical thought itself is

undertaken alongside its recognition as a poor concept. Moreover, in *What is Philosophy?* utopia and fabulation are mentioned side by side, with utopia, warts and all, remaining undeleted from the text in its final publication. Deleuze and Guattari are notorious for representing conceptual prototypes in their major works and later noting their insufficiencies without redaction. Their works often reflect a palimpsest-like quality, with concepts being worked on in different ways in different texts, always invoking their alternative renditions but acknowledging their respective nuances and departures. This raises the question of how to read and characterize a body of work inclusive of sketches which may have otherwise been elided from the work. Whatever their authorial intent, their methodology nevertheless provides us with a conspicuous and unpaneled view to an extensive genealogy of various figures at work in their creative process.

A People To Come and the Future of Politics

The enactment of fabulation, as we have discussed, fosters the rise of a people to come. A people to come are an emergent community which is called into being by the creation of new aesthetic modalities and the embodiment of new narratives. Their imminent arrival inaugurates the future, but does not secure a fixed vision of it, as a new people eschews utopian notions. They are called upon to speak in new political discourses which diverge from dominant ones. Might these new people also reject democracy? In view of the intolerable conditions which are said to precipitate their rise, will a new people continue to tolerate existing forms of the modern democratic state? What is the defining political character of a people to come, and what are forms of governance they might envision? Will these forms of governance operate more viably than their predecessors? Balibar, in a discussion of emergent political communities, offers the following:

> The political community as it is claimed and as it takes shape on the horizon of actions of resistance or insubor-

> dination is not given, but is *always still to come or to be invented*. It is not complete or self-sufficient but incomplete, conflicted, exposed, to the intrusion of the other, *which it needs to constitute itself and which nonetheless most often disrupts it and calls its identity into question*. (Balibar 2014, p.289, emphasis added)

Balibar's community of citizens (a self-described analog of Deleuze's "a people to come") is marked by the ongoing manifestation of its identity in relation to both its internal and external antagonisms. The community Balibar suggests begins as a gathering of excluded bodies and voices, "multiple histories, heterogeneous identities, and explosive combinations of archaism and modernity that make up the reality of [a people]" (Balibar 2014, p.290). Such communities are brought about in collective resistance to the mutual intolerability of living under domination. The sustained imposition of dominant narratives on an emergent political community induces the creation of new shared ways of thinking collectivity, merging fragments of the past with new prospects. As Balibar suggests, the character of such a group is conditioned through its porousness, its malleability of its internal limits, and its disposition towards that which threatens its provisional constitution. An important accent in Balibar's appropriation of Deleuze's concept is the notion that a political community is sustained

> by recognizing what it is constitutively missing or prevents it from being complete, exclusive, self-sufficient that it paradoxically rediscovers the capacity to exist and act as a political body (ibid.)

Such open-ended political configurations remain under dual threat. Not only do communities driven by provisional revolutionary goals exist in opposition to majoritarian forces, but they may also succumb to stigmas cast upon them by the factions of resistance movements which begin to develop stricter ambitions (a dynamic highlighted in *Anti-Oedipus* in Deleuze and Guattari's

discussion of *subject-groups*). However, a capacity for self-criticism and collective shifts in its praxis is, as Balibar notes, also a capacity of a political community to reconstitute itself as a provisional unity and act upon the present horizon.

Let's turn now to the objections Bogue raises against Mengue's reading of a people to come. Mengue lambastes the micropolitics of Deleuze and the implications which he reads into them: a people to come is a reflection of an "unconscious idealism", a trite leftover of modernism resentful towards the impossibility of achieving its aims. Mengue believes that a more effective non-foundational politics is available in modern forms of democracy "because [they are] the only possible logic in the fact of the plurality and rivalry of opinions and political actions" (Bogue 2007, p.85). Deleuze, in his later life, made disparaging comments about democracy, which Mengue claims to express Deleuze's "hidden longing for the foundation of a stable, ideal political order" (ibid.). Both Mengue and Bogue identify this trend in Deleuze's thought, though Bogue takes a different path. In Bogue's account, Deleuze's criticism was not intended as a total rejection of democracy, but was issued largely as a polemic against the "complacency and self-satisfaction" of those who uncompromisingly support democracy as it is practiced in the West (Bogue 2007, p.86). Bogue is accurate in dismissing the claim that Deleuze calls for a politics which manifests as nothing more than enclosed avant-gardism which myopically celebrates a retreat from the political mainstream. If we take Balibar's extension of a people to come to be congruent with Deleuze's original envisioning of the concept itself, it follows that an extraordinary level of political engagement would be required in sustaining a community embodying the commitments which they both suggest. In its real manifestation, it would likely act multi-tendentially, expressing itself among a host of political venues both inside and outside the mainstream.

What Deleuze disparaged about modern democracy, as Bogue notes, was the relatively narrow spectrum of debate it fostered, one intolerant to the kind of discourse that Deleuze valo-

rized. The force of *doxa* imposes itself on all manner of political conceptions, making the public forum a less than ideal venue to foster the kinds of political transformation for which Deleuze and Guattari advocate. Modern democracy, as is implied in Deleuze's negative comments on democracy, registers poorly in its ability to generate more than a "mere contestation of competing ideologies" within a constrained spectrum of debate (Bogue 2007, p.86). Moreover, modern parliamentary forms of democracy reveal their own inherent limitations in their being beholden to the preservation of the state as an apparatus which secures the conditions of the modern economic form.

In the end, what Deleuze agitates for is the creation of something new, which often involves a break with clichés and exaltations of familiar political discourse. This is best exemplified in Deleuze's remembrance of the events of May '68, in which he talks about how despite protesters not achieving their host of demands, the effort produced a new kind of person, a new subjectivity. The intolerability of social conditions reached a perceived critical limit inducing not only widespread direct action efforts, but it also brought about new forms of relations with the world. This generation of a new political narrative, a kind of fabulation, also generated new possibilities. What's more, it fostered an obstinate collective voice which enunciated its permanent refusal to interface with social reforms which merely appear as extensions of the system which the era's malcontents railed against. Deleuze ascribes much importance to this resoluteness and maintains that such an attitude allows us to keep vigil over what is still possible in this world (Deleuze 2007, pp.11).

Bibliography

Balibar, Etienne. (2014) *Equaliberty: Political Essays* (Durham: Duke University Press)

Bogue, Ronald. (2011) "Deleuze and Guattari and the Future of Politics: Science Fiction, Protocols and the People to Come" *Deleuze Studies*, vol. 5, pp. 77-97

Deleuze, Gilles and Felix Guattari. (1994) *What is Philosophy?* (New York: Columbia University Press)

Deleuze, Gilles. (1995) *Negotiations*, 1972-1990 (New York: Columbia University Press)

Deleuze, Gilles. (2007) "May '68 Did Not Take Place" *Autonomia*, pp. 11-12 (Los Angeles, CA: Semiotext(e))

Deleuze, Gilles and Felix Guattari. (2007) *A Thousand Plateaus* (Minneapolis: University of Minnesota Press)

Deleuze, Gilles and Felix Guattari. (2009) *Anti-Oedipus: Capitalism and Schizophrenia* (New York, NY: Penguin Books)

Deleuze, Gilles. (2009) *Cinema 2 the Time-Image* (London: Continuum)

Deleuze, Gilles, Barbara Habberjam, and Hugh Tomlinson. (2011) *Bergsonism* (New York: Zone Books)

Deleuze, Gilles and Paul Patton. (2014) *Difference and repetition* (London: Bloomsbury)

Marx, Karl and Frederick Engels. (1848) "Manifesto of the Communist Party" Retrieved on March 17, 2017 from https://www.marxists.org/archive/marx/works/1848/communist-manifesto/

Horkheimer, Max. (2002) *Critical Theory: selected essays* (New York: Continuum)

Marcuse, Herbert. (1970) *Five Lectures; Psychoanalysis, Politics, and Utopia* (Trans. Jeremy J. Shapiro and Shirley M. Weber) (Boston: Beacon Press)

Marcuse, Herbert and Jeremy J. Shapiro. (2009) *Negations Essays in Critical Theory* (Colchester: MayFly)

Parr, Adrian and Ronald Bogue. (2014) "Fabulation" In *The Deleuze Dictionary*, pp. 99-100 (New York, NY: Columbia University Press)

The Human as the Object to be Eaten

Andre Agacer

Introduction

In *The Simpsons* episode "Lisa the vegetarian", Lisa, the middle child of the titular family, undergoes the realization of the violence of meat-eating and the human treatment of nonhuman animals. The episode features an educational/propaganda film titled "Meat and You: Partners in Freedom". The title suggests that meat-eating is a patriotic act foundational to our western liberal democratic ideals. The film justifies meat eating by arguing that the human consumption of other animals is a natural process ("in nature a creature invariably eats another creature to survive") in virtue of the "food chain" (a diagram is shown with a human figure at the center with various depictions of animals pointing towards the human mouth). But this "anthropocentrism" is not just justified as a natural hierarchy but an act of defense at the species level ("if a cow ever got the chance, he'd eat you and everyone you care about").

The episode parodies the dogmas of meat-eating (viz. the anthropocentric food chain) that naturalizes the "animal industrial complex" and satirizes the critique made against meat-eating (viz. the cow as a carnivorous human-eater). Although the critique of an anthropocentric worldview (i.e. the world-for-us), as presented in the film's diagram of the "food chain," must not be understated, the ridiculous nature of the claim that a cow would "eat you and everyone you care about" if they "ever got the chance" contains a serious concern the carnivore presents to the vegan/vegetarian. Would you still hold on to your vegetarian/vegan ideals even if humans could be *mere* meat to nonhuman animals? *The Simpsons* writers were clearly making a joke about the impossibility of a

cow's carnivorous desire, but, in doing so, assumed that humans can't be rendered *as meat*. This tendency to deflect the possibility of rendering humans into *mere* meat, as an object to be eaten by an animal other, is not isolatable to a joke in an animated sitcom. From utilitarianism to deontology, ethical frameworks that prioritize human actions fail to conceptualize humans as objects subjected to nonhuman animals. The impossibility of rendering humans into *mere* meat commits veganism to an anthropocentric ethical framework that rejects the possibility of the objectification of humans. The question arises—what would a non-anthropocentric ethical veganism look like?

Practical and health concerns aside, the usual criticisms made against veganism can be reduced to two categories: (1) the *selective indignance* of veganism and (2) veganism's unintentional anthropocentric ethics. By selective indignance, I am referring to veganism's tendency to focus on the violence committed against particular nonhuman animals (e.g. cows and dogs) and therefore ignoring the violence committed against nondomesticated animals or other animals we cannot sympathize with. Various animal rights ethicists have attempted to resolve these criticisms by formulating non-speciesist ethical frameworks (e.g. Singer's "equal consideration of interests" and Regan's "subjects-of-a-life") but speciesism is not equivalent to anthropocentrism. Speciesism is the unquestioned assumption that nonhuman animals should be excluded from the realm of rights, privileges, protection, and ethical consideration we grant human beings. Anthropocentrism, on the other hand, is the assumption that the human perspective has the unique and privileged capacity to interpret the world. Although vegan ethical frameworks can avoid the trappings of speciesism, veganism can still be anthropocentric. The purpose of this paper is an attempt to provide a foundation for a non-anthropocentric veganism. Rather than discussing veganism from an economic perspective, I will be discussing the practice of meat-eating and the process that renders animal bodies—human and nonhuman—into meat.

In recent years, much work has been done to approach

ontology, ethics, and political problems in non-anthropocentric terms. The object-oriented ontology (OOO) of Graham Harman, and Ian Bogost, among others, has attempted to provide a "flat" ontology that gives equal consideration to all "actants" and rejects the privilege philosophy has granted the human perspective. Although OOO has focused on establishing a non-anthropocentric ontology, there has been very little consideration of the ethical and political implications OOO could have on nonhuman beings. Critical animal studies (CAS) theorists have attempted to question, and rethink, the human-animal binary to arrive at transformative political and ethical practices in our engagements with nonhuman animals. Matthew Calarco notes that non-anthropocentric ontologies, such as OOO, could be useful for CAS unless these ontologies are merely attempts "to transfer quintessentially human epistemological and phenomenological perspectives onto the whole of the nonhuman world at the ontological level" (Calarco 2012, 59). My aim of this paper is to unpack Calarco's concerns of OOO—particularly the ethical debate surrounding veganism—and to demonstrate that OOO's non-anthropocentric ontology is complicit to, and reinforces, anthropocentric practices.

In the first section, I will establish the differences between CAS's ethical and political critique of anthropocentrism from OOO's ontological and epistemological critique of anthropocentrism. Particularly, I will focus on how CAS and OOO differ in their formulation of a non-anthropocentric ontology. In the second section, I will discuss Ian Bogost's critique, from an OOO perspective, of veganism and its "selective indignance" towards the consumption of animal flesh. In other words, Bogost reveals the anthropocentrism of veganism. As a response to Bogost's critique, in the third section I will discuss Matthew Calarco's (CAS) concept of "indistinction" as a general approach—theoretical and practical—of reconceptualizing the relation between humans and nonhuman animals. Not only will Bogost's critique be shown to be unmotivated, but I will also demonstrate that a veganism that adopts "*indistinction*" as a foundational concept of its ethical framework, and reframing it in non-anthropocentric terms, will be

immune to Bogost's critique. In the fourth and final section, I will show Bogost's critique to be complicit with an ethico-political anthropocentrism despite having a non-anthropocentric ontology by demonstrating its parallel with "difference approach" of CAS. Finally, I will argue that the concept of "indistinction" must be adopted by OOO theorists to achieve a non-anthropocentric ethics and politics consistent with their philosophic ideals.

Section 1: The Critiques of Anthropocentrism

While both CAS and OOO provide us with critiques of anthropocentrism, the motivations for their critiques differ. CAS is concerned with the institutional, political, and ethical consequences produced by the human-animal ontological distinction. CAS's theoretical and practical consideration is not limited to nonhuman animals, but to all beings—human and nonhuman (e.g. ecosystems, the subaltern, minoritarian groups)—affected by the human-animal binary reinforced by anthropocentric ideologies and practices. OOO, on the other hand, critiques anthropocentrism to avoid, as Quentin Meillassoux calls, the *correlationist* pitfalls that plague post-Kantian philosophy. Correlationism is the dogma that the world is inseparable from human thought. Consequently, the correlationist is ontologically committed to the impossibility of conceiving the world without humans and humans without the world. OOO avoids correlationism by deprivileging the human perspective in favor of a "flat" ontology that gives equal consideration to the perspective of all objects whether they be doors, cats, error messages, gears, humans, or cotton. OOO attempts to reorient our understanding of "perspective" and "subject-hood" in nonhuman terms. By granting "subject-hood," to inanimate beings, OOO's hope is to build an ontology that doesn't privilege, and centralize, one perspective over another—namely, the human perspective.

OOO's critique of anthropocentrism is motivated by ontological and epistemological concerns to extend our philosophic

scope beyond human thought. While CAS theorists are still concerned with these consequences of anthropocentrism, their primary concern lies in the ethical, political, and institutional *effects* of anthropocentrism. I will now focus on CAS's ethico-political critique of anthropocentrism.

CAS's Ethico-Political Critique of Anthropocentrism

In "Identity, Difference, and Indistinction" (2012), Calarco identifies three approaches in critical animal/animality studies and the strengths and limitations of these approaches. Calarco gives a critical analysis of the two dominant approaches—identity and difference—and identifies the recent emergence of another approach he calls "indistinction". Rather than "eliminate the first two approaches in favor of the third, nor [to] establish a dialectic in which the first two approaches are subsumed in a third, higher form," Calarco aims to bring increased attention to modes of thought and practice that utilize the indistinction approach. For Calarco the identity and difference approach are useful in ways that the indistinction approach is incapable of, but the indistinction approach will be useful when the identity and difference approaches fall short. I will now focus on the strengths and limitations of the identity-based approach.

The identity approach has been endorsed by the analytic animal ethics tradition associated with philosophers and theorists influenced by, to name a couple, Singer's utilitarian equal consideration of interests and Regan's deontological view that (some) nonhuman animals are "subjects-of-a-life". They are primarily concerned with extending moral consideration to nonhuman animals that have the relevant moral characteristics. This extension of moral consideration is done by two moves: (1) "a rigorous application of Darwinian [or a naturalistic] ontology" and (2) "normative impartiality" (Calarco 42). The recognition of a naturalist ontology grounds humans as a purely natural and biological species with a "fundamental relatedness to other animals". Consequently, our moral criteria for normative impartiality requires the

"rational moral agent to extend equal moral consideration to all beings who have interests". Irrespective of which moral characteristics one deems fundamental (e.g. sentience, subjectivity), philosophers who employ the identity approach view species differences as morally irrelevant in the normative domain. Just as an ethical framework that excludes consideration based on race or gender can be deemed racist or sexist, an ethical framework that excludes consideration based on species difference would be deemed speciesist. For Calarco, the greatest contribution of the identity approach has been its "effort to be especially critical and vigilant about any attempts to draw sharp, clean, binary distinctions at the ontological level between the human species and nonhuman animal species" (Calarco 43).

The identity approach has inspired practices and changes in legal policies with the explicit goal to reject our dogmatic speciesist ideologies. For example, the Great Ape Project, founded by Peter Singer and Paula Cavalieri, advocated for the conference of basic rights to nonhuman great apes. Granted that they were successful in granting rights to great apes (e.g. Spanish parliament in 2008), the extension of rights is severely limited to nonhuman animals closest to the human animal—the great apes. Thus, this approach has been limited by a "stubborn form of *logocentrism* and a persistent *anthropocentrism*" (Calarco 44). By *logocentrism*, Calarco argues that the identity approach theoretically and practically centralizes speech, rationality, and knowledge in their analysis of the "animal encounter" (45). By solely employing the "space of neutral rationality" for transforming ethical frameworks, it fails to account for the other means of transforming thought and practice that's not reducible to rational discourse (e.g. emotional affect, the ethical encounter). In other words, rational discourse is not enough for people to alter their practices—we must include emotional responses and an ideological critique of anthropocentrism. Not only is this approach persuasively limited, its *logocentric* tendencies are deeply problematic by basing *logos* (i.e., the privileging of speech, rationality, and knowledge) as the ground for ethical consideration. As the name suggests, the

identity approach begins with the premise that the criterion for defining the scope of ethical consideration (i.e., who's interests and actions must be ethically evaluated) is dependent upon the relevant moral characteristics that we have regarded, historically speaking, as *uniquely* human (e.g. sentience, subjectivity). Therefore, extending the scope of ethical consideration on the grounds of human *logos* is always an "anthropocentric extension."

This "anthropocentric extension" determines ethical priority to nonhuman animals that resemble human beings on *logocentric* grounds. Ironically, by using the identity approach in animal ethics to grant ethical consideration to nonhuman animals who resemble human beings, this criterion also serves as a criterion to justify the devaluation of nonhuman animals and, in some cases, human beings (e.g. racism, sexism, ableism) who fail to qualify as "rational moral agents." For this reason, the critique of speciesism—the uncritical bias of one's own species over another species—cannot provide us with a robust enough critique of the human-animal binary. Therefore, the real problem is not speciesism but the anthropocentrism that is in the methodological core of the identity-approach. By qualifying ethico-political subjecthood on a criterion of resemblance to the "ideal" human, the identity-approach reinforces a logic of exclusion that creates a paradigm of identity and difference that marginalizes the "Other."

Calarco defines anthropocentrism as "the privileging of that class of beings who best fulfill the conception of what is considered to be quintessentially human over and against all nonhuman others" (46). Calarco provides five primary conceptual characteristics of anthropocentrism: (1) a "specific form of human exceptionalism," (2) a binary human-animal ontology, (3) strong moral hierarchy that privileges the human over other animals, (4) the reproduction of sub-or extra-human zones of exclusion, and (5) the employment of "a wide variety of institutions to found and reproduce a privileged space for the human" (Calarco 2014, 416).

The characteristic of "human exceptionalism" identifies the tendency to "filter, measure, and relate [other worlds] through quintessentially human perspectives and concerns." This excep-

tionalism entails a human-animal ontology in which there are discrete distinctions between humans and animals on the grounds of establishing uniquely human characteristics (e.g. consciousness, awareness of death, subjecthood) that are absent in other nonhuman animals. This establishes a hierarchy of human superiority over the deficiency of the animal (i.e., animals are beings in a "lack"). For example, we can kill chickens just because they cannot contemplate on their own death the way humans can. This discrete ontological distinction creates a strong moral hierarchy that values the human over the animal/nonhuman/subaltern. This "violent hierarchy" operates under a "missing premise," or logical jump, that justifies the ethical devaluation of the animal/nonhuman/subaltern. By "missing premise" I am referring to the reasoning that we can mistreat nonhuman animals simply on the grounds that they lack certain characteristics. This logic of inclusion/exclusion is not only used in the human-animal distinction, but it has been used to justify racism—among other forms of exclusions—by identifying the "animality within human beings" (418). This "anthropological machine," as coined by Giorgio Agamben, identifies only a portion of biological human beings into the political and ontological category of "human." In turn, this inclusion excludes human beings who fail to qualify as "human" (e.g., mentally incapable, comatose). This aspect of nonhuman zones of exclusion shows that anthropocentrism does not only affect nonhuman animals, but it also affects the marginalized humans who are denied full subject-hood and recognition from the dominant culture. From this theoretical exclusion of the nonhuman/animal, it's clear to see this logic of anthropocentric exclusion is enmeshed institutionally. As such, this discloses the way in which we view "ourselves" and how we grant the "quintessential" human being a privileged subject position.

By privileging the human subject, the identity-approach is incapable of providing a conceptual framework that recognizes subjectivity in nonhuman terms. In terms of meat-eating, the identity approach is incapable of conceiving the human as *mere* meat and an object to be consumed by nonhuman animals. This is

largely due to the insistence that humans are qualitatively different from nonhuman animals. By framing the human as a "rational moral agent", we divorce human beings from their animal bodies that could be consumed by an animal other. The identity approach merely asks us to find the human (i.e., *logos*) in the nonhuman animal but not the animal (i.e., embodiment and product of the reduction to the body—meat) in the humans. Additionally, considering the identity approach's insistence on the quintessential human subject, this approach is uninterested in the ruptures that forces thought to disclose the animality in the human—namely, the ever-present possibility of humans to be reduced to *mere* meat.

Unlike the identity-approach, the difference approach understands the inherent ontological problems of *logocentrism* and *anthropocentrism* in the human-animal binary. The difference-based approach "tends to derive a more Continental style of philosophical orientation and is characterized primarily by an exploration of the non-anthropocentric dimensions of post- or anti-humanism" in order to create "radicalized notions of difference" (Calarco 2012, 48). Although there have been various versions of the difference-approach, the dominant version has been associated with Jacques Derrida and deconstruction. In order to rethink human-animal issues, deconstruction attempts to complicate human subjectivity and the animal by a "play of *difference* out of which all singularities and relations (both human and nonhuman) emerge" (49). By paradoxically insisting the differences constituted, instead of the similarities, between the human and the animal, neither are shown to have ever been the homogenous categories we assume them to be—the human and the animal were *always already* made up of differences. All categories (e.g. the human, male, hetero/straight/cis, white) are only understood and recognizable in relation to, and in response to, its Other (e.g. the nonhuman/animal, female, queer, black). In other words, the "human" and the "animal" are emergent beings that are neither static nor reducible to the other.

Ethically speaking, the difference approach calls for an "open-ness" to the "unanticipated ethical force" by the arrival

of the Other—in this case, the "Animal". Extending ethical consideration to the animal, Derrida argues that the *arrival* of the animal in the scene urges us into an ethical relation by the "call of the Other" and, because of the inherent inaccessibility of the Other, our obligation to the Other is equally infinite. Although this ethical "open-ness" of the "encounter" expands the scope of ethical consideration by allowing for an "infinite" reevaluation of ontological distinctions, the difference-approach fails to provide us with a concrete account for a non-anthropocentric politics. By merely avoiding the homogenizing anthropocentric practice that limits the identity approach, merely complicating and problematizing the human/animal distinction fails to radically transform the socioeconomic institutions that reinforces the violence of the binary. By insisting on the differences between the human and the animal, this approach could be complicit in anthropocentric practices reinforced by the human-animal distinction. The privilege conferred upon humans is enmeshed in the history of the distinction itself. By not attempting to look beyond the distinction, this approach can reproduce our heritage of anthropocentric practices. A veganism that uses the difference-approach, by asserting the discrete nature of the human-animal distinction, will be incapable of, or at the very least uninterested in, recognizing a zone of indistinction between the human and animal (viz. fleshy embodiment). Although the difference-approach can recognize the animal in the "human" and human in the "animal", the stubborn assertion of the discreteness of the distinction blinds this approach from the material reality of the disproportionate nature of humans (i.e. the inedible disembodied animal) and the nonhuman animal (i.e. the edible embodied animal).

Considering the anthropocentrism of the identity approach and the political and transformative impotence of the difference approach, Calarco provides a "third-way" that is neither a synthesis of the two nor an approach that supersedes the two—the indistinction approach. Focusing on the "zones of indistinction" between human and animal, the indistinction approach finds the site of radical transformative possibilities in the ways in which

humans become indistinct from the animal other. Instead of either reducing the animal Other to the qualified subject position of the human or maintaining the human-animal distinction in order to deconstruct it, and ignoring the ethical and political baggage that comes with the "human", the indistinction approach avoids the anthropocentrism and logocentrism of the identity approach and the political and transformative inefficacy of the difference-approach. By blurring the lines of the human-animal distinction, this approach not only presents a non-anthropocentric ontology, it also provides a theoretical framework to discover the practical possibilities to transform the institutional effects of the human-animal distinction. In terms of the problem of the reduction of animal life into meat, this approach provides veganism with a more robust framework that eludes the other two approaches.

I will now discuss the non-anthropocentric ontology of Object-Oriented Ontology and how it differs from CAS.

OOO: Non-Anthropocentric Flat Ontology

Object-Oriented Ontology rejects the ontological and epistemological priority traditional western philosophy has granted humans and the human perspective by asserting the equal ontological status to all objects—living and non-living. Under OOO's "flat ontology," the perspective of cats, cotton swabs, CPUs, and the earth's moon has equal philosophical significance as the human "perspective." For OOO, "perspective" is not limited to the subjective point of view of sentient, or cognizant, beings. Rather, perspectives are the relational point of view of one object with another object. OOO's assertion of the ontological equality of all objects attempts to bypass the anthropocentric consequences produced by metaphysical frameworks that begins on human terms (e.g., sensation, reflection, logic). One might think that OOO is rejecting the value and insight provided by the human perspective, but, rather, OOO is rejecting the privilege we grant the human perspective. This rejection allows us to recognize the finitude of human thought and the nonhuman perspectives we can only merely attempt to speculate. OOO's non-anthropocentric ontology attempts to conceive

reality from the nonhuman perspective by providing an account of inter-object relations.

Graham Harman, the founder of OOO, analyzes inter-object relations in his "tool-being" interpretation of Martin Heidegger's broken hammer analysis in *Being and Time*. Heidegger argues that Dasein's, his term for humans, primary interaction with objects in the world is through "using" them as equipment that serves a function to fulfill meaningful projects we're engaged in. As equipment/tools, we never recognize the objects themselves beyond their use in projects. This "readiness-at-hand" relation with objects prevents Dasein from recognizing the environmental background network of objects that support our projects. But even after a tool becomes "present-at-hand" and no longer "ready-at-hand," the entire supporting network of equipment (e.g. the nail, board, the laws of physics) is still *withdrawn* from our awareness because equipment is forever in practical action itself and to make equipment present-at-hand is to reduce them to their "objective" status as a tool-object but not as the tool enmeshed in a network of equipment.

Harman takes Heidegger's conception of equipment, withdrawal, and the interplay between "presence-at-hand" and "ready-at-hand" as a general ontological theory of objects he calls *tool-being*. Tool-beings are "vacuum-sealed objects" distinct from their modes of being as equipment in relation to other objects. Unlike Heidegger, Harman argues against the assertion that "*human existence* is the hero that frees entities from the present-at-realm" (Harman 2002, 19). Even when Dasein is aware of objects in their "presence-at-hand-ness", their "veiled performance or execution become concealed behind some present-at-hand configuration". In other words, awareness of tools outside of their use (e.g. the positive sciences) will allow us to see the tools as equipment but we will be incapable of recognizing the tool-being of the object-in-itself because the tool is still tied to our human relation to the tool as equipment. Rather than interpreting Heidegger's tool-analysis as a means of critiquing "the notion of independent objects, as if to champion instead a subjective human realm of gadgets

or linguistic signs," Harman's interpretation of tool-analysis shows that inanimate objects themselves are not "*just* manipulable clods of matter, not philosophical dead weight best left to 'positive science'" but, rather, objects are "already aflame with ambiguity" and never reducible to our human use or examination. For Harman, Dasein's ontological engagement with the world is not uniquely human. Rather, Dasein's *tool-being* is found in *all* objects/beings. In the case of veganism, the human perspective of suffering may differ from, for example, a cow's perspective of suffering in terms of quantitative differences between "levels" of sentience, but this difference does not entail profound qualitative differences that justify an ontological hierarchy.

In OOO's flat ontology, objects—as tool-beings—relate with other objects *as* equipment but not the object-itself. In reference to the debates on causality in classic Islamic philosophy, Harman uses the example of the relation the interaction between cotton and fire. When cotton makes contact with fire and disintegrates in a matter of seconds, we intuitively assume that the fire consumes the entirety of the cotton but OOO claims that the fire can never fully consume the cotton even when there's no physical trace of the cotton. Unlike Bruno Latour's *actor network theory*, which argues that the "actants" in a network are reducible to the relations between actants, OOO states that the object-in-itself is metaphysically inaccessible despite being enmeshed in a network of relations with other objects. When the fire comes into contact with cotton, a fire-cotton relation emerges in which the fire and cotton present the qualities relevant to the other. As cotton is being disintegrated by the fire, the fire-in-itself will be inaccessible to the cotton—and vice-versa. Objects only "sense" the qualities other objects reify that are relevant in the relation (e.g. the cotton as a flammable object). The object-in-itself can never be reduced to the qualities they present in their relations with other objects. For this reason, Harman argues that the object-in-itself *withdraws* from other objects. However, if the cotton in-itself *withdraws* from the relation between the cotton-fire, then the cotton in-itself doesn't burn, but rather the "caricature" qualities of the cotton enmeshed

in the cotton-fire relation is the one burning.

This *withdrawal* suggests that objects are incapable of directly interacting with each other. Harman calls this "shifting communication and collision between distinct objects" the *carpentry of things*. The carpentry of things is not concerned with the "physical but the *metaphysical* way in which objects are joined or pieced together, as well as the internal composition of their individual parts" (Harman 2005, 2). In the case of the fire and cotton, the qualities we perceive as cotton may cease to exist but the cotton-in-itself continues to exist, metaphysically, alongside the fire. Since the objects-themselves *withdraw* from qualities they animate in their relations with other objects, objects only interact with the "caricature" of the other (i.e. the qualities relevant to the relation). Interaction via a "caricature" of the other entails that objects causally affect the other indirectly. Harman calls this indirect causation *vicarious causation*. The inaccessibility of objects-themselves implies that objects inherently resist conceptual capture. This inaccessibility is not just limited to human sense-perception, but all sciences and logic—human thought in general has a finite scope. Human thought, or any other inter-object relation, can never objectively understand objects because the object as it is presented to us is already always a "caricature" of the object. In the case of meat production, we conceptually and materially reduce animals to *mere* meat and reject any possibility for animals to be more than meat. Meat is the caricature of animal bodies.

Although I cannot presently do justice to Graham Harman's robust ontology, this discussion will suffice to show the parallels between OOO and CAS. Besides asserting the necessity of a non-anthropocentric ontology, we can find a lot of parallels between CAS and OOO. I find three key parallel principles shared between CAS and OOO: (1) animals and objects resist conceptual capture, (2) a relational ontology that resists privileging the perspective of one type of being from another, and (3) the "encounter" between beings/objects as the site of possibilities. I will now discuss how CAS and OOO diverge despite their similarities. Particularly, I

will discuss Ian Bogost's critique of CAS and veganism.

Section 2: Critique of CAS & Veganism

Despite their mutual commitments to overturn anthropocentrism, OOO has criticized CAS's "arbitrary specificity." Bogost argues that despite expanding the domain of inquiry, "[animal studies] stops short by focusing on a single domain of 'familiar' actants—dogs, pigs, birds, and so forth—entities routinized thanks to their similarity in form and behavior to human beings" (Bogost 2012, 8). CAS urges us to limit our scope to relations "from the vantage point of human intersubjectivity, rather than from the weird, murky mists of the really real." Therefore, ethical veganism has a selective indignation in terms of which alimentary acts are prohibited. Although these critiques are not exclusive to OOO, something suspicious is going on when OOO's critique of veganism is vulnerable to reinforcing, or at the very least complicit to, anthropocentric practices and ideologies despite having a nonanthropocentric ontology.

In *Alien Phenomenology*, Ian Bogost takes Harman's object oriented ontology and the concept of vicarious causation between objects and extends it to include ideas, concepts, relations *qua* relations, and processes. Consequently, objects can be further separated by immaterial objects. This extension to immaterial objects allows Bogost to conceptualize interobject relations that are not reducible to carnal qualities—even considering Harman's broad conception of sense-perception that includes inanimate objects. For example, a theist has a conception of a deity that frames their metaphysical conception of reality. In the theist's relation of reality, can we say that reality is effected by the theist's conception of reality? Based on our best empirical theories on reality, we can assume that deities, as conceived by the theist, do not exist. The theist's metaphysical conception of reality is their *interpretation* of reality derived from the fact of being thrown into the world.

Bogost calls this the interpretative relation with objects *metaphorism*. Any attempt to understand the subjective experi-

ence of another being/object will always elude us because "the very idea of experience requires this 'being-likeness,' a feature that eludes observation even if its edges can be traced by examining physical properties" (Bogost 62). Bogost's metaphorism mirrors Thomas Nagel's *objective phenomenology*, which claims that physical reductionism can never fully explain subjective experience, but instead of rejecting empathy and the imagination as Nagel does, Bogost argues for an *alien phenomenology* which "accepts that the subjective character of experiences cannot be fully recuperated objectively, even if it remains wholly real" (64). Alien subjective experiences (e.g. bat's use of sonar) can never be understood but we can only apprehend them metaphorically (e.g. the bat operates *like* a submarine). Rather than reject the distortion created by the metaphor, an alien phenomenology accepts metaphor and distortion because the other's subjective (alien) experience should remain alien and withdrawn from our understanding. As constructed metaphors, object relations should not be taken as reality but, rather, as caricatures and tropes of the object. The only reality we can derive from object relations is the metaphoric nature inherent to object relations. But how does metaphorism effect our ethical relations with nonhuman beings (viz. nonhuman animals)?

Bogost argues that ethics is a uniquely human practice that, when in a relation with nonhuman objects/beings, reduces them to the all too human interpretation of the relation. "Metaphorism is necessarily anthropomorphic" thus it would be an "egoistic practice" to impose our ethical frameworks onto other objects (76). This poses a problem to vegan ethics in two ways: (1) ethical codes are always "ethics *for us*" (for example, the ethics of meat eating often overlooks the ethical and practical problems that arise for veganism when considering its incompatibility with carnivorous companion animals) in which "moral standards sit on the inside of the unit *human being*" (i.e. veganism as anthropocentric) and (2) its "selective effrontery"/indignation towards the violence committed against animals ignores the "suffering" of plant-life (73). Although the sentience of plant-life is still a controversial

topic ever since botanist Jagdish Chandra Bose's plant-nervous-system hypothesis, the mere assertion that plant insentience is axiomatic prevents us from viewing plants as being more than senseless organic life. For Bogost, the vegan cannot argue away (1) and (2) unless they resist framing their ethical principles (e.g. non-violence, interests) in anthropocentric terms. It is unclear if Bogost believes that veganism can resist its anthropocentric tendencies, but it's safe to assume that he believes that veganism is an anthropocentric practice because of his insistence that ethics is *always for us* humans.

Ethical relations are *more* alien than sensual relations because in the ethical relation the relation with the Other is mediated by another relation, namely the ethical relation. When a vegan eats a soybean, they enjoy the sensual qualities of the soybean relevant to the relation (e.g. vegetability, nutritive content), but they add moral qualities to the soybean (viz. being ethically permissible food) that is external to the soybean's relation to the vegan. In other words, the first-order relation between the vegan and the soybean are the sensual qualities the vegan enjoys when eating the soybean, but when the vegan adds moral qualities to the soybean a second-order relation is created. The ethical relation is determined by the effects the vegan experiences from their relation with the soybean. The vegan and soybean relation is immanent in the first order relation but not in the ethical second order relation. Thus, it can be said that the soy bean doesn't *bathe* in the vegan's ethics. The one-sidedness of the vegan-soybean relation entails that not only does the vegan project human moral qualities onto the soybean but they also conceptually capture the soybean by ignoring the possible violence the vegan commits against the soybean and other plant life.

Although metaphorism poses a problem to anthropocentric ethics, including vegan ethics, metaphorism can help us understand what a true ethics of objects will look like. If we were to speculate the possibility of the ethical relations between, for example, a piston and fuel, we would be projecting *our* understanding of violence onto the alien relation between piston's and

fuel's that is inaccessible to human thought. Therefore, if we were to posit an object ethics it would be deferred "to an ethereal beyond, a logic that lives inside of objects, inaccessible from without" (79). Ethical judgement of the other (e.g. tofu, the relation between piston and fuel) proves to be a metaphorism that attempts to understand the Other in *our* ethical terms. For Bogost, the vegan mistakes the animal's "call" for an ethical relation that is coming from the animal's withdrawn essence. Rather, that ethical call is a projection of human values. Ethics is therefore a *hyperobject*, "a massive, tangled chain of objects lampooning one another through weird relation, mistaking their own essences from that of the alien objects they encounter, exploding the very idea of ethics to infinity" (79). This "explosion to infinity" discloses ethics, and our ethical relationship with nonhuman animals, as constant reminder of the inaccessibility and irreducibility of the other. As nonhuman animals *withdraw* from human ethics, our ethical codes must infinitely change to do justice to their irreducibility to our ethical frameworks.

Bogost's critique shows ethic's inherent egoism doesn't necessarily negate ethics, but, rather, forces us to view ethical relations as a projection of our values onto the Other. However, this emphasis on the radical difference and *withdrawal* of the Other is prone to the same complicity to anthropocentric practices as the difference approach. But what does this reveal about OOO and anthropocentric ontologies that deprioritizes the practical and institutional effects of anthropocentrism? Both OOO and the difference approach paralyzes ethics in their shared insistence of difference—Otherness (deconstruction) and the "thing-in-itself"/withdrawal (OOO). I will argue that Calarco's indistinction approach can free us from the ethical paralysis posed by Bogost.

Section 3: Indistinction and "Being Toward Meat"

Bogost's critique of veganism parallels Calarco's critique of the identity approach to animal ethics. They both criticize vegan

ethics that base ethical consideration on the degree of similarity nonhuman animals have to the *quintessential* human (i.e. rational moral agents) and the identity approach's *conceptual capture* of the animal other. But which approach would OOO use if it were to attempt to answer the "question of the animal"?

I would argue that OOO would use an approach similar to the difference-based-approach of deconstruction. As Derrida insists on the maintenance of the human-animal binary despite the underlying relational structure of *difference* that precedes the categories, OOO and Bogost would argue that, despite the multitude of relations that constitute the *carpentry of things*, the object-in-itself is forever inaccessible from all objects—including humans and the object itself. Despite the ontological and ethical differences between OOO and deconstruction (e.g. the ethical call of the Other, ethics as a *hyperobject*), OOO's concept of *withdrawal* and deconstruction's *infinite transcendence* (i.e., inaccessibility) of the object/other views the human-animal distinction in politically and ethically problematic ways. Just as in Calarco's criticism of deconstruction, Bogost's object ethics is equally complicit to the institutional effects of anthropocentrism. OOO liberates objects from conceptual reduction and capture but ignores the material capture, exploitation, and ontological reduction (e.g. meat, research material) of nonhuman animals. Deferred to an "ethereal beyond," human ethical values are limited to the human domain and any attempt to understand the ethical relations of nonhuman animals, if there is even such a thing, are speculative and metaphorical. By resisting any attempts at anthropomorphizing the nonhuman animal, OOO is forced to be silent on how the nonhuman animal views and treats the human. The ontological and epistemological abyss separating the human from the animal commits object-ethics to reinforcing the human-animal distinction and its institutional effects. A vegan ethics that accepts OOO's concept of *withdrawal* can't provide an ethical account of the potential reduction of human bodies to *mere* meat. Under OOO, even if we accept our *becoming-meat* we cannot assume that the rendering of cow, pig, chicken, and other factory farmed animal

bodies to *mere* meat is identical or, at the very least, isomorphic to the rendering of human bodies into *mere* meat.

In "Being Toward Meat: anthropocentrism, indistinction, and veganism", Calarco discusses the indistinction approach to human-animal issues by rethinking the works of Friedrich Nietzsche and eco-feminist Val Plumwood. Calarco provides an analysis of Nietzsche's reversal of anthropocentrism ("man is the cruelest animal") and Plumwood's story of the "shocking reduction" she suffered from "being a unified human subject to being a piece of meat" (424).

Throughout Nietzsche's writings he attempts to naturalize humanity by not placing human beings on the equal status as nonhuman animals (as we find in the identity-approach), but, rather, he argues that "human beings should actually be seen as occupying a *lower* rank than animals" (Calarco 2014, 420). Nietzsche overturns human narcissism and exceptionalism by reinterpreting the characteristics that we typically regard as signs of human superiority (e.g. rationality, language, fear of death) as *actually* being signs of human frailty and weakness. By recognizing the "false rank" we give to ourselves, we will be open to the existence of "other perspectives and openings onto the world" (421). Similar to OOO's flat ontology, Nietzsche urges human thought to seek the radical potentiality in the other-than-human world that "lies *beyond* the human horizon". By displacing human knowledge and values, we are open to "engage in this 'overrich' world" that requires the creation of new ideas, practices, and relations to emerge.

Analyzing the human-animal distinction through the lens of Nietzsche's displacement of anthropocentrism, Calarco argues that this displacement "undercuts the human-animal distinction and places what we call 'humans' and 'animals' in a zone of profound identity called *indistinction*" (423). Unlike the *anthropocentric extension* of the identity-approach, this zone of *profound* identity places animals and humans alike in a *shared space* in which the proprietary attributes/characteristics that determines either categories are dissolved. Pulling from Gilles Deleuzes' writings on

Francis Bacon's "meat paintings," Calarco identifies *meat* as the shared space where humans and animals are "caught up beyond their control in a shared space of exposed embodiment" (423). One may object that this "shared space" is just a zone of unarticulated difference, but by focusing on rearticulating and reestablishing differences we fail to recognize the sites and moments in which differences become irrelevant. More importantly, the emergence of the human as an animal that is no longer exempt from the process of becoming meat displaces human exceptionalism. Thus, this "shared space" should not be seen as an unilluminated site that needs to be resolved by rearticulating new lines of differentiating humans from other animals, but rather as site of transforming our theoretical and practical engagement with *all* fleshy vulnerable bodies. If we differentiate man's reduction to mere meat to, for example, a cow's, we run the risk of providing justification for a strong moral hierarchy and its institutional effects.

Although it's obvious that humans are natural beings, to conceptualize the human as a vulnerable body that could be rendered into meat we must approach ethics and ontology in non-anthropocentric terms. If we don't, we are prone to reproducing the limitations and problems that plague the identity-approach. In order to do so, Calarco provides an analysis of ecofeminist Val Plumwood's "Being Prey".

In "Being Prey" Plumwood provides us with a concrete account of our *becoming towards meat*. She recounts a near-death experience she faced when she was repeatedly attacked by a crocodile while kayaking in the Kakadu National Park. A vegetarian before the attack, Plumwood was able to reorient her place among animals and nature—legitimizing her vegetarianism even further. The visceral experience of being reduced to *mere* meat and prey for another animal exposes the ever-present possibility of our vulnerability as embodied beings. Caught within the jaws of a crocodile, she was thrown into an alien world indifferent to the value of human life. In this alien world of "raw necessity," human and animal life become indistinct in a "joint existence" as *mere* meat. Unlike the inaccessibility of the withdrawn alien worlds of OOO,

Plumwood's glimpse into the alien world of "raw necessity" gave her the joyful affirmation of the "world understood as lacking in human meaning and value but not as lacking in meaning and value altogether" (426). As prey and *mere* meat, human beings "*are*, in principle and as a permanent virtual possibility, meat for others;" but, as she learned from the crocodile attack, human and animal bodies are not exclusively meat. Her simultaneous reduction to *mere* meat and resistance to be reduced to meat discloses the alien/nonhuman experience of being-meat. This ontological contraction of being prey and more than prey, discloses the fact that animals rendered into meat were always already *more* than meat.

Section 4: Response to Bogost's Critique

By using the indistinction-approach, an ethical veganism can be consistent with the idea that "both humans and animals are fundamentally and ontologically edible creatures" (427). We live in an anthropocentric world where the ever-present possibility of being food for another is an ontological impossibility. Likewise, when we reduce animals to *mere* bodies/meat it is impossible for us to view the cow, pig, or chicken as being more than *mere* bodies/meat. But by "world"-traveling to an alien world indifferent to the human-animal distinction, a vegan ethics can still maintain the ideal of respecting animal life without rejecting the ontological possibility of edible bodies.

Bogost's critique of veganism does not include an indistinction approach to veganism, but I can *speculate* that he will argue that the *alien* world of "raw necessity" Plumwood finds herself in is still a *human-all-too-human* caricature of the worlds of rats, birds, fish, and gazelles. Humans can be simultaneously *mere* meat and more than meat, but for Bogost this would be a metaphor. Since the Other *withdraws* from us, it would be a mistake to confuse this metaphor, which is based on our human experience and understanding, for the reality only known by the *withdrawn* Other. But isn't OOO's concept of object *withdrawal* consistent with Plumwood's and Calarco's view that vulnerable bodies reduced to meat actively *resist* this reduction? But why would Bogost insist on

the inaccessibility of the Other in his ethical encounter with the nonhuman and ignore the sites of possible glimpses into an alien nonhuman world?

Due to the ubiquitous effects of the human-animal distinction and anthropocentrism, we are blinded to the fact that animals, nonhuman and human (e.g. minoritarian groups), are not subject to the *bare life* of being *merely* bodies/meat. Although Bogost's ontological motivations insist that the inaccessibility of the object-in-itself is justified, the material reality of the human-animal distinction erases the presence of the animal in many ways—from the rendering of animal bodies to meat to the ritual of sacrificing animal bodies for the advancement of science and medicine. For OOO, one of the primary motivations in formulating a non-anthropocentric ontology is to escape the epistemological and ontological problems that plague correlationism. As evidenced in his critique of veganism, Bogost reduces the political and ethical motivations of veganism to an ontological analysis of intersubjectivity. I'm unsure what the political and ethical motivations of arguing that the soybean or the animal doesn't participate in the vegan's ethical project are. CAS approaches ontology as a means of radically transforming ideas and practices in order to push back the institutional effects of anthropocentrism. My point is not to argue that politics/ethics has philosophical precedence over ontology, or vice-versa, but to argue that Bogost misrepresents CAS and veganism by viewing them on his ontological terms and not on their own terms.

Although this might be beyond the scope of the present discussion, I would like to suggest the possibility of a nonhuman ethics that accepts the ontological framework of OOO and an indistinction approach to ethics. By nonhuman ethics I do not mean the possibility of ethical relations between objects. Rather, a nonhuman ethics would be the human ethical approach towards the nonhuman—living or otherwise. In most relations, metaphorism is used to make the alien familiar by creating a caricature of alien, but in zones of indistinction the familiar is made alien—notably the human becomes sub-or-extra-human. In doing

so, entering the alien gives us a glimpse into an alien nonhuman world. For example, Plumwood's "shocking reduction" made the familiar, her human subjectivity, alien by reducing her subjectivity to a vulnerable exposed fleshy body. This *shared space* of becoming *mere* meat can be located at the threshold of the human and nonhuman animal. But can we say this about the soybean or the gourd? We can metaphorize the suffering plant-life may experience when violently ripped out of the soil to the violent act of beheading a cow or human, but unless we find ourselves in a *shared space* with plant-life we are not obligated to invent new ethical practices.

Metaphorism could be used as a transformative engagement with nonhuman beings. Instead of allowing the difference inherent to metaphorism to deter ethical praxis, metaphorism should be viewed as a site of indistinction—a shared space—that discloses the encounter of the alien/other as an ethical imperative. Metaphorism as a site of indistinction gives us an empathetic ethical framework that gives an inversion of our ethics where the human is not the privileged subject but, rather, the object.

Bibliography

Bogost, Ian. (2012) *Alien Phenomenology, or What It's Like to Be a Thing* (University of Minnesota Press)

Calarco, Matthew. (2014) "Being toward meat: anthropocentrism, indistinction, and veganism" *Dialectical Anthropology*, vol. 38, no. 4, pp.415-429

Calarco, Matthew. (2011) "Identity, Difference, Indistinction." *The New Centennial Review*, vol. 11, no.2, pp. 41-60

Harman, Graham. (2005) *Guerilla Metaphysics: Phenomenology and the Carpentry of Things* (Open Court Publishing)

Harman, Graham. (2002) *Tool Being: Heidegger and the Metaphysics of Objects* (Open Court Publishing)

More Democracy Everywhere

Fernando Cierra

Introduction

In this paper, I will present the political theory of Radical Democracy developed by Ernesto Laclau and Chantel Mouffe. My aim is to do so by outlining the deep conceptual and historical tie their theory has with the Marxist tradition. In doing this I will illustrate why the institutional criticism of their theory misunderstands their project's transformative goals. Namely it misunderstands that their theory seeks to both destabilize what has been taken to be one of the defining aims of democracy, stabilization and security for capitalist markets, and show how this unstable democratic terrain allows for the dramatic transformation of social and political bodies. I will then pose a criticism of the Rawlsian Theory of Justice from this perspective, as well as an attack of a separate consensus approach to democracy outlined by George Vassilev.

The Germs of Radical Democracy in Marxist Theory

In order to fully understand Radical Democracy it is necessary to briefly look at its conceptual roots in the Marxist dialectics of the 20th century. In particular, one must consider how this dialectic responded to the expansion of classical liberal democracy following World War II. Classical liberal democracy maintained many of the commitments it had since first being developed in the Enlightenment. In particular there was the democratic commitment to conceiving individuals as able to govern themselves and collectively will their political demands. Classical liberalism differed most from its inception during the Enlightenment in how the economy functioned, since the time period after World War II marks the creation of international institutions which solidified

and further expanded democratic forms of government in more parts of the world.

The idea of a rational self-governed populace, coming together to will their demands on the body politic, remained a defining aspect in most democratic governments. This idea was proposed most popularly by Jean Jacques Rousseau through his notion of a "general will." For Rousseau the individual was self-interested and free, and as such the individual could will. This freedom was naturally given to man. These same principles extended into the state since it was constituted of these men. The state was thus defined as a collective will, its authority coming from the people. Rousseau articulated the relationship between the state and the people as paternalistic. Rousseau thought that the "family, may be called the first model of political societies" (Rousseau 1920, p. 55). During this same time period an economic shift away from feudalism towards capitalism was changing the way society was structured at an economic level. These new theories of government, being formalized concurrently with the change in economic relations, created a deep tension between democratic politics and a capitalist economy. In particular, there was a tension in maintaining an ever expanding democratization of the state and maintaining the economy in relation to labor, where labor was criticized as being undemocratic. This became especially prominent during the industrial revolution, when capitalism became the dominant economic structure in some parts of the world.

In response to this tension and contradiction in applied democratic practices and institutions, Marx and Engels developed a dialectical theory of history. By this they meant that "The history of all hitherto existing society is the history of class struggles" (Marx and Engels 1906, p. 12). The "class struggle" was for them defined in opposition, describing it as taking the general form of "oppressor and oppressed". This struggle, occurring throughout known history, was a determining factor to the development of history. In particular, the class struggle that developed in capitalism at the time was between the *bourgeoisie* and the *proletariat*. These classes were formed by how capitalism structured these

individuals, with the bourgeoisie having control of the means of production, while the proletariat was exploited by them for their labor which produced and maintained capital. Marx and Engels argued that this exploitative relation would come to a brink at some future point in history. The proletariat would unite and overthrow the bourgeoisie, and in so doing would also move away from a capitalist stage of history. This can be called the proletariat's historical task, taken on by the proletariat in virtue of their class identity. However, as the 20th century came and capitalism only continued to grow stronger, Marx's and Engel's historical theory had to be reconciled with how human affairs had actually developed. Instead of a unified class, what had developed was a fractured proletariat.

From this historical point, Laclau and Mouffe begin to draw from Marxist concepts, developing their theory with an examination of the reaction in Marxism against this dilemma: how to reconcile the smooth linear transition from a capitalist structured society to a socialist society with what seemed like an increasingly more powerful capitalist structured society and a proletariat that did not unify. This historical transition was a move in history that was to be the effect of a united proletariat revolting and toppling the bourgeoisie. The proletariat class was supposed to unite against the capitalist because of their shared interest against capitalist oppression. Here Laclau and Mouffe start with an examination of Rosa Luxemburg. Luxemburg's theory is of interest because she developed the notion of *spontaneism* to explain the lack of class unity and how it came to be actually constituted. Luxemburg thought that class unity was only formed concretely within the context of revolution. Here revolution is taken to be a process of rebellion, and it is only within this context that the proletariat is completely able to take on their historical task. There is for Luxemburg a *plurality* of relations and interest which can constitute the working class outside the capitalist power structure. As such this plurality is not a fully enclosed unified whole, and it is only through spontaneity within a revolutionary context that this plurality of economic, political, and social interest can be consti-

tuted into a cohesive class unity. It is only at this juncture that the proletariat can become conscious of their class position and materialize concretely through action their historical task. Laclau and Mouffe describe this as showing the differing locations, or *subject positions*, political agents can find themselves in. It is a reference to the diversity of the form a struggle may take and the relationship these struggles make among each other to form a unity. There is, however, no natural unity between these struggles as subject positions are so varied and so are formed at points of *antagonism*. They are formed antagonistically because these subject positions are not equivalent to each other, as such a subject position or some set must be favored in order to unify class.

This serves to illustrate for Laclau and Mouffe that, "The unity of the class is...a symbolic unity" (Laclau and Mouffe 1985, p. 5). This is because mobilizations of the proletariat within a revolutionary context for Luxemburg come to have more than one *meaning*. In the revolutionary context this mobilization does not just signify *particular* demands but also signifies the process of revolution in *general*. Luxemburg maintained that this always came to unify class identity but Laclau and Mouffe reject this outcome as necessary. As an example we can quickly look at the Women's March on Washington which took place on January 22nd, 2017. This March originated in the U.S and its particular demands were respecting and demanding expansion of women's rights. The people who participated in this march however were vastly diverse, which is to say that its participants occupied a multitude of subject positions. Yet at this march some of these different subject positions were excluded and the people participating mobilized around a feminist politics. This exclusion however did not destroy those subject positions, these other subject positions were still present and if they had not been excluded they were linked with it to articulate a particular feminist politics. Among the particular demands made by the organizers of the march was an end to racial profiling by police, this demand did not *have* to be made but was done in order to link their feminist politics with the cause of racial minority groups. It is worth noting of course

that there are many poor whites who also face police harassment, but that the organizers excluded these people in their rhetoric. The march at the same time, however, also signified a process of revolution in general in that it was an act of resistance. This other meaning, beyond just its very particular demands, allowed the march to also be articulated outside its nationalist goals, into an international day of resistance. This overdetermination of the meaning of the march is what constitutes the unity that came to be formed; that is, it allowed a notion of feminism to center and constitute collective action.

Now regardless of its ultimate success or failure, all this is just to illustrate what Laclau and Mouffe mean when they say class unity is a symbolic unity. They can take on more than their literal meaning and it is this *overdetermination* that allows a unification of multiple subject positions to be articulated and *hegemonize* around one, or a set of, subject positions. Laclau and Mouffe thus find that Luxemburg can be read as proposing overdetermination as the mechanism for constituting class unity. The logic of spontaneity in Luxemburg's theory still functions alongside the original Marxist notion that the proletariat would unify and revolt. But the logic of spontaneity is novel in that it functions by disrupting this original notion, allowing it to more fully explain the Marxist dilemma. But here Laclau and Mouffe criticize Luxemburg for not abandoning the insistence that what must be constituted through this process is a class unity (i.e. it unifies the proletariat). Since the logic of spontaneity already acknowledges a plurality of subject positions it makes no sense to insist on this logic only ever constituting a class unity.

Laclau and Mouffe draw on two other tendencies they find in Marxist theory which they term *Marxist Orthodoxy* and *Marxist Revisionism*. The Orthodox view attempted to explain this gap between a unified class taking on a specific historical task and the actual world by viewing history as completely abstract. It is abstract because for the Orthodox Marxist the gap between classical Marxist theory and the actual state of affairs could be explained away by placing the actual world within a certain

historical location within their historical theory. For instance if someone were to ask, "Why hasn't the proletariat overthrown the capitalist yet?" the Orthodox Marxist would simply reply, "Because it's not time yet." The Orthodox Marxist appeals to the objective and predictive historical laws found in their theory. The actual state of affairs is then only an *appearance* of those laws' failure. By saying that we are simply not *right now* at the historical juncture in which the capitalist state is to be overthrown, class unity becomes fixed to some abstract future location in a deterministic historical narrative. In contrast to this Orthodox position was the Revisionist position which articulated a firm division, albeit artificial, between an autonomous political sphere and the economic sphere. It is artificial because though politics had a role in advancing Marxist ideology, the economic sphere was still what unified class identity. Revisionism answered the problem of unifying a disparate class with politics. It is participation in the political sphere, usually democracy, which unifies the proletariat and not just their position within the economic sphere.

For the Orthodox Marxist the disunity of class is only a false surface level appearance. There is still for the Orthodox Marxist the notion of history and capitalist structures operating on and unifying class identity. The Revisionist seems to concede that identity is constituted contingently, since it does away with centering class unity solely on economics. But it still maintains that transforming society, by overthrowing the capitalist system, is the special task of the proletariat. Laclau and Mouffe find that even though both acknowledge that class unity does not happen naturally but is contingent, neither can abandon privileging the proletariat completely.

Antagonistic Democracy

I have thus far been discussing Marxism and very little of democracy itself, but here the two meet. Some Marxists, like Eduard Bernstein became aware that organized capitalism along with the expansion of liberal democratic states led to a proliferation of

subject positions. It became clear to Bernstein that because people occupied multiple subject positions the proletariat was no longer just a proletariat. The proletariat was also because of the democratic expansion, through the right to vote, a *citizen*. At the same time because of capitalism's increasingly organized structure and the rising wealth of the proletariat in some parts of the world, the proletariat was also a *consumer*. Here we are confronted with the democratic question: if in a democracy one occupies a diversity of subject positions with no *a priori* unifying force that constitutes a unified political subject with particular political tasks how do people come to act politically?

Laclau and Mouffe ground political action through articulating subject positions, which is to configure these dispersed positions towards a particular aim. Articulation refers to any practice establishing a relation among elements such that their identity—or subject position—is modified as a result. This modification hinges on relations being constituted around *difference* and is discursively made. Discursivity here means not just language but also includes behavioral action. What unifies this discursive formation through articulation is regularity in dispersion. Dispersion unifies because in order to articulate something a point of reference is needed. From this point of reference unarticulated differences can be articulated and form a chain. This point of reference can be, and is for Laclau and Mouffe, dispersion itself, which is to center the reference of unarticulated subject positions on an acknowledgement of its very plurality. This means that articulation in a democratic terrain allows one to both fix and disrupt a system of differences. For example at an address to members of the American Federation of Labor and Congress of Industrial Organizations (AFL-CIO), Dr. Martin Luther King Jr. formed a link between the struggle for civil rights and labor. He stated, "The duality of interests of labor and Negroes makes any crisis which lacerates you, a crisis from which we bleed" (King 1961, pp. 203). By saying this King articulated a chain of equivalence between the two different interests of labor and civil rights. By articulating this equivalence he is also dislocating the racist chains that were articulating the labor

movement at the time. King both fixed and dislocated a system of differences to advance anti-racism. People self-define for themselves and others the political task to take on, and it is not given already simply because of the subject positions in which they find themselves in. Any articulated discourse, then, only partially fixes dispersed identities/subject positions because it is always possible to dislocate and reconfigure them.

Democracy being constituted around difference means it does not privilege one unified subject over another. Democracy instead provides a discursive field which centers around notions of equality or liberty in which people can articulate demands to further expand or retract democracy into other domains. But because subject positions are constituted around difference, they are articulated through exclusion. What it means to occupy some subject position means to be excluded or exclude another. This entails that Democracy operates within an us/them duality. We articulate subject positions and then connect them to other positions by equivocating them around a particular position, like "human rights," for example. In this way it becomes possible to unify a large bloc of people around some cause, but this chain of equivalence never fully collapses one's discrete subject positions into a unifying whole because it is never fully enclosed. This inability to completely unify identity allows articulation to hegemonize this diverse subject positions. An example of a hegemonic articulation would be, for example, the two opposing factions on gay marriage within the gay rights movement. One side making an equivalence between marriage as equality, articulating the issue of gay marriage in terms of "human rights" exposes an inconsistency. However some people in the same movement articulate the very institution of marriage as heterosexist and thus demand it be deserted for being inconsistent with "equality". But it was the latter articulation which succeeded in hegemonizing and thus advancing their agenda. These real world examples show the way identity is always to some degree antagonistic in the Democratic terrain through those it excludes "because every object has inscribed in its very being something other than itself and that as

a result, everything is constructed as *difference*... every identity becomes purely contingent" (Mouffe 2009 p. 21).

Rawls, Institutions, and Radical Democracy

I have given a brief overview of how Laclau and Mouffe came to form their theory through an analysis of the Marxist tradition. In particular, I discussed how their theory developed from analyzing Marxist theory attempting to reconcile class unity when social agents occupy diverse social positions beyond just economic class. How Laclau and Mouffe found that maintaining only class could or should be unified was not consistent with the contingent and antagonistic character that makes up identities like class identity being social/political position-and as such should be abandoned. I did this so that it was clear that Radical Democracy, like the Marxist tradition it draws from, aims not just at maintaining a stable society but to *transform* it as well.

Radical Democratic theory is usually contrasted to another democratic theory, the Theory of Justice advanced by John Rawls. Rawls developed an ideal theory where we should construct the terms of organizing society by starting from the "original position" (Rawls 1980, p. 522). This original position is inhabited by people who are rational but not reasonable. Rawls defines rationality as having the ability to make choices, planning, and being self-interested. Reasonableness is defined as making choices based on morality, or from some non-self-interested position. Lastly he says the people in this original position are behind a *veil of ignorance*, which means that they are unaware of their social positions. This position, Rawls argues, should be the starting point in developing a just society because it has built into its framework justice already.

Democratic society for Laclau and Mouffe operates on a totally contingent terrain, where it is possible for a wide spectrum of differing and antithetical positions to come to power. This terrain operates around an antagonistic relationship between some

group articulated as "us" and another as the "them." Rawls' position is that we can construct the ideal society by doing away with this antagonistic terrain. This move seems to be what makes the Theory of Justice untenable, because once these antagonisms are removed how is one outside the original position to be motivated to act in accordance with it? If antagonisms in the actual world are what constitute the political task we come to take, and the Theory of Justice has none, it cannot be constituted in the actual world. But Rawls' theory, it may be said, is in fact just meant as an *ideal* to be used as the meter in relation to which we can judge whether a society is justly organized, an ideal to be *approximated*. If the ideal state accepts that liberty is just, and so guaranteed to some extent, then it makes no sense to start from a position that lacks antagonism. Actually free political subjects, in virtue of the fact that they are free to engage with each other politically and free to hold opposing political goals, must encounter antagonism. So it makes no sense to attempt to fix the political subject in such an abstract state because it becomes inconsistent with the very state they are trying to form.

One of the major benefits of the Theory of Justice that cannot be said about radical democratic theory is that it works from the outset to form institutions for society. For instance the theory of Justice works from scratch taking the original position to develop basic principles to organize society, and then from that same position works to codify the principles through law and then institutions. Radical democratic theory on the contrary has been said to fail in giving attention to "the need to institutionalize democratic arrangements" (Norval 2007, p. 54). This is certainly true: there is no positive outline to forming and maintaining democratic institutions. In addition, how can institutions which help in structurally maintaining the democratic state be formed, maintained, or function at all if every relation is an antagonistic one? My reply to this is that we have clear examples of antagonisms *in* democratic institutions and *between* them in actuality. These institutions, however, do not fall merely because of antagonism, and some are built explicitly to be antagonistic. For example, the three branches

of American government were explicitly designed to be antagonistic, and within all three there are articulations between actors. For instance, in 2009-2010, while congress was developing healthcare reform, the belief was that, since a new liberal government had just been elected, some form of universal healthcare would be passed. The image which had been cultivated was that the Democratic Party had taken on this task, and since it was in power it would fulfill healthcare reform through legislation. However, as it turned out, some members of this same liberal party did not desire to pass such legislation. There was contention by various members for various reasons. There was, to put it another way, an articulatory failure around the cause for universal healthcare which led to a breakdown between congressional representatives. Yet still, regardless of these antagonisms, this great institution lived on.

Radical democratic theory allows us to make antagonisms more *visible*. It gives us tools so that we can better advance our political goals within a democratic terrain. Rawls original position, and the liberal democratic state which forms from it, is less concerned with change than with security and stability. To be sure that is a goal of both but because the original position wishes to do away with antagonisms to arrive at a solution that is just to all participants involved (i.e. the people), it will fail to be realized. Making antagonisms visible gives people a way to *strategize* and then change social and political bodies; it allows people to advance a plurality of visions for the state. Radical democratic theory positions institutions as too contingent one can say. Yes, but this contingency is a strength and not a weakness.

The Consensus Approach

This antagonistic approach, what some term an *agonist* position, indicates that a consensus-oriented approach to analyzing democracy fails in fully capturing how relationships are formed in democracies. A consensus oriented approach is one in which through some deliberation consensus is reached. How consensus is to be reached can take as its starting point a Lockean state of

nature or Rawls "original position." Though these theories differ, they share in common the idea that antagonism between people can be resolved or should be. Some people like George Vasilev go so far as to claim that, "Rejection of consensus is unsustainable from the very pluralist perspective agonists seek to uphold" (Vasilev 2015, p. 74). For Vasilev, this is the case because a "conceptualization that reduces consensus to systems of signs is inadequate owing to its detachment from human experience" (Vasilev 2015, p. 82). It is my contention that (1) Vasilev misunderstands Mouffe's and Laclau's philosophy of signs by equivocating it with the views held by Michael Foucault, and (2) that his distinction between degrees of consensus does not show that difference is ever resolved but instead that it is itself a discursive articulation which excludes.

Vasilev's argument that the agonist's philosophy of the sign is too abstract for analyzing social relations does not apply to Laclau and Mouffe. It misunderstands their divergence from Foucault. Their Agonist framework "rejects the distinction between discursive and non-discursive practices" (Laclau and Mouffe 1985, pp. 93). Their philosophy then acknowledges that human behavior constitutes discourse along with language. Vasilev seems to think they operate under such a distinction and so "discursivity" does not include human action. No such distinction exists for Laclau and Mouffe unless it is articulated between two positions, but not in the discursive field itself. Language, itself being seen as an action, modifies identities in the world. Discursivity is never just some abstract notion that remains in thought but also something that produces concrete action.

Vasilev's alternative to an agonist democratic theory is to accept that *some* level of reconciliation is possible. For this aim he advances a notion of consensus developed by Dryzek and Niemeyer which distinguishes between two planes of consensus, the *simple* and the *meta*. Simple level consensus is when there is agreement over specifics, such as two people agreeing that murder is wrong. Meta level consensus is a disagreement over specific issues but agreement that the general framework of this disagree-

ment can be acted on politically. For example, Republicans and Democrats disagree on how to handle specific issues like taxes, but both agree that taxation is politically viable. These planes can be further subdivided into *normative*, *epistemic*, and *preference* consensus. Normative consensus concerns values, epistemic consensus concerns which sort of epistemological views are valid, and preference consensus concerns "what should be done" (Vasilev 2015, pp. 83). At the simple level of consensus, there is agreement on at least one of these other three forms, while at the meta level there is no agreement whatsoever over the specific positions held. This multi-dimensional perspective of consensus illustrates for Vasilev that antagonism between political agents really does, at least sometimes, collapse.

As an example Vasilev says, "Catholics value the traditional heterosexual family... radical feminist fiercely reject this ideal as the source of women's subordination. Nevertheless, both support bans or restrictions on pornography." This demonstrates that though both the Catholic and the radical feminist share no normative consensus they share a preference consensus. However the consensus claimed to be shown in this example is only possible by positing "radical feminist" as a unitary subject that necessarily prefers to advance this specific policy goal. This claim however shows an articulation being formed discursively and linking two antagonistic parties. This link being created by constituting the identity of "Catholics" and "radical feminists" through exclusion, excluding Catholics who are opposed to the heterosexual family and radical feminists who advance a notion of sex positivity incompatible with banning pornography. There is no consensus as described by Vasilev, just another site of antagonism that is being articulated.

In this site one's difference is not resolved but in fact excluded from the articulation of "radical feminist" or "catholic." Vasilev wants to maintain that compromise can be reached through consensus and difference reconciled, if not always absolutely. But he fails to give a compelling reason for us to believe that difference, and as such the antagonism between two parties, has disap-

peared. Instead he has merely shown another method by which one can articulate two differences and form a hegemonic totality.

Conclusion

Rawls' Theory of Justice and the consensus approach in general are very intriguing proposals. But in the main I think they are found to be too idealistic. A political fiction can be useful but when it discards a basic precept, like antagonism, it already becomes too ideal. A radical democratic theory begins and finds power in what consensus wishes to do away with and because of that it should be seen as a more tenable theory from the outset.

Bibliography

Rousseau, Jean-Jacques. (1920) *The Social Contract & Discourses* (E.P. Dutton & Co)

Marx, Karl, and Engels, Frederick. (1906) *Manifesto of the Communist Party* (Chicago: Charles H Kerr & Company)

Laclau, Ernesto, and Mouffe, Chantal. (1985) *Hegemony and Socialist Strategy Towards a Radical Democratic Politics* (London: Verso Books)

King, Martin Luther Jr. (1961) *A Testament of Hope: The Essential Writings and Speeches of Martin Luther King Jr.* (Harper Collins: Harper One)

Mouffe, Chantal. (2009) *The Democratic Paradox* (London: Verso Books)

Rawls, John. (1980) "Kantian Constructivism in Moral Theory" *The Journal of Philosophy* vol. 77, pp. 515-572

Norval, Aletta J. (2007) *Aversive Democracy* (Cambridge University Press)

Vasilev, George. (2015) "The Uneasy Alliance between Consensus and Democracy" *The Review of Politics* vol. 77, pp.73-98

Contributors

Andre Agacer. B.A. Philosophy, California State University, Fullerton. Social political philosophy, Continental Philosophy, Critical Animal Studies, Posthumanism. Andre plans to apply to PhD. programs in the near future.

Fernando Cierra. Currently pursuing a B.A in Philosophy at California State University, Los Angeles. Social Political Philosophy, Ethics, Feminist Philosophy.

Van Doan. B.A. Philosophy, California State University, Sacramento. Ethics, Meta-ethics, Applied Ethics, Philosophy of Law, Aesthetics, and Metaphysics. Van plans to teach at a community college after earning his M.A.

Taylor Dunn. B.A. Philosophy, University of Missouri. Metaphysics, Philosophy of Language, and History of Philosophy. Taylor will be applying to Ph.D. programs in 2018.

David Fonth. Currently pursuing a B.A. in Philosophy, California State University, Los Angeles. Philosophy of Language, Logic, Metaphysics. David plans to apply to Ph.D. programs in the fall of 2017.

Craig Laubach. B.A. Philosophy and East Asian Studies, Pennsylvania State University; M.A. Teaching (Single-Subject English), Point Loma Nazarene University. Social and Political Philosophy, Existentialism, Phenomenology, Aesthetics. Craig is currently a teacher and musician, and he intends to continue researching Marx, Nietzsche, and Deleuze and Guattari.

Jonathan Poh. B.A. Economics, University of California, Los Angeles. Metaphysics, Ethics and Philosophy of Mind. Jonathan is looking towards pursuing a career in the legal field, and hopefully teaching in some capacity thereafter.

Eduardo Salazar. B.A. Philosophy and Minors in Comparative Literature and Comparative Religion, California State University, San Diego; M.S. Nursing, California State University, Fullerton. 19th and 20th Century Philosophy, subjectivity, Critical Animal Studies, Philosophy of Language, Consciousness Studies. Eduardo plans to teach in the fall 2017 semester and apply to Ph.D. programs in the fall of 2018.

Cameron Takeda. B.A. Philosophy, University of California, Berkeley. Metaphysics, Philosophy of Fiction, Ethics. Cameron intends on teaching Philosophy at the Community College level.

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