



2021 Annual Undergraduate Philosophy Conference Journal

Conference Organizers: Paul Jin, Andy Liu, Lucie Lu, Ryan Wang, Pacy Yan, Lin Ye, Heyan Xiao, Chloe Zhu

Review Committee: Nicole An, Joshua Arrayales, Jaime Andres Fernandez, Kira (Yuxin) Liu, Krishang Nadgauda, Juan N. Pava, Julian Perilla, Rachel Roberts, Ali Sina Sahin, Yingcan Sun, Pacy Yan

Special thanks to the NYU Department of Philosophy

5.22 - 5.23
2021

Philosophizing (in) Uncertainties

 @ NYU
Annual Undergraduate
Philosophy Conference

Table of Contents

Pornography and Authority Silencing

by Arabella Adams, Wellesley College

2-10

Against Singer: Ethics and the Limits of Obligation

by Gabriel Sanchez Ainsa, University of Chicago

12-23

What Counts as Reliving?

by William Choi, Johns Hopkins University

24-32

Arguments Against Arguments Against Affirmative Action (AAAAAA)

by Long Do, Brown University

34-44

Empedocles' Notion of Universal Kinship as an Important Consideration In Ecological Thought

by Josh Doland, University of California, Los Angeles

45-58

Can a Kantian Count? Kant on Aggregation

by Andrew Garber, Harvard University

60-69

Double Standards of Rape: The Paradox of She Said, He Said Cases

by Charlotte Goldberger, McGill University

71-78

Carnap and Impredicative Definitions

by Junyan Jiang, New York University

82-91

Which Came First, the Camera or the Photograph?

by Aashutosh Mukerji, Boston University

92-97

5.22 - 5.23
2021

Philosophizing (in) Uncertainties

 @ NYU
Annual Undergraduate
Philosophy Conference

Pornography and Authority Silencing

Arabella Adams

Wellesley College

May 2021

Abstract

In “Speech Acts and Pornography,” Jennifer Hornsby argues that pornography silences women by interfering with men’s ability to recognize a woman’s intention to refuse sex. According to Hornsby, *reciprocity* (that the hearer understands what the speaker is trying to communicate) is necessary for a speech act to be successful. If a man does not understand that a woman intends to refuse sex, then the woman *fails to refuse* and is therefore silenced because she cannot secure reciprocity from her listener. In this essay, I use the framework put forward by Hornsby in conjunction with Mary Kate McGowan’s article, “Debate: On Silencing and Sexual Refusal,” to argue for another kind of silencing known as *authority silencing*. *Authority silencing* occurs when the hearer takes a speech act to be a misfire because they do not recognize the speaker’s authority to perform the given speech act. For this argument, I propose a revised definition of reciprocity that I call *reciprocity-plus*, which requires that the speaker’s authority be recognized in addition to comprehension. If men fail to recognize women’s authority to refuse sex, then all sexual refusals by women will be taken as misfires. Since their speaker authority is not recognized, women are unable to issue a successful refusal and are therefore silenced. First, I

explain why Hornsby's aforementioned example is considered silencing, and then discuss how failure to recognize speaker authority functions. Ultimately, I conclude that failure to recognize speaker authority indeed constitutes silencing.

Pornography and Authority Silencing

In her paper “Speech Acts and Pornography,” Jennifer Hornsby claims that pornography silences women¹ (Hornsby). According to Hornsby, consumption of pornography interferes with men’s ability to recognize a woman’s intention to refuse sex. If a man is unable to recognize that a woman intends to refuse sex, then according to Hornsby, the woman has failed to refuse² and is therefore silenced. Using Hornsby’s framework, I argue for the existence of another kind of silencing known as *authority silencing* (McGowan). In *authority silencing*, the hearer understands that the speaker intends to refuse³, but they take the speech act⁴ to be a misfire because they do not believe that the speaker has the authority to refuse. In this paper, I argue that failure to recognize the speaker’s authority constitutes silencing. First, I will explain why Hornsby’s example of failure to recognize the speaker’s illocutionary intention to refuse is considered silencing. I will then explain how failure to recognize speaker authority functions, and will ultimately conclude that failure to recognize speaker authority also constitutes silencing.

In order to discuss illocutionary silencing, it is important to first define what illocution is. According to J.L. Austin, speech can have three different forces: *locutionary*, *perlocutionary*, and *illocutionary*. The *locutionary* force of an utterance is the content or the meaning of what is said (Austin).

¹ In this paper, I will use the example of a man and a woman, but it is important to note that these issues do not only take place between men and women.

² Hornsby believes a concept called *reciprocity* (discussed further below) is required for successful illocution. If reciprocity fails, the speaker has failed to perform the intended illocutionary act.

³ Or perform any other illocutionary act. In this case, I will use refusal as an example given that we are discussing the sexual refusal case.

⁴ In the case of pornography and sexual assault, the speech act is refusal.

Perlocution refers to the causal effect the utterance has via the hearer's⁵ recognition of what is said (McGowan). For example, if I tell a story and you are inspired, your being inspired is the perlocutionary effect of my utterance (Austin). Finally, *illocution* is the action constituted by an utterance. When I say, "I apologize," I am not only describing what I am doing but the words "I apologize" actually constitute the act of apologizing⁶ (McGowan).

Some philosophers argue that *uptake* is necessary for illocution. *Uptake* is the hearer's appreciation of the intended illocution of the speaker (Bird). The argument is that if the locution does not secure uptake, then it fails to be the illocutionary act intended by the speaker⁷. For example, if an actor on stage tries to warn the audience of a real fire in the theatre, the audience may think he is still acting. His illocutionary intention in yelling "Fire!" is to warn the audience of the danger, but if they believe he is still acting, they will not take his utterance as a warning (Langton). Similarly, Hornsby believes that the hearer has to recognize that you are trying to tell them a certain piece of information for you to successfully communicate the information. If the hearer does not recognize that you are trying to tell them this particular thing, then you have not fully succeeded in telling it to the person. Therefore, Hornsby asserts that telling invokes what she calls *reciprocity*. *Reciprocity* is secured when people are able to recognize another person's speech as it was meant to be understood, therefore ensuring the

⁵ Terms like "listener" and "hearer" are ableist but they are used in this paper to be consistent with the terms the literature uses.

⁶ "I apologize" constitutes an action as opposed to a different utterance like "I'm reading." When you say "I'm reading," you are describing what you are doing, but saying "I'm reading" does not constitute the act of reading. However, in saying "I apologize," there is not only the locution or the meaning of those words, but those words themselves constitute the act of apologizing.

⁷ Failure to secure uptake of the speaker's illocutionary intention means that the illocutionary act does not take place.

success of the speaker's attempts to perform speech acts. If reciprocity is obtained, speakers can do things by simply being heard and understood as doing them (ie. apologizing, warning).

In "Speech Acts and Pornography," Hornsby argues that pornography silences women. We can define silencing as illocutionary disablement: something disables the speaker from being able to perform an illocutionary act (McGowan). It is also important to note that silencing must be systematic,⁸ that is, "cumulative" and pervasive (Hornsby). According to Hornsby, *reciprocity*, or that the hearer understands what the speaker is trying to communicate, is necessary for a speech act to be successful. Hornsby believes consumption of certain kinds of pornography⁹ undermines *reciprocity* by interfering with men's ability to recognize a woman's intention to refuse sex (Hornsby). Because the man does not recognize the woman's intention to refuse, due to practices observed in certain kinds of pornography, there is no reciprocity of the woman's refusal. As a result, if the man does not understand that the woman intends to refuse, then according to Hornsby the woman *fails*¹⁰ *to refuse*. She is therefore silenced because she is unable to successfully refuse (Hornsby).

Hornsby's framework can be used to discuss another kind of silencing known as *authority silencing*. *Authority silencing* is defined as silencing that occurs when a hearer fails to recognize the speaker's authority to perform a certain illocutionary act (McGowan). However,

⁸ If I yell "Duck!" and you don't understand I mean "Get down," Hornsby would argue I have failed to perform the act of warning because I didn't successfully communicate a warning. However, I would not be *silenced* because this is an isolated incident - there is nothing systematic about this communicative interruption.

⁹ Hornsby, Langton, and MacKinnon are specifically referring to pornography in which women are degraded and/or treated violently, and portrayed as enjoying this treatment.

¹⁰ Although she may have said "No," the woman fails to *successfully* refuse because there is no reciprocity, and according to Hornsby reciprocity is required for illocution.

some question whether this communicative issue counts as a form of silencing. I argue that failure to recognize a speaker's authority *does* indeed constitute silencing. To demonstrate this, I will discuss the issue of speaker authority in the context of pornography and sexual assault cases. For the purpose of this paper, I am taking it to be true that certain speech acts require authority. For example, in order for someone to marry a couple through the speech act of pronouncing them married, they must be a member of the clergy, judge, or another legally ordained individual with the authority to marry people. In the case of sexual assault, I take McGowan's assertion—that everyone has authority over their own body simply by virtue of being a human being and can therefore decide who has “sexual access” to their body—to be true (McGowan).

Nonetheless, the authority of the speaker is not always recognized. Consider the case where a woman refuses sex and a man understands¹¹ she is intending to refuse, but the man does not believe the woman has the *authority* to refuse him. The man recognized the woman was intending to refuse sex, but he did not believe she had the authority necessary to perform the illocutionary act of refusing. Therefore, the man takes her refusal as a “misfire.” As McGowan asserts, this is comparable to an employee trying to order their boss to give them a raise. The boss would recognize the employee's intention to order them to give them a raise, but the employee would still fail to order because they lack the authority necessary to do so (McGowan). The employee's utterance would therefore be considered a failed attempt to order—a misfire. In the case of sexual assault, however, the woman *does* have the authority necessary to refuse sex,

¹¹ Unlike in Hornsby's case, in this case, the man understands the woman is refusing. The issue is that the man does not believe that the woman has the *authority* to refuse.

so the man *mistakenly*—i.e., incorrectly—takes her refusal to be a misfire. Although a woman has authority over her own body and therefore has the authority to refuse sex, if the man does not believe she has the authority to refuse sex, he will consider her refusal to be a misfire (McGowan). I argue that this failure to recognize the woman’s authority constitutes silencing.

For my argument, I will use an adaptation of Hornsby’s definition of reciprocity. Hornsby states there is reciprocity when the hearer is capable of recognizing the speaker’s speech as it was meant to be taken (Hornsby). According to Hornsby, you do not communicate if your hearer doesn’t recognize what you’re trying to communicate. In the example above, there would be reciprocity according to Hornsby’s definition because the man recognized the woman’s intention to refuse. However, the refusal was still taken as a misfire because the man did not believe the woman had the *authority* to refuse¹². In response, I propose a revised definition of reciprocity that I will call *reciprocity-plus* based on McGowan’s assertion that *the hearer’s recognition* of the speaker’s authority is crucial for a refusal to succeed (McGowan). Therefore, I believe reciprocity-plus should include that the speaker’s authority must be recognized in addition to the illocutionary act being recognized.

Under my revised definition of reciprocity, there is *no* reciprocity in the case of the man who is mistaken about a woman’s authority to refuse sex. The man failed to recognize the woman’s authority, so her refusal was mistakenly understood to be a misfire¹³ when in fact she had the authority to refuse. Therefore, women are silenced because certain kinds of pornography

¹² However, she did have the authority to refuse! The woman had the authority to refuse sex and she did, so her refusal was successful. The problem is that her illocution is not recognized as successful by the man.

¹³ The woman *can* and *did* refuse successfully. She has the authority to refuse, it just isn’t *taken* as a refusal.

undermine reciprocity-plus by interfering with men's ability to recognize a woman's authority to refuse sex (McGowan). If men cannot recognize a woman's authority to refuse sex, there is no reciprocity—under my definition— of the woman's refusal and therefore her refusal becomes “defective” due to the man's inability to recognize her authority over her own body (McGowan). She is therefore silenced because she is unable to successfully refuse.

One objection is that the woman is *not* silenced because the man understood she was intending to refuse, even though her authority was not recognized. If she managed to communicate her intention to refuse well enough that the man understood it, then it can be argued that she was not silenced. However, I argue that this example still constitutes silencing. We previously defined silencing as illocutionary disablement: something disables the speaker from being able to perform a certain illocutionary act (Hornsby). According to Hornsby's definition of silencing, this case certainly constitutes silencing. If men fail to recognize women's speaker authority, then there will never be reciprocity according to my definition of the term. They will recognize the intention to refuse, but *think the woman didn't manage to succeed in refusing* because she didn't have the authority (McGowan). This means all sexual refusals will be taken as misfires, even though women *do* have the authority to refuse. If there is no reciprocity, and as a result, all sexual refusals are taken as misfires, then women will fail to have their successful illocution recognized. If the woman is unable to perform a successful refusal due to the man's failure to recognize her authority to refuse sex, she is illocutionarily disabled and, according to Hornsby's definition, she is silenced (McGowan). Therefore, the issue with speaker authority caused by the consumption of certain kinds of pornography leads to a systematic illocutionary disablement for women which constitutes silencing.

Based on Hornsby's framework and her assertion that pornography undermines reciprocity—which leads to women being illocutionarily disabled and therefore silenced—I have argued for another type of silencing known as *authority silencing*. *Authority silencing* depends upon a revised definition of reciprocity that I have called reciprocity-plus, which includes that the authority of the speaker must be recognized by the hearer for an illocutionary act, such as a refusal, to succeed. Authority silencing occurs when a speaker's authority is not recognized, and as a result, reciprocity-plus is not obtained. In this case, a man thinks that the woman didn't manage to succeed in refusing because she didn't have the authority to when she did in fact have the authority to refuse and refused successfully. The woman is capable of successfully refusing, but she is illocutionarily disabled (and therefore silenced) because her successful illocution is not ever recognized due to the man's failure to recognize her authority as a speaker. Therefore, I argue that failure to recognize speaker authority constitutes silencing, specifically authority silencing.

Bibliography

Austin, J. L. *How To Do Things With Words*. 2nd ed., Harvard University Press, 1975.

Bird, Alexander. "Illocutionary Silencing." *Pacific Philosophical Quarterly*, vol. 83, no. 1, 2002, pp. 1–15., doi:10.1111/1468-0114.00137.

Hornsby, Jennifer. "Speech Acts and Pornography." *Women's Philosophy Review*, no. 10, 1993, pp. 38–45., doi:10.5840/wpr19931021.

Langton, Rae. "Speech Acts and Unspeakable Acts." *Philosophy and Public Affairs*, no. 22, 1993, pp. 293–330., doi:10.1093/acprof:oso/9780199247066.003.0002.

McGowan, Mary Kate. "Debate: On Silencing and Sexual Refusal." *Journal of Political Philosophy*, vol. 17, no. 4, 2009, pp. 487–494., doi:10.1111/j.1467-9760.2009.00346.x.

Against Singer:

Ethics and the Limits of Obligation

Gabriel Sanchez Ainsa

University of Chicago

May 2021

Abstract

In this essay, I respond to Peter Singer's 1972 essay "Famine, Affluence, and Morality", in which he argues that individuals and governments in developed nations have a moral obligation to give our financial resources to alleviate suffering and death worldwide to the point of *marginal utility*—the point at which, were I to provide more aid and efforts, I would cause more bad to myself and others as I am alleviating. In the first part of this essay, I will reconstruct Singer's argument. In my interpretation, Singer defends the moral necessity of marginal utility through the tacit assumption of two significant meta-ethical theses: a "law conception of ethics" (i.e. that morality is a matter of norms in such a way that agents experience the normative force of moral judgements as if from the law) and "the principle of impartiality" (i.e. that morality is agent-neutral as demonstrated by his drowning child thought experiment). I will then proceed to provide some intuitive rounds for rejecting his meta-ethical views and examine the extent to which a different meta-ethics leads to a shift in Singer's practical conclusions. I will argue that once we abandon those commitments, although a rational agent may still find it desirable to

provide much more material aid than most do now, Singer's point of marginal utility does not become ethically necessary.

Against Singer: Ethics and the Limits of Obligation

Having lived through a global pandemic, our generation cannot deny the deep interconnection of the worldwide human community; still, how come that I, living in the United States, have significantly greater access to health-care and the COVID-19 vaccine than my grandfather in rural Mexico? Can this inequality be morally justified? If not, what do the global rich owe to the global poor? In his famous essay “Famine, Affluence, and Morality” Peter Singer answers, a lot more than we currently do. Even when we do not consider the question of reparations and ongoing exploitation, an increase in foreign aid and philanthropic efforts by developed nations could greatly alleviate suffering and death across the world. Singer claims that it follows that we morally ought to provide aid to the point of *marginal utility*—the point at which, were I to provide more aid and efforts, I would cause more bad to myself and others as I am alleviating.¹ In the first part of this essay, I will reconstruct Singer’s argument. In my interpretation, Singer defends the moral necessity of marginal utility through the tacit assumption of two significant meta-ethical theses: a “law conception of ethics” (i.e. that morality is a matter of norms in such a way that agents experience the normative force of moral judgements as if from the law) and “the principle of impartiality” (i.e. that morality is agent-neutral as demonstrated by his drowning child thought experiment). I will then proceed to provide some intuitive rounds for rejecting his meta-ethical views and examine the extent to which a different meta-ethics leads to a shift in Singer’s practical conclusions. I will argue that once we abandon those commitments, although a rational agent may still find it desirable to provide much more material aid than most do now, Singer’s point of marginal utility does not become ethically necessary.

Singer’s argument is relatively straight-forward and *prima facie* sound and valid:

¹ Peter Singer, “Famine, Affluence, and Morality” *Philosophy and Public Affairs* 1, no. 3 (Spring 1972)

- (1) If x is an evil, and if there exist means for A to prevent x , and if preventing x does not sacrifice another significant moral value, A morally ought to prevent x
- (2) Unnecessary suffering and death are evils
- (3) There exists unnecessary suffering and death across the world
- (4) There are means for persons and groups in developed nations to prevent unnecessary suffering and death across the world
- (5) *Therefore*, persons and groups in developed nations morally ought to employ their wealth and power to prevent unnecessary suffering and death across the world

When Singer wrote this paper in 1972, he was concerned about the famine in East Bengal, which was the subject of (3). I suggest that in 2021 the scope of this argument includes not only famine worldwide, but also the refugee crisis caused by the Syrian civil war, the humanitarian crisis in Yemen, the Climate Crisis, the COVID-19 pandemic, global poverty and famine, etc.

Furthermore, I will grant to Singer and the ‘effective altruists’ he has inspired that most people in developed nations have the financial means to alleviate much of this unnecessary suffering. It is just true that most of us could choose to give our excess wealth to save as many lives as possible rather than spend it on coffee or better education. Western governments, too, have the resources required for significant foreign aid and policy approaches, for example, to share vaccine formulae across the world and actively fight Climate Change much more rigorously. However possible, this better world we can all envision has not become actual. As Singer accused Western countries in 1972 of moral failure, this accusation stands now.

I take the last two premises for granted and shall focus on the first two. (2) is uncontroversial—even Stoics, famous for rejecting that anything but vice is an evil, concede that suffering and death are to be avoided, whether in oneself or others. Note that (2) is a value-judgement, that is, a simple, descriptive statement about what is good and bad. On the other hand, (1) complicates matters. Of the three conditions, only the latter seems analytically interesting and not a mere empirical question. “if preventing x does not sacrifice another

significant moral value” contains the ambiguity of what constitutes “significant moral value.” Singer distinguishes between two possible interpretations, the “strong” and “moderate” versions, both of which show that most of us morally ought to do more. Singer himself prefers the “strong” version which interprets premise (1) as,

(1a) “If x is an evil, and if there exist means for A to prevent x , and if preventing x does not cause another evil y such that y is a greater evil than x , A morally ought to prevent x .”

The strong version inevitably leads to the obligation of marginal utility, that is, we morally ought to give our resources to the point that, were we to give more, we would cause more harm than good.

Now, note that this interpretation presupposes that x and y are commensurate. It is not hard to imagine, I suppose, that a person who gives this much will lead a relatively aesthetic lifestyle. Of course, not everybody will concede that all morally relevant values are commensurate (e.g. Kantians and moral absolutists), and part of Singer’s aim is to convince as many as possible to give to aid.

Hence, Singer offers a “moderate” interpretation to which most will assent:

(1b) “If x is an evil, and if there exist means for A to prevent x , and if preventing x does not cause any evil or violate any obligation, A morally ought to prevent x .”

Even moral absolutists can agree to this. This interpretation is moderate because whoever accepts it will feel obliged to donate as much as possible but not necessarily (though possibly) to the point of marginal utility. It seems to me that Singer correctly asserts that any interpretation carries a much more heavily more demanding imperative than most people and governments currently fulfill, including a drastic redistribution of wealth and power across and within nations.

I point out this distinction to bring attention to how far-ranging Singer’s argument is. He wants to convince as many people as possible that they can do more to make this the best of all possible worlds. He is like Diogenes the Cynic standing before the Athenian crowd, showing us how we all are moral failures and should be better. For this reason, although a consequentialist

himself, in this paper, he declines to openly commit himself to any moral theories or meta-ethical views. However, in my reading of his paper, it is his Diogenic inclinations which betray his meta-ethical commitments. In the first place, he openly criticizes our cultures for our confused moral categories of “duty” and “charity”, which his arguments “upset” (235):

“It follows from what I have said earlier that we ought to give money away, rather than spend it on clothes which we do not need to keep us warm. To do so is not charitable, or generous. Nor is it the kind of act which philosophers and theologians have called “supererogatory”—an act which it would be good to do, but not wrong not to do. On the contrary, we ought to give the money away, and it is wrong not to do so.”

The moral worth of supererogatory actions, those which are good to do but not obligatory, is an open question. At a surface level, the difference between labeling an action as “obligatory” versus “supererogatory” amounts to a motivational issue—a speaker who says that to x is a duty and wrong to fail to x will have greater force on her audience than one who says that x is charity and good to do but not wrong not to do. Imagine if the President talked about paying our taxes in the latter manner. But, I suggest, there is a deeper meta-ethical judgment at play. The difference in content between the proposition “to x is right and not to x is wrong” and “to x is good” is that the former employs predicates which imply norms while the latter employs those which imply value and practical desirability. This is to say, which concept, “good” or “right,” should constitute ethical judgements? Of course, Singer employs both concepts, for Premise (1) and (2) both are judgements, but the former evaluates the goodness and badness of such-and-such while the latter the rightness of an action. Nevertheless, Singer still considers that we have arbitrarily and wrongly made “goodness” the main attribute describing foreign aid and philanthropy rather than “right” and we should change these harmful categorizations.

I, on the contrary, argue that they are “good actions” rather than “obligatory actions.” On the surface, I agree that “You ought to prevent evil” compels an audience more than “It is good for you to prevent evil”—the former conveys the so-called ‘pull of normativity,’ that is, we

experience that judgement as if it were mandated by the law. But its force does not make it true. The experience of normativity, as Elizabeth Anscombe suggests, belongs to a ‘law conception of ethics’ inherited from the Abrahamic tradition which lost its meaning with the Enlightenment secular project.² This tradition imagines morality as a set of commandments from God, analogous to the decrees of a king. Christians, moreover, create the concept of ‘sinful acts’—for them, the meaning of “ought” is that if I do not do as I ought, I will be eternally punished. Anscombe argues that without a divine law-giver setting obligation, as it became the case of moral thinking in Europe post-Enlightenment, “ought” becomes meaningless. And in fact, reflection on the word “ought” suggests as much—can we actually pinpoint its meaning? One might argue that we can give “ought” meaning through contracts, that if I promise to do *x*, then I ought to do *x*. After all, this approach does not trace the concept of obligation back to a God. Still, I argue, there is something worse hiding behind this “ought”—a punisher. Why is it that people mock the UN’s decrees? Because they are not enforceable, no punishment. Laws and justice, on the other hand, require a penal code, punishment. If the word “ought” requires such a system of contracts, then it also requires some sense of punishment, that the person for whom you are obliged to do *x* can somehow punish you for not doing *x*, even if merely through shaming and pointing. Moral obligations, therefore, are mere analogies of the law, and without a God to serve the role of judge and punisher, or without some form of moral realism justifying that this “ought” belongs to the categories of nature or whatever, philosophers from Kant to Singer employ reference-less and empty concepts. It follows that to assent to Singer’s argument, we require a certain world-view under which the moral “ought” has meaning and force. This is

² Elizabeth Anscombe, “Modern Moral Philosophy,” in *Philosophy*, Jan., 1958, Vol. 33, No. 124 (Jan., 1958), pp. 1-19

not a mere linguistic peeve—behind his apparently innocent syllogism lies a whole metaphysics of morals or something worse.

As a close-reading of Kant's *Groundwork* or Mill's *Utilitarianism* or mere reflection on the concept of "obligation" would show, this approach to morality brings a second meta-ethical view, a commitment to impartiality in ethical thinking. Singer openly states that premise (1) implies "impartiality, universalizability, equality, or whatever" (232). This principle of impartiality or agent-neutrality has a long tradition in Western moral philosophy from Adam Smith to Rawls. Indeed, (1) contains a statement of moral obligation which does not register any individualities of the agent in so far as she is a person with a certain character, status, relationships, or world-view. Of course, the person must meet certain preconditions (i.e. possession of certain material and social goods), but the obligation itself does not causally or essentially depend on who I am (for example, "It is good for me to buy my sister a birthday gift because I am her brother and I love her" seems quite different from "I ought to give 10% of my income to effective altruist causes because anybody who can give 10% ought to do and I happen to be such a type of person"). The 'moral point of view' in which Singer's agents deliberate demands literal selflessness.

The principle of impartiality has two important functions in Singer's argument. First, it shapes how much I should give—the principle of impartiality allows Singer to arrive at the point of marginal utility. The point of marginal utility is that at which I would cause more harm than good if I gave more. By this we do not mean "more harm than good to myself or the things I care about" but "to the world at large." To not respect impartiality is to say that I could introduce personal and arbitrary preferences in moral deliberation; if I did, I could justify causing more good to myself and loved ones at the expense of the world at large. The second function is to

determine in which way and how we should give. This is brought out by the Drowning Child thought experiment. Suppose that Jake is walking by a pond and sees a child drowning there. Would we not think that Jake is morally required to save the child, no matter how expensive his suit is? But, Singer cleverly remarks *a la* Diogenes, is there not a child dying from famine and war somewhere in the world right now whose life we could save by giving our excess resources to humanitarian or activist causes? Will future generations not suffer much more from Climate Change than we enjoy our fossil fuels? Singer proudly concludes that proximity and distance do not matter from the moral point of view, showing that we have duties to others whom we never meet. But he also, even if he does not remark it in this paper, shows that we arbitrarily choose which drowning children we save. There is no difference, he would say, between the lives of those who die at the hands of police brutality in the US from those who do so at the hands of famine in East Bengal or drown from rising sea levels in the Pacific Islands. Would all the money that many of us gave to Black Lives Matter and bail funds not have done more good in the pockets of effective altruists and global activists? Therefore, the principle of impartiality is necessary for Singer's argument and incredibly consequential if we accept it.

The argument thus far has shown that Singer's conclusion relies on a law conception of ethics and a principle of impartiality. As I read Singer's paper, these meta-ethics are more necessary than they appear. However, I argue that we can reasonably see ourselves as the sort of beings for whom these meta-ethical assumptions do not apply. Human life requires neither a law-conception of ethics nor the principle of impartiality. In the remainder of this paper, I shall invite the expansion of the range of Singer's "significant values" to consider what I call "*practical values*," by which I roughly mean all things which contribute to a good life for us, whether of moral, aesthetic, social, physical, or hedonic kind. This practical standpoint allows for

the evaluation of action without a law-conception of ethics or a principle of impartiality. I will suggest that if we grant that agents act by deliberating from this point of view, the conclusions of Singer's arguments also change.

Firstly, we only need to look around to realize that there are a plurality of reasons which justify action. We act according to our own projects, self-identity, and notions of what is good for us. While the concept of "ought" characterizes the judgements made from the moral standpoint, that of "value" or "goodness" characterizes the beliefs adopted in the practical standpoint. The former carries a dimension of normativity as a force, while the latter that of desirability. Ethical theories with this latter point of view, like ancient Greek and Roman eudaimonisms, do not require a law-conception of ethics. It is clear, too, that the practical standpoint does not have agent-neutrality or the principle of impartiality. All appropriate actions result from my personality and what is good for me.

The Singerian objects, "is this not egoism?" I argue that genuine other-concern can easily be incorporated into this view, and often is. Many of us who actively tried to contribute through financial means to the Black Lives Matter protests in the Summer of 2020 did it out of a sense of justice, but most of us did not do so out as impartial agents maximizing the good. We cared about it at a personal level. The point here is that a rational person can easily be persuaded that the demands of justice and generosity are justifiable projects, and often for their own sake. Deontologists like John Rawls will say that moral action is independent of goodness; consequentialists since G.E. Moore will say that moral action is only a function of things which are good; the view which I advocate, instead, claims that just and generous actions *are* good.

Some might claim that my view does not have practical consequences. On the contrary, I claim that once we inhabit the practical standpoint, Singer's "strong" conclusion that we ought to

donate to the point of marginal utility turns into a “soft” one that we should donate as much as our rational plans of life allow and suggest. The strong and soft versions are not necessarily different in quantity—for some, my soft version might still lead to marginal utility. It does, however, challenge Singer’s critique of our moral categories. If he wants to abandon all talk of “supererogatory” actions, I want to abandon all talk of “obligatory” actions. It does not follow that I turn all duties of justice into those of generosity and love. Although the view I advocate for includes friendship as a significant value in deliberation, it does not exclude concern for justice, whether in my own personal community or in the global stage. On the contrary, I agree with Singer that the realization that we *can* do more entails that a good person *will* do more, much more than most people do now. Some effective altruist groups only require that its members give 10% of their income to effective charities (1% if a student), and, for some, “to do more” means to fight the systems which reinforce global inequality itself. Moreover, in my view, good governments would do much more than they currently do, whether giving more than 10% of their GDP in foreign aid or actively trying to change structures which cement the status of the global poor. However, this quantity does not necessarily need to be the point of marginal utility. My proposal, then, is that a different view of meta-ethics and moral theory, like an eudaimonist view, will have substantive consequences for this applied ethics problem; still, my view can present an attractive case for having global justice, alleviating suffering, and granting rightful reparations as an attractive aim.

My argument first showed that Singer’s argument relies on certain meta-ethical assumptions about the content of the concept implied by the moral ought and then suggested a potential alternative with substantial consequences. At this point, he might object that even if his view is not as innocent as he pretends, neither is mine; furthermore, I have offered no reason to

prefer my assumptions over his. And in fact, we should prefer his, he would claim, for my way of looking at the ethical universe is fundamentally selfish, for it positions my experience of self at the center. And a moral theory which begins with egoism is not a moral theory at all.

In response, I remark that I have given enough reasons as to why other-concern can still be part of the practical standpoint. But more importantly, I should also note that human beings are simply not impartial, even if we could behave *as if* we were. We are the sort of beings that have a sense of what it means to live well for us. When ethics begins here, with human existence as a human way of life, our theorems will look significantly different. To conclude my argument for this “soft” version, I invite my reader to picture the best version of ourselves possible, a human being whose life most of us, excluding Singer maybe, would choose as a goal: this person is good and altruistic, contributing as much aid as possible given her circumstances, but not literally self-less. She would not reach a point of destitution, not even of net positive utility, sacrificing it all to do all the good she could, as if she were a mere vessel of the forces of utility trying to be maximized; instead, she would leave room in her budget for her own personal flourishing, her loved ones, and justice in her community, and she cares for these for their own sake.

5.22 - 5.23
2021

Philosophizing (in) Uncertainties

 @ NYU
Annual Undergraduate
Philosophy Conference

What Counts as Reliving?

William Choi

Johns Hopkins University

May 2021

Abstract

Autonoesis as a temporally spanning sense of self has been deemed by Endel Tulving a fundamentally distinguishing feature of episodic memory, as opposed to semantic memory. This theory aligns with most people's intuitive understanding of how they relive episodic memories as a form of mental time travel but seems to neglect a small yet significant number of cases in which episodic memories are experienced without this auto-noetic sense of self. This paper will argue against Tulving's notion of autonoesis as a necessary feature of episodic memory by presenting counterfactuals of people who lack a sense of self while reliving episodes from their past. In order to empirically challenge Tulving's presuppositions, I will first explore the case studies of Patient R.B., who lacks a sense of ownership over his memories, as well as victims of past abuse who display dissociative memories. After addressing some potential objections that may be raised by the Tulving advocate, I will then develop a hypothetical case of the Reformed Criminal (RC) that can provide us with moral reasons to include self-less episodic memories as legitimate memories rather than non-memorial imaginations. The case of RC will ultimately demonstrate the importance of recognizing one's episodic memories as recollections of past

events that carry moral responsibility, even without the feeling of identification with one's past self.

What Counts as Reliving?

Introduction

Our memories can be divided into declarative memories and non-declarative memories, in which declarative memories involve explicit representational content, such as recalling my scuba-diving experience in Hawaii, while non-declarative memories, such as remembering how to ride a bike, do not. Declarative memories are further divided into episodic memories and semantic memories, in which episodic memories entail recollections of events that occurred in one's past, while semantic memories concern more general and abstracted facts, such as remembering that Paris is the capital of France. Endel Tulving, who coined the terms episodic and semantic memories, distinguished them by arguing that episodic memories involve the reliving of past events, which is in turn imbued with a "distinctive flavor of conscious phenomenology" called *autonoesis* (Tulving 3). In this paper, I intend to argue against Tulving's claim that *autonoesis* fundamentally distinguishes episodic memory from other types of memory. I will first examine the relevant terms of "episodic memory" and "autonoetic consciousness" before demonstrating that *autonoesis* is not a *necessary* feature of episodic memories and thus not a fundamentally distinguishing feature.

Defining Key Terms

I will be defining episodic memory as a memory of events experienced by the person who can retrieve it through mnemonic simulation. First, it is necessary to determine what "episodes" mean. Episodic memories typically involve the recollection of specific life events, and Cheng and Werning assume "that events are concrete (rather than abstract), [meaning] that

each event occupies a distinctive region in space-time, which no other therefrom independent event occupies” (Cheng & Werning 1353). Thus, episodic memories concern spatiotemporally unique events, in contrast to the general and abstracted information of semantic memory. In order to distinguish episodic memories from fabrications, these episodes must maintain some fidelity to an originally experienced event. In other words, it would not be possible to “remember” events that never actually occurred to me; they would instead be mere scene constructions or imaginations. However, it is important to note that episodic memories are psychological reconstructions of the event invoked through the reactivation of a “memory engram” or “memory trace,” which is a localized modulation in the neural wiring of the brain in response to an external event (Josselyn & Tonegawa). Due to this reconstructive nature of episodic memories, they may be prone to include confabulatory details, but as long as they remain faithful to the general “gist” of the event, they may count as episodic memories. Determining the threshold of accuracy for a memory to qualify as a genuine memory is an important but difficult task that we will save for another time as it is not critical for the purposes of this paper.

I will be interpreting auto-noetic consciousness as a temporally spanning sense of self that often accompanies the reliving of a recollected event. Tulving argued that episodic memory is imbued with auto-noetic consciousness because it involves mental time-travel back to the past event, which requires the conscious experience of the event *as experienced by myself*. For example, in order to relive my past experience of hiking in a nearby park, I must consciously experience the sight of the foliage, the coolness of the air or the rustling of the leaves as *I* had experienced while walking through the park. While I need not retain all of the sensory details or in their original richness, I must retain a sense of self within that moment of the scene in order to

recognize that I am reliving an event that I had experienced before. Tulving described auto-noesis as this particular kind of self-aware consciousness that is “necessary for the remembering of personally experienced events. When a person remembers such an event, he is aware of the event as a veridical part of his own experience” (Tulving 3). In order to ascertain the veracity of the past experience and its rightful belonging in one’s personal history, Tulving asserted that auto-noesis allows one to recognize that the past self is continuous with the current self. Therefore, episodic memory involves a sense of auto-noetic consciousness that maintains a sense of self, which in turn allows one to relive the event as experienced by the past self.

Responding to Tulving

I will now argue that episodic memories do not *necessarily* require auto-noesis by deconstructing Tulving’s presupposition of the sense of self when reliving memories. Before we begin, it is important to establish our precise point of dispute. While Tulving and I both accept that episodic memory involves some form of reexperience of the event, I diverge from his claim that an auto-noetic sense of self is necessary for this reliving. The Tulving advocate must demonstrate that reexperiencing episodic memories does in fact require a sense of self in order to prove that episodic memories fundamentally involve auto-noesis. Through a series of counterfactuals, I will attempt to demonstrate that while auto-noesis within episodic memories is a common and intuitive experience for most people, it is not *necessary* for reliving a past event and therefore cannot be a fundamentally distinguishing feature of episodic memory.

Tulving’s argument suggests that consciously experiencing a memory already presupposes that you have an auto-noetic sense of self connecting your current self to your past self. However, we can look to the counterfactual case of Patient R.B., a patient who suffered from a bilateral lesion in his hippocampus, to challenge this presupposition. While R.B. was able

to intentionally recall and relive through specific and spatiotemporally unique events in his past, he did not feel as if those experiences belonged to him. He described a memory in which he “could clearly recall a scene of me at the beach in new London... [but it was] as if I was looking at a photo of someone else’s vacation” (Klein and Nichols 686). R.B. demonstrated a capacity for episodic memory since he could reconstruct and reexperience scenes that are presumably accurate to events he actually experienced in the past, and his lack of a sense of self or ownership of the memories did not preclude his ability to relive that episode. Therefore, even if a person loses a sense of self-continuity with their past self, they may still relive these events, thus retaining episodic memories without auto-noesis.

Here, the Tulving advocate may argue that R.B actually does have a sense of self by virtue of reexperiencing these memories of his past, and his lack of “ownership” over his memories does not discount the sense of self that must necessarily accompany his conscious reliving of an event. However, it is necessary to distinguish the *self* that currently relives a past event from the auto-noetic *sense of self* that feels connected to that past self and believes the episode is a veridical part of one’s personal history. For example, most people can visualize a scene from their memories from a particular spatiotemporal perspective, yet the mere conscious visualization of the episode itself does not necessarily entail that the rememberer has a sense of self *within* that memory. I may be able to visualize the specific spatiotemporal event of myself lying in bed, drowsily reaching for my cough medicine, yet lack any sense of self-continuity with that person lying in the bed (perhaps because I believe I could never get sick because of my youthful vigor.) I may still recall the episode by visualizing a first-person perspective of a hand weakly grasping at the medicine bottle, or the sun filtering through the shuttered window, but

lack a sense of self while reliving this memory because I do not feel as if *I* had actually experienced the event.

Considering Objections

The Tulving advocate may be skeptical about the possibility of detachment from my sense of self within a memory, and argue that surely I must only be verbally proclaiming that I do not recognize myself while experiencing the memory all along as myself lying sick in bed. However, not only does R.B. present the realistic possibility of this counterfactual case, but these cases in general are not as implausible as they might first appear. When people experience events that are humiliating or traumatic, they may employ psychological defense mechanisms to personally detach themselves from the memory while recalling and reliving the event. One psychiatric study on victims of childhood abuse demonstrated that some participants reexperienced fragmented flashbacks of the past event through visual, auditory or kinesthetic reliving, but “when [they] have the memory, it is still not [theirs] or it belongs to somebody else, to an alter” (van der Hart 62). Instances of such dissociative detachment from traumatic memories demonstrate the possibility of reliving an episode without the accompanying sense of self within that memory. Thus, whether or not the recollected episodes *cause belief* in the remembering person as veridical parts of their own past does not change their categorization as episodic memories, as long as they originated from real events.

The Tulving advocate may persist and object that if one does not have a sense of self in these episodic memories, they are not memories at all but rather some non-memorial representation of an event. After all, if you feel like an imposter peering into someone else’s memories, you are no longer remembering but rather experiencing some intrusively inserted

scenes. This stance may appear as a convenient recategorization, but it will make the Tulving advocate accept some ethically problematic consequences of this objection. The Tulving advocate may quickly interject here that memory as a scientific concept should not be beholden to a moral debate. And while this argument belongs to a lengthier debate that I will not explore in this paper, I assert that what we choose to legitimize as memories or reject as non-memories do not exist in a scientific vacuum but are instead entwined with moral consequences and social duties that are worth sincere consideration. Let us imagine a Reformed Criminal (RC) who, after several years of education and spiritual healing, has positively transformed his outlook on life such that he no longer appreciates what his past life as a criminal was like. However, he occasionally experiences flashbacks to his previous acts of crime in vivid visual detail, such as the fear on his victims faces or shards of broken window glass, but denies any sense of self-continuity within that memory. It is not just that RC no longer endorses his past actions; RC expresses confused horror at the actions of his apparent doppelgänger and remains disassociated from his past self.

If the Tulving advocate agrees with RC that these could not be episodic memories due to RC's lack of connectedness to his past self, they withhold RC's ability to take moral responsibility for his past actions. Recharacterizing RC's episodic memories as non-memorial imagery contravenes the moral impetus for RC to hold himself accountable as the individual who committed those crimes, which denies an appropriate attitude of personal atonement towards the victims who suffered from his actions. By treating these visualizations as unrelated non-memorial imagery, RC cannot reflect upon his past mistakes and may be inclined to act irresponsibly in the future by failing to assume moral ownership over his personal history. It is critical to recognize that RC's recollections are still constructed from actual events experienced

by him and thus carry moral baggage even when devoid of RC's continuity of a *sense* of self.

This example demonstrates why the lack of auto-noesis need not relegate one's memories as non-memorial representations, but rather as episodic memories that lack a sense of self but are nonetheless reconstructed from actual events that occurred to one's self. Thus, by demonstrating the possibility of episodic memories without a sense of self through a series of counterfactuals, I have demonstrated that auto-noesis is not necessary for reliving episodic memories.

Conclusion

At first glance, auto-noesis as a temporally spanning sense of self seems integral for reliving episodic memories. Indeed, it is common for most people to mentally time travel through their memories with a continuous sense of self, and Tulving presents an intuitive advocacy for auto-noesis as a fundamental and necessary feature of episodic memories. However, this claim is based on mainstream introspection of episodic memories and neglects a small yet significant number of cases in which people lack this sense of self when reliving their past. By presenting counterfactual cases of patient R.B. as well as victims of past trauma, I have demonstrated the realistic possibility of reliving through spatiotemporally unique events that occurred to oneself without possessing a sense of self. To deny that these are episodic memories due to their lack of auto-noesis not only dismisses the testimonies of these individuals, but also harms their ability to understand and accept these memories as veridical episodes of their personal history. Even without a sense of self, these episodic memories ultimately ought to be accepted as genuine memories attributable to real experiences that occurred to their self in order to help people come to terms with past traumatic events or assume moral responsibility for past wrongdoings.

Works Cited

Cheng, Sen, and Markus Werning. "What Is Episodic Memory If It Is a Natural Kind?"

Synthese, vol. 193, no. 5, 2015, pp. 1345–1385., doi:10.1007/s11229-014-0628-6.

Josselyn, Sheena A., and Susumu Tonegawa. "Memory Engrams: Recalling the Past and

Imagining the Future." *Science*, vol. 367, no. 6473, 2020, doi:10.1126/science.aaw4325.

Klein, S. B., and S. Nichols. "Memory and the Sense of Personal Identity." *Mind*, vol. 121,

no. 483, 2012, pp. 677–702., doi:10.1093/mind/fzs080.

Tulving, Endel. "Memory and Consciousness." *Canadian Psychology/Psychologie*

Canadienne, vol. 26, no. 1, 1985, pp. 1–12., doi:10.1037/h0080017.

Van der Hart, Onno, et al. "Memory Fragmentation in Dissociative Identity Disorder."

Journal of Trauma & Dissociation, vol. 6, no. 1, 2005, pp. 55–70.,

doi:10.1300/j229v06n01_04.

5.22 - 5.23
2021

Philosophizing (in) Uncertainties

@ NYU
Annual Undergraduate
Philosophy Conference

Arguments Against Arguments Against Affirmative Action (AAAAAA)

Long Do

Brown University

May 2021

Abstract

This essay addresses several of the arguments laid out by Louis P. Pojman in his influential essay, *The Case Against Affirmative Action*. His argument is structured as a series of responses to the most popular arguments in favor of affirmative action. This paper responds to those responses. His claims are largely based around a distinction between weak affirmative action, which he supports, and strong affirmative action, which he rejects. This difference is characterized roughly as the difference between equalizing opportunities and equalizing results. This paper begins with a response against this distinction, arguing that it is a rather meaningless division, as results for certain opportunities are opportunities for other results, so one cannot equalize opportunities without also equalizing results. For example, admission into college is a result in one sense but it also serves as an opportunity in another sense, such as for graduate school or certain jobs. It would be impossible to give all students similar qualifications in applying for graduate school without having equalized the result of who received admission into undergraduate schools. This essay also addresses Pojman's multiple rebuttals against the argument for diversity which states that admitting students on the basis of their potential ability to expose other students to other cultures is treating them as a means to an end and that

competence should be weighed much more heavily than their contribution to diversity. This paper argues that this treatment of students is not of a different kind than a criteria that takes into account the effect of a student's admission on anyone or anything that is not themselves. It also contends that the contribution of competence to the fitness of a candidate varies by position. Finally, this paper responds to Pojman's rebuttals against the argument that no one deserves their merit by pointing out some insufficiencies in his counterexamples.

Arguments Against Arguments Against Affirmative Action (AAAAAA)

Louis P. Pojman, in his essay *The Case Against Affirmative Action*, divides affirmative action into two tiers: weak affirmative action and strong affirmative action. The weak version is an argument for the equality of opportunity, which would include policies such as providing scholarships for the underprivileged or in other ways increasing the opportunity for “disadvantaged people to attain social goods and offices” (Pojman 2). Conversely, strong affirmative action seeks to tilt the scales for results instead; this includes, for example, policies that preferentially treat applicants on the basis of race. Pojman argues that weak affirmative action is morally required and that strong affirmative action is morally reprehensible. He examines six arguments in favor of strong affirmative action and concludes that they all fail. Moreover, he also argues in favor of three arguments against strong affirmative action. This essay will serve two main functions. The first is to criticize Pojman’s division between weak and strong affirmative action. The second is to respond to his rebuttals to two arguments in favor of strong affirmative action, specifically his remarks on the diversity argument and the “no one deserves their talent” argument, which I’ll shorten to the ‘undeserved merit’ argument. I’ll begin by taking a deeper look at Pojman’s view on the differences between weak and strong affirmative action and argue that they are not particularly meaningful. Then, I will give an overview of the argument for diversity, explain Pojman’s objections to those arguments, and rebut, accordingly. I will follow the same structure for the underserved merit argument.

I have already stated some differences that Pojman lays out between weak and strong affirmative action in the introduction. Further details include stipulations that weak affirmative action can take the form of policies such as scholarships, specifically for underprivileged classes, but that these should be based on socio-economic background and not race. He also characterizes strong affirmative action with examples such as black female students being accepted into college over white male students with superior qualifications and provides similar examples applied to the professional domain. I think the main difference Pojman stakes out between the two types of affirmative action is succinctly captured in his statement, “the goal of weak affirmative action is equal opportunity to compete, not equal results” (Pojman 2). All the other examples he gives seem to align under this maxim. We should not favor one applicant over another on the basis of race in regard to college admissions in order to tilt the scales when it comes to results. We should, however, ensure that students from all backgrounds have equal access to the tools that might make them strong applicants for college.

The fundamental flaw with differentiating strong affirmative action from weak affirmative action is that there is no clear way to define what counts as an ‘opportunity’ and what counts as a ‘result’. Even more strongly, the situations which Pojman sees as illegitimately equalizing results could be easily interpreted as an opportunity to be legitimately equalized so that the agent could compete for some higher result. For example, in the situation of college admissions, Pojman believes it is unfair for certain groups of people to be advantaged over others. He views the result here as being accepted into the university in question and the opportunity as some baseline of schooling and resources that make a student a strong applicant. If you shift the window through which you are looking at this hypothetical applicant’s life, perhaps their real goal is being a doctor. Equal opportunity for the result of being a doctor would

include admission into this college. If it is a prestigious school, an education at that institution would improve their likelihood of being accepted into and graduating from medical school. If other applicants who also want to be doctors are accepted into the school, then it seems like society is failing to guarantee the original student a fair opportunity to compete.

Of course, Pojman can respond by saying the type of opportunity he means is primary schooling. Access to all other opportunities after primary schooling can be regarded as a result that should not be equalized under his framework. However, it seems arbitrary to choose a moment in time when inequality in opportunity is suddenly permitted. Without a sufficient reason behind this cutoff point, the distinction between these two tiers of affirmative action seems meaningless. As an extension, the moral distaste of strong affirmative action and the moral acceptance of the weak version are misplaced.

After formulating this distinction between weak and strong affirmative action, Pojman spends the next part of his essay discussing certain arguments for strong affirmative action. Despite my rejection of this delineation, I will address his responses to the arguments for strong affirmative action on its merits. He refers to one of the arguments he analyzes as ‘the argument for diversity’. This argument states that there is value in encountering people from different backgrounds as they can expose other students to different values and ideas. Encountering a variety of perspectives can help someone scrutinize their own views and refine their opinion with the aid of a more complete context. A diverse student population will also share a diverse culture. There is often “aesthetic and moral value” to diverse customs (Pojman 6). Additionally, in the pluralistic world that we live in, it is important to know how to get along with people of different ethnicities and races. A racially diverse student body grants students the opportunity to practice how to interact with the diverse world they will graduate into.

(1) Pojman rejects these benefits on the grounds that they do not seem substantial enough to override a principle of treating people with equal respect. (2) Admitting students simply for the sake of educating other students about different cultures or exposing them to interactions with students of different races is treating minority students as a means to an end. (3) He also argues that competence far outweighs diversity. He would much prefer that the surgeon who operates on him is skilled than that he adds to the diversity quotient of the hospital staff. He makes a similar argument about how basketball fans would much prefer a high performing team than one whose players accurately reflect the demographic makeup of the country they live in.

I'll reply to each of these points in order. It does not seem that taking diversity into account interferes with the maxim that people should be treated with equal respect. For example, in the case of college admission, admissions officers are making decisions based on their preference of certain traits over others. This practice is not one that Pojman rejects in its entirety so I will assume that it still counts as treating people with equal respect. Adding diversity to the list of criteria that admissions officers are trying to maximize does not fundamentally change the type of task they are doing. If preferring an applicant who was a high school debater over one who was a mechanic's apprentice is treating people with equal respect, so is preferring an applicant whose culture is not as well represented in the student body over one whose culture is. Since Pojman has established a principle that selecting one applicant on the basis of a preferential trait is treating people with equal respect, that principle could be applied to the preferential trait of diversity and maintain its status as a practice that treats people with equal respect.

His second argument is that minority students would be treated as a means instead of an end. He does not explain why such a classification is so damaging to the argument for

affirmative action. To revisit the example of college admissions, students are often admitted not just on the basis of their academic achievements insofar as they benefit the student. Their academic achievements are valued partially because of what their potential success in their field of choice could mean for the school's reputation or endowment. Looking at potential students through such a lens is also treating them as a means to raise the reputation of the school. Examples of this type of potential student analysis extends past just calculations of possible future donations. For example, excellent athletes are often advantaged in college admissions because they will benefit the performance of the college's team. Other students might enjoy having a high performing team represent their school. A consideration of how a student's presence on campus benefits people other than themselves is at least partially treating them as a means for an end that is not themselves. Students in the current state of college admissions are often treated as a means. A rejection of affirmative action on these grounds would have to be followed by a rejection of all such criteria in college admissions.

His final argument is that the competence of students should heavily outweigh what they add to the overall diversity of the school. I have two responses to this point. First, I think that this argument does not refute affirmative action policies. It is often the case that people accepted through affirmative action are quite well qualified anyway. Even if they are not as well-credentialed as other applicants the competence trade-off seems small, though this claim is an empirical one. In either case, an acceptance of diversity as a possible criterion to choose one applicant, even if weighed lightly, over the other is an intrinsic acceptance of some level of equalizing of results.

Second, the definition of competence that Pojman uses is unclear. If competence is defined as the amalgamation of the necessary traits that would benefit a person in performing at a certain job, it seems obvious that competence is the only quality that one should be maximizing. In this case, I would argue that an applicant's potential contribution to diversity would be included in this overall competence metric for the reasons outlined in the positive arguments above. If competence is a more specific quality, having to do with intelligence and hard skills applied to whatever the occupation is, then the amount to which that quality contributes to overall potential job performance varies from position to position.

In the case of a surgeon perhaps it is relatively more important that they are competent. For other positions, however, like being a college student, competence of this kind seems to constitute a lower proportion of their overall quality as a candidate. Qualities such as friendliness, speaking skills, ambition might contribute significantly to one's candidacy. Since it is the case that competence contributes different amounts to the quality of a candidate, it is wrong to say that competence always heavily outweighs contribution to diversity. Pojman's attacks on diversity fail in several ways. Any acceptance of diversity as a possible criterion for accepting a candidate into a position is an example of taking race into account in the results stage and is an acceptance of strong affirmative action.

The final argument for strong affirmative action that Pojman looks at is what I'm calling the underserved merit argument. The argument which Pojman rebuts revolves around the idea that all the qualities which we currently possess were bestowed upon us at birth, as a part of the genetic lottery. While it is the case that a person who is currently a professional ballerina was probably incapable of performing pirouettes at birth, her potential to become an extremely good dancer was the result of some natural gift for dexterity in her limbs or work ethic. This argument

presupposes the idea that all skills or qualities we currently possess can be traced to some trait we had or were destined to have from birth. Since it was purely by chance that we ended up the way we currently exist, as it was the result of chance that we were given the traits we did at birth, we have not truly earned any of our achievements. Since we do not deserve our credentials, we don't have a legitimate claim to any position. Therefore, society can distribute positions in any way the society desires; the process of distribution can include strong affirmative action.

Pojman gives a formal formulation of the argument which he will use to dispute its validity. The argument is constructed as follows: (1) Society ought to distribute positions on the basis of whether or not an individual has a claim to that position, (2) In order to have a claim to something, one must have earned it, (3) Individuals have not earned their natural intellect, work ethic, industriousness, or any other quality that leads to superior qualifications (4) People who have not earned what produces something, have not earned the result. (5) People with superior qualifications do not deserve positions over less qualified individuals (Pojman 10).

Pojman argues against the undeserved merit argument by disputing the 4th premise. He believes it can be the case that people deserve the products of tools which they did not deserve. He gives an example of two people who are given a gift of \$100. Person A chooses to save the money by burying it in a box underground. Person B chooses to invest the money into stocks. After 10 years, person A digs up his box and still has \$100, while person B sells his stocks, which have increased in value over time, and has \$200. Pojman argues that it would be unreasonable for person A to be entitled to the gains from person B's initial investment into the stock market. Despite the fact that person B did not deserve the initial \$100, as it was a gift, he can still deserve the extra \$100 that the initial \$100 produced. (Pojman 11)

The issue is that the scenario does not accurately recreate the stipulations of the premise. Pojman argues that person B should be rewarded for making the better financial decision to invest the gift. He is implicitly creating two separate categories for analysis. The first is the gift, which represents undeserved merit and the second is the decision to invest, which constitutes deserved merit. The words 'gift' and 'decision' are meant to appeal to the reader's rational/intuitive basis for desert. However, it seems wrong to create such a contrast. Person B's decision to invest the money could be the result of some savvy business sense or some propensity for long-term thinking. A proponent of the undeserved merit argument would claim that such traits can be traced to some genetic gift that person B did not deserve. In this case, the gift of money and the decision to invest are both representative of undeserved merit. A reasonable reframing of the scenario with these observations in mind would be: person A is undeservedly given a gift of \$100 and person B is undeservedly given a gift of \$200. The \$200 is meant to represent the financial worth of the original \$100 added to whatever undeserved skills person B utilized in deciding to invest the money. In this case, it seems much less clear that Person B is deserving of the extra \$100. Since the example Pojman uses to refute the 4th premise of the undeserved merit argument does not accurately reflect the situation, it fails to refute the point. As a result, the undeserved merit argument is still a valid argument in favor of strong affirmative action.

Ultimately Pojman fails to create a meaningful distinction between weak and strong affirmative action. His refutations against the argument for diversity and the undeserved merit argument have serious flaws, such as failing to address how diversity of backgrounds is fundamentally different than any other acceptance criteria and unsuccessfully refuting the 4th premise of the undeserved merit argument. Pojman agrees that the goal is equality, and the

question, of course, is how to get there. Affirmative action is a complicated policy, but one that should not be removed from the arsenal of tools that move us closer to our ideal society based on the grounds that Pojman lays out.

References

Pojman, Louis P. "The Case Against Affirmative Action." *Oxford University, International Journal of Applied Philosophy*, 1998.

Empedocles' Notion of Universal Kinship as an Important Consideration In Ecological Thought

Josh Doland

University of California, Los Angeles

May 2021

Abstract

When thinking about the proper way in which humans ought to relate to plants and non-human animals, it seems that we have a default preference for answers which are grounded in a form of human exceptionalism. This is to say that humans tend to view themselves as being in an antagonistic relationship with nature, where we, being the superior and civilized force, are justified in dominating and exploiting our non-human competitors. This anthropocentric perspective drives us to alienate plants and non-human animals, making them others from which we can contrast ourselves.

We can think of many of our modern practices as embodying this perspective: we create factory farms that cause an astronomical amount of animal suffering, we deforest at an unsustainable rate, and even in our daily lives we alienate our pets by projecting our human sensibilities and expectations upon them.

Within the following paper I take up the work of Empedocles, a pre-Socratic, in an attempt to show how his notion of universal kinship could positively impact our relationships with other members of the biosphere. In order to do this, I first offer a brief characterization of Empedocles' cosmology and anthropogony in order to ground his notion of universal kinship. Next I show how adopting this ecological mindset would affect our practices of deforestation and

factory farming. Finally I offer a reinterpretation of a modern ecological thinker, Donna Haraway, in reading her arguments for the restructuring of our relationships with companion species as following in the tradition of Empedocles.

Empedocles' Notion of Universal Kinship as an Important Consideration In Ecological Thought

When thinking about the proper way in which humans ought to relate to plants and non-human animals, it seems that we have a default preference for answers which are grounded in a form of human exceptionalism. This is just to say that humans tend to view themselves as being in an antagonistic relationship with nature, where we, being the superior and civilized force, are justified in dominating and exploiting our non-human competitors. This anthropocentric perspective drives us to alienate plants and animals, making them others from which we can contrast ourselves.

We can think of many of our modern practices as embodying this perspective: we create factory farms that cause an astronomical amount of animal suffering, we deforest at an unsustainable rate, and even in our daily lives we alienate our pets by projecting our human sensibilities and expectations upon them.

Our current insistence on separating ourselves from nature in order to dominate it is characteristic of the sort of strife-dominant world that one pre-Socratic, Empedocles, warned against embracing. In forcing our destructive wills on plants and non-human animals, Empedocles suggests that we have forgotten the intimate history we have with them, and so have strayed from the ethical path of unity and love.

In order to argue for this shift in ecological thinking, in the first section I will offer a basic presentation of Empedocles' cosmology and anthropogony, and in doing so, contextualize our current era of aloof ecological thought within Empedocles' wider story of human emergence. In the second section I will show how Empedocles' understanding of the world ties in with his central notion of universal kinship and metempsychosis. In the third section I will

briefly examine our current practices of factory farming and deforestation and suggest how we might reconfigure these practices in light of our new understanding of universal kinship. In the final section, I will offer an Empedoclean interpretation of the work of Donna Haraway, in order to show how some modern ecological thinkers are already taking up the idea of universal kinship in their work, in this case, by re-examining lines drawn between human and animal.

SECTION I: EMPEDOCLES' ANTHROPOGONY AND COSMOLOGY: HUMANS AS THE FOURTH GENERATION

In this section it is my goal to put forth Empedocles' basic cosmology and anthropogony. In doing so I hope to make clear how Empedocles would characterize our current relationship to plants and non-human animals.

Within his theory Empedocles thinks of all beings as being composed of the same four basic elements: earth, air, fire, and water which he refers to as Zeus, Hera, Aideon and Nestis.¹ These elements are thought to mix and interchange depending on the forces that act upon them but are never actually generated or destroyed.

Along with these four elements, Empedocles introduces two antagonistic forces: love and strife. These forces act upon the elements and account for their apparent change as they whirl between bodies. Love is thought of as a sort of unifying force which brings like things together while strife is a separating force which breaks things apart. Unlike another pre-Socratic, Heraclitus, who viewed strife as a sort of necessary good, Empedocles characterizes strife as an accursed evil which is not to be embraced.² With a basic understanding of Empedocles' elements and the forces of love and strife we can now turn to his formal cosmology and anthropogony.

Empedocles thought that the process by which humans and the world emerged was a circular one, brought about by the natural waning and waxing of the forces of love and strife.

At the beginning of our world, love strongly dominated and so all the elements were unified together into a singular sphere of being. Gradually love began to wane and strife became increasingly dominant, leading to the sifting off of some of these elements into separate bodies. Empedocles distinguishes four stages of this sifting off as follows:

¹ Barnes. 2001. pg. 130.

² Ibid. pg.133, 135.

- 1) In stage one elements initially began sifting off of the sphere. There were no whole beings, only pieces of beings, with “naked arms [wandering] devoid of shoulders” and eye straying alone “in search of foreheads”.³
- 2) In stage two more of the sphere sifts off, allowing for the creation of monstrous forms with “man-faced oxen and ox-headed men”.⁴ Love still dominated during this stage, but there were enough sifted off elements for separate whole beings to coalesce outside of the sphere.
- 3) In stage three, strife begins to dominate love and we see the emergence of recognizable humans, plants and non-human animals.
- 4) During the fourth stage, strife has become far more powerful than love. The humans, plants and non-human animals which emerged in the previous stage begin to reproduce, and forget their initial state of unity now viewing each other as separate beings.

Empedocles thinks that modern humans fall into this fourth stage, in which we, “trusting in mad strife...wander in the darkness on the meadows of ruin.”⁵ Essentially we as humans have

³ Ibid. pg. 142.

⁴ Ibid.

⁵ Ibid. pg. 121.

forgotten our initial unity with all other beings, and have begun to regard ourselves as exceptional in comparison to other non-human animals and plants.

Though Empedocles only notes four stages, it is implied that strife would become increasingly dominant past the fourth stage, leading to the dissolution of human bodies until we reach a sort of perfectly separated world, at which point love would start to become more dominant and the world would collapse back into a sphere of unity.⁶

Empedocles' diagnosis of humans as falling into the fourth stage seems strikingly accurate. We do view ourselves as separate and superior to other non-human animals and plants, and we use this superiority in order to alienate and enact violence on these other beings.

While within this section I have merely presented some of the basic moving pieces in Empedocles' picture, in the next section I will examine how his background understanding of the way in which humans emerged ties in with his greater idea of universal kinship.

SECTION II: EMPEDOCLES ON UNIVERSAL KINSHIP

As shown previously, Empedocles thinks that all beings are composed of the same fundamental elements which were initially unified in one central sphere of being. As strife separated these elements out they became modern humans, but also “plants and fish of the

⁶ This section deals with Empedocles' notion of double birth and double death, in which the world is thought to continually cycle through these periods of absolute unity and absolute dissolution. We can think of an analogue in the modern astrophysical theory of the big bang/big crunch, in which the universe is thought to cycle through a similar series of expansions and contractions. For the sake of relevance and space constraints, this idea will have to be explored further elsewhere. Barnes. 2001. pg. 120, 121.

watery halls, and beasts of the mountain lairs and flying gulls.”⁷ The elements which compose the plants, fish, and humans are all the same, they simply differ in terms of proportion and location.

This underlying unity of material leads Empedocles’ to his theory of metempsychosis, a theory which states that the soul in some way transcends the body and can travel, taking many different forms. Empedocles argues that he was once “an immortal god”, now mortal, and that he had in past lives been “a boy and a girl and a bush and a bird and a silent fish in the sea.”⁸ Since all beings are composed of the same elements, and since they had once been unified in a singular sphere, the soul is able to travel across bodies in different lives.

Along with the idea of metempsychosis comes a rather strict ontological hierarchy of beings, in which Empedocles thinks that laurel trees are better than other plants, lions are better than other non-human animals, and gods are better than mortal men.⁹ Depending on how you act during your life, your soul would be transferred either up or down this ontological hierarchy.

Since Empedocles thought that the soul could be transferred across bodies, and could inhabit not just humans but also plants and non-human animals, it seemed unethical to him to eat animals or to harm plants. When one did so, Empedocles thought we were unknowingly murdering one of our own, as if “a soon [took] his father, [or] children their mother [and] they

⁷ *ibid.*

⁸ *ibid.* pg. 157.

⁹ *ibid.* pg. 114, 157.

bereaved them of life and [ate the mother and father's] dear flesh."¹⁰ To consume animal flesh was then a sort of cannibalism, as a human could be inhabiting the animal body in question. Similarly, harming a tree was to potentially harm a soul within a tree, and so Empedocles advised to "keep altogether the leaves of the laurel" and to "keep your hands from beans".¹¹

It is this understanding of plants and animals as related to humans that informs Empedocles' notion of universal kinship. All things are related in that they are constituted of the same materials, and so in some sense all things are kin to one another. In the same way we condemn cannibalism, Empedocles condemns the abuse of non-humans, as we all belong to a greater, unified being, even if we are currently in separate forms.

While this notion of universal kinship seems attractive in providing a strong framework for animal and plant rights, one could reasonably be concerned that it builds within it certain anthropocentric notions. After all, it may seem like Empedocles is only concerned about plants and animals insofar that a human soul inhabits their bodies.

While this is a reasonable interpretation of Empedocles' work, I would argue that it overemphasizes the importance of the human soul in particular. As mentioned above, the human is not at the top of the ontological hierarchy, but merely a step on a long ladder towards divinity. I would suggest that when Empedocles raises concerns over cannibalizing souls embedded in animal or plant forms, he isn't concerned that humans are eating humans, but rather that a soul is eating another soul. Since all souls contain within them the same elements, and these elements

¹⁰ *Ibid.* pg.158

¹¹ *Ibid.* pg. 159.

all initially come from a unified being, we can reasonably interpret the concern being not over the humanness of animals and plants, but rather of the inherent value of beings constructed of the same elements.¹²

One may still be unsatisfied with Empedocles' account, as it does preserve within it an ontological hierarchy which puts humans above other animals and plants. This is also a reasonable concern, though one which needn't interfere with his notion of universal kinship. Each plant, animal, human and god are made of the same elements which were united in the same sphere of being. Even if Empedocles ranks lions as greater than other animals, and perhaps animals above plants, each being still has its own inherent value in being constituted of the same, shared, elemental components. It's also worth saying that while humans are placed above both plants and animals, Gods are placed above humans. This is just to say that while a degree of human exceptionalism may be read into Empedocles' ontological hierarchy, we need not interpret it in this way.

It seems then that Empedocles, in acknowledging universal kinship, requires us not to eat or harm animals or certain plants. These prohibitions seem very strict and difficult to follow; it will be my task in the following section to take up our current treatment of plants and animals and suggest how Empedoclean lines of thought may lead us to alter these relationships.

SECTION III: APPLYING EMPEDOCLEAN ETHICS

Within this section I will examine how Empedoclean lines of thinking could affect our current practices of deforestation and factory farming. In doing so I hope to both clarify what

¹² *ibid.*

sort of ethical obligations flow from claims of universal kinship while also offering a realistic path forward. To be clear, it is not my goal to offer a comprehensive solution for the issues being addressed, but rather to gesture towards how taking the idea of universal kinship seriously may lead us to change some of our current practices.

For our purposes we can understand deforestation to refer to the permanent, mass removal of trees either for the sake of harvesting tree-related products like palm oil or for making room for agriculture or the grazing of animals. If we took the Empedoclean notion of universal kinship seriously, we would need to consider the moral status of the trees which we are harvesting and the animals affected by deforestation efforts.

With this in mind, we might take various approaches to reduce the number of trees we need to harvest. This could be accomplished through a number of different strategies, including finding replacements for some non-essential products like palm oil or by making sure our harvesting techniques are efficient. By shifting our perspective as humans from an exceptional one, to one which is situated within the wider biosphere, we have the groundwork to care about the wellbeing of plants. Even if we can't eliminate the harm we do to plants, we can be mindful of the impact we are having, and the ways in which our actions affect the wider ecosystem. This seems preferable to the kind of indiscriminate practices we currently engage in.

Similar considerations bear on our practice of factory farming, a method of harvesting meat which crowds animals together in poor conditions, only to slaughter them, typically in a brutally efficient manner. For Empedocles, it seemed clear that any eating of meat was a form of cannibalism and so should be avoided. Even if we aren't ready to commit to strict vegetarianism,

it seems clear that if animals are kin to us, we should not be creating facilities which cause such great animal suffering. By thinking through the lens of universal kinship, we have the framework to advocate for vegetarianism, or at least the harvesting of meat which either is lab-grown in a way which doesn't harm animals, or is harvested in a way which treats animals as beings whose moral status matters.

While the examples taken in this section are on a larger scale, the impact of taking seriously the notion of universal kinship also is relevant to our daily lives. Within the next section I will offer a brief introduction to the work of Donna Haraway, seeking to examine the ways in which she applies a sort of kinship concept to our relationships with our pets.

SECTION IV: THE COMPANION SPECIES MANIFESTO AS A NEO-EMPEDOCLEAN WORK

In this section my task will be to briefly introduce some of the work of Donna Haraway, a contemporary ecological thinker who takes seriously a notion similar to Empedocles' universal kinship in addressing the sorts of relationships we have with certain plants and animals like dogs.

In her work, *The Companion Species Manifesto*, Haraway focuses unsurprisingly on companion species, which are certain plants and animals that “make the leap to bio-sociality” and so enter into direct relationships with humans.¹³ Haraway suggests that our relationships with companion species are special in that they are formed by a sort of ontological dance, in which both species co-constitute each other. In arguing this, Haraway makes permeable the line between human and animal, or plant, and so engages in a project which is similar to Empedocles universal kinship. According to Haraway, a dog and a human can only be denoted as such when

¹³ Donna Haraway. 2003.pg. 106.

they stand in a relationship with each other. Neither species pre-exists their relationship, reminiscent of how individual beings didn't pre-exist the initial sphere of unified being.¹⁴

Once Haraway has established her subject matter and described the ways in which our relationships to companion species are distinctive, she makes a number of prescriptive claims about the pet-owner relationship, more specifically dealing with owning a dog. While Haraway ultimately thinks the pet-owner relationship is not something we need to do away with, she does think that certain forms of projection onto dogs fail to treat them as significant-others, but rather alienates them, and so sets them up to be euthanized.

Haraway in particular has in mind the practice of treating one's dog as a "furry child" with expectations that the dog should in turn offer unconditional love to its owner.¹⁵ When we engage in this practice, we fail to recognize our dog's distinctive dog nature, and instead project upon it our human sensibilities. For Haraway, it is important that we stand in a relationship of mutual respect, or mutual kinship, and as part of that we need to recognize the capacities and temperaments of those we engage with.

In this section we can read Haraway as arguing that the insistence on anthropomorphizing one's animals is to set them up to bite children and subsequently be killed.¹⁶ She thinks we need to shift our perspective from one of human exceptionalism and projection, to one of universal kinship and significant otherness, in which we recognize and respect the status of other members

¹⁴ *Ibid.* 103.

¹⁵ *Ibid.* pg.128.

¹⁶ *Ibid.*

of our biosphere without projecting upon them our distinctly human set of expectations and demands.

In articulating the companion-species relationship, I suggest we can read Haraway as continuing a distinctly Empedoclean project of breaking down seemingly strict distinctions between human, animal and plant, in order to foster relationships of mutual recognition and regard.

In closing, it would seem then that Empedocles' work and notion of universal kinship have the potential to reconfigure the ways in which we view ourselves in relationship to other members of the biosphere. It will be a long process to move from our current aloofness to this more situated perspective, but one which I would suggest will be fruitful not only for us, but for those whom we cohabit the earth with.

Works Cited

Barnes, Jonathan. 2001. *Early Greek Philosophy*. London: Penguin.

Haraway, Donna Jeanne. 2003. *The Companion Species Manifesto: Dogs, People, and Significant Otherness*. Chicago, Ill. : Bristol :Prickly Paradigm ; University Presses Marketing.

5.22 - 5.23
2021

Philosophizing (in) Uncertainties

@ NYU
Annual Undergraduate
Philosophy Conference

Can a Kantian Count? *Kant on Aggregation*

Andrew Garber

Harvard University

May 2021

Abstract

We often face tradeoffs between helping many people and helping only a few. Intuitively, if we can help the many without causing the few any more harm, then we should help the many. One supposed virtue of consequentialism is that it explains and justifies straightforwardly why we ought to help the many in such situations. Non-consequentialist ethical theories seem less likely to come to this conclusion, since they often reject that goodness can be aggregated. I aim to show that Kantian ethics can account for our duty to the many over the few in various circumstances. First, I argue that the Formula of Universal Law and the Formula of Humanity both yield a duty to help someone in these tradeoff situations. Second, I show that only the Formula of Universal Law can establish whom we ought to help. Third, I show that if we understand the Categorical Imperative as such, then saving the many would be virtuous, and saving the few would be neither virtuous nor vicious. Finally, I conclude that these two formulations of the Categorical Imperative are inequivalent.

Can a Kantian Count? *Kant on Aggregation*

According to Kant, there are two main ways to test whether agents act on good principles: The Formula of Humanity and the Formula of Universal Law. The Formula of Humanity stems from Kant's value theory. He claims that if there is a moral law, there must be some unconditional good. Everything we usually take to be good derives its value from this unconditional good. For Kant, the unconditional good is humanity, which is agents' ability to set and pursue ends. Ethics involves respecting this humanity by valuing our and others' ends. The Formula of Universal Law, on the other hand, focuses more on the form agents' maxims must take. For Kant, that form is universality: rational agents must be able to will their maxims so that they can be universal laws. If a world in which the maxim is inconceivable as a universal law or undermines the agent's will in such a world, then such a maxim would fail the test.

Here is the scenario. You are hiking on a trail atop a hill. In the distance, you see an empty trolley hurtling down its track. Ahead of the trolley, there are two sets of people. One group comprises five people tied to the track; the other contains just one person. The groups are far apart, so you only have time to save one group. As a Kantian, you ask yourself: "What should I do?"

You might immediately ask, "must I do anything?" In other words, do you have a duty to save someone? Both formulations of the Categorical Imperative affirm. The Formula of Humanity requires that we take others' ends to be at least somewhat valuable (*Groundwork* 4:430). But if you ignore the people on the tracks, you take their most important end—their existence as persons—as worthless. Thus, to respect humanity as an end, you must help them to stay alive. The Formula of Universal Law provides a more roundabout argument. Agents necessarily act in ways sometimes consistent with self-love. But "since our self-love cannot be separated from our need

to be loved (helped in case of need) by others as well, we therefore make ourselves an end for others” (*Metaphysics of Morals* 6:393). Thus, when we need help, we will to be helped by others. Hence if your maxim is not to help others when you are the only one who can help, then the universalization of this maxim generates a contradiction in the will (*Groundwork* 4:423). Thus, you have a duty to help those who could not get help otherwise. You therefore have a duty to help the people on the tracks.

According to both formulations, you must save someone on the tracks. Both of my arguments are drawn from Kant’s arguments for the duty of beneficence. One important note here is that the duty of beneficence is usually *imperfect* or *broad* – that is, it is a duty to have a certain end (the happiness of others) and directed at rational agents impersonally and impartially, not at some agent in particular. As a result, agents have leeway in how they implement beneficent behavior, both in how much to contribute, and how they contribute. But the duty of beneficence *can* also be *perfect* and *strict* – that is, the duty can be owed to particular people and can involve the execution of a specific action, leaving no leeway. The argument from the previous paragraph shows that in the trolley case, the duty to save someone on the tracks is perfect and strict—an example of the latter type of duty of beneficence.

We have now shown that you can have a perfect, strict duty to save someone. Any plausible moral theory would answer this way. The more interesting question is whether you have a duty to save the five over the one. I set aside arguments favoring the one over the five. Most Kantian arguments for saving the one would also apply to any of the five individually. Hence, at best such an argument would make it permissible to save either group. Thus, I will analyze whether the Formula of Humanity and Formula of Universal Law establishes a duty to save the five. I begin with the Formula of Humanity.

One tempting thought with the Formula of Humanity is simply to add up the reasons and compare them. Five people have five times as many ends as one person; they have five times the humanity. Thus, if you want to respect humanity as much as possible, you should help the five. This answer misunderstands what it means to value humanity. In the *Groundwork*, Kant draws the distinction between price and dignity. Things with a price, such as objects you can buy at the store or “wit, lively imagination, and humor” are good because they are good for us by their effect (*Groundwork* 4:435). Because things with a price are only good through their effects, they are not good in themselves. Therefore, they can be replaced without any moral qualms. Moreover, having more of these priced things produces more goodness, since the resulting effect increases. For example, five books bring me more joy than one book does.

On the other hand, things with dignity are good in themselves, not for the sake of something else, like the effect produced from them. Thus, they are irreplaceable. Also, their value does not add up in the same way as the value of things with a price do. Kant thinks that we must value humanity unconditionally because rational agency is intrinsically good through its willing, not through its effects. Thus, humanity involves dignity. Hence its value should not be aggregated. This view contrasts with the act utilitarian view of the worth of agents. For the act utilitarian, agents have no intrinsic worth; rather, it is good to preserve agents because they bring about good effects, such as pleasure or the fulfillment of desires. In other words, the act utilitarian treats agents as things with a price; if the agents have dignity instead, one cannot aggregate agents’ worth in the way the act utilitarian does.

We can make another argument from the Formula of Humanity for saving the five without aggregating the value of agents. Suppose that you save the one. One of five might complain: “You’re treating four of us as though we weren’t here! If it had been one against one, it would be

reasonable for you to use your discretion to choose between us, as our reasons conflict. But now there are five of us. You are acting as though the other four do not matter.” Her complaint does not assume that we ought to “sum up” the value of rational agents and decide accordingly. It merely assumes the more modest claim that it is wrong to ignore people’s humanity. Thus, one would do wrong by saving the one rather than the five.

Reasoning that way, suppose you then choose to save the five. Now the one seems justified in retorting: “Now you’re ignoring me! Had I not been tied to this track, you still would have saved the five. Nothing about your action now acknowledges my presence.” His complaint is that you are not respecting his humanity, just as you would fail to respect the humanity of the five when saving the one.

We are now at an impasse. Whichever side you choose, you fail to respect someone’s humanity. There are some creative attempts to weasel out of this situation. One is to flip a coin: heads, save the five; tails, save the one (Taurek 303). Now you give everyone an equal chance of living, thereby recognizing their equal value. This still does not solve the problem. The people’s value is not equal; it is incommensurable. When you take them to be equal, you mean that a world in which one group survives rather than the other has the same value as a world in which the other group survives. You take these worlds to be interchangeable. If these worlds are interchangeable, however, you seem to treat those people as replaceable, as they are the only difference between the worlds. But as noted before, persons are irreplaceable.

Let us now examine arguments from the Formula of Universal Law. To test whether you can save the five, you make it your maxim to neglect helping some people, even though you are the only one who can help them, in order to help a greater number of others. You can well conceive

of a world in which this maxim is universalized; it is the same as the world of beneficence, just one in which people must make tradeoffs. If we posit that the world of the universalized maxim has tradeoffs, then there is no contradiction. If you cannot get help because people are already helping others, this is no contradiction, because such a tradeoff is inevitable. Now, we have shown that the Formula of Universal Law, unlike the Formula of Humanity, does not run into paradoxes.

We can do even better. The Formula of Universal Law establishes a broad, imperfect duty to help more people rather than fewer. Suppose that you make it your maxim to help the few rather than the many when their needs conflict. This would not generate a contradiction in conception, because Kant thinks that even a world without beneficence is conceivable (*Groundwork* 4:423). Nevertheless, it would generate a contradiction in the will. Recall that you must will to be helped by someone. But some of our projects require the help of many others rather than just a few others. Thus, to pursue our ends, we often require the help of many people. But the only way to ensure that we receive the help required is if people are willing to help more rather than fewer people when such a dilemma emerges. Thus, to ensure that our ends are open, we must make it our maxim to help more rather than fewer people.

That argument might seem too slippery. I have not shown that there is a contradiction, only that there might be one. This is not a serious objection for me. In the argument for beneficence from the Formula of Universal Law, we did not guarantee that there would be a contradiction either. After all, it could be that you choose to only pursue things that you can get on your own. But the point is that because, as a limited being, you are overwhelmingly likely to pursue something that requires others' altruism, you then still run into a contradiction. There are cases in which you will to have the help of many others, not just one other; on the other hand, there are few, if any, cases in which you will to be helped by just one person. Although it is likely that you will not need the

help of many others in pursuing most ends, this point still shows that there is a stronger reason to help more people rather than fewer. . The lack of certainty here means that your duty to help many others is weaker than your duty to help someone. This dovetails with my following point, which is that a weaker duty need not to be fulfilled at every instance.

The argument from the previous paragraph follows the argument for the duty of beneficence from the *Groundwork*. As such, the argument grounds an imperfect and broad duty. Therefore, the duty does not require that in any particular situation you ought to save the many. It rather requires that you have a general policy of saving the many rather than the few. Thus, the duty to save the many over the few is weaker than the duty to save someone. Although the latter duty is also a corollary of the duty of beneficence, it establishes a necessary condition for what it means to be beneficent. In other words, you cannot claim to be beneficent unless you help those who need *your* help. Helping many rather than few, however, is merely being beneficent more effectively.

This conclusion might seem counterintuitive because it does not require you to always save the many. But I take this to be a benefit to the theory. To draw the point closer to intuition, it is helpful to invoke some Kantian terminology. It is virtuous for you to act in a way exemplifying your requirement by a broad duty. In this case, it is therefore virtuous for you to save the five rather than the one, because it corresponds to a broad duty. But because you have flexibility in fulfilling this duty, you do not act viciously when you fail to save the many. Your action of saving the one would only indicate “want of virtue, lack of moral strength,” since you do not strictly violate a duty, but nevertheless fail to do as much as you can (*Metaphysics of Morals* 6:390).

The distinction is therefore this. If you save no-one, you act viciously, since you act against duty. If you save the one person, you act neither viciously nor virtuously, since you act according to the requirements of duty but not beyond them. If you save the five, you act virtuously. This accords with common moral intuition. On one extreme, suppose that you must decide to save one person or two. We would fault you little if you chose to save the one rather than the five (e.g., if he were your friend). On the other extreme, say that now you can either save a billion people or just one person. If you were to choose to save the one, then we would fault you greatly. This is because it is hard to see how you can both be committed to a policy of generally helping more rather than fewer people while at the same time saving so few people. At this point, we could say that your duty is almost perfect and strict to save the many. Cases in between these extremes are scaled similarly; the strength of your duty to save the five varies. And this is exactly what we would expect from our intuition.

Granted, I have not yet argued that there exists an adequate Kantian account of reasons that can justify a spectrum of duties between perfect, strict duties and imperfect, broad duties. Providing such an account is beyond the scope of this paper. Nevertheless, that there should (or must) exist such an account is intuitive. There are cases between one in which you are the only person who can help (such as the situation examined in this paper), from which we derive a perfect, strict duty, and one in which you are one among many who can help (such as donating to charity), from which we derive an imperfect, broad duty. A Kantian account of reason should be able to account for the difference in strictness of duties falling between these two extremes.

The foregoing discussion also suggests that the Formula of Humanity and the Formula of Universal Law are inequivalent, contrary to what Kant claims. The two formulations generate different duties, and hence cannot be equivalent. This result is unsurprising, because the Formula

of Humanity and the Formula of Universal Law might generate different duties of beneficence. For example, Barbara Herman argues from the Formula of Universal Law that the duty of beneficence might be limited to certain subsets of rational agents who share the same vulnerabilities (such as humans). In this way, the duty of beneficence is really a duty of mutual aid between vulnerable rational beings. Angels, who lack many of the needs and weaknesses of humans, might not have a duty to help humans (Herman 590). By contrast, the argument for the duty of beneficence from the Formula of Humanity requires agents to help each other just because they are agents, not because they have needs or limits. Thus, because the duty to help the people on the tracks is at least related to the duty of beneficence, it is unsurprising that we arrive at different results from the two formulations.

We might view the Formula of Humanity as a more ideal formulation of a moral principle; it expresses a laudable ideal of human relationships, but due to the imperfection of the world, it might not provide an adequate account of action in the world real agents inhabit. Christine Korsgaard argues for this point using the example of the murderer at the door. In that case, the Formula of Humanity definitively grounds a duty to not lie to the murderer; by contrast, the Formula of Universal Law might permit one to lie. Her conclusion is that the Formula of Humanity is an ideal theory; that is, theory for conduct within the Kingdom of Ends, in which evil no longer exists (Korsgaard 349). The Formula of Universal Law, by contrast, is non-ideal—it guides conduct in a world with evil. My conclusion illustrates a different aspect of how the Formula of Humanity is ideal than Korsgaard does through the murderer at the door example. Tradeoffs between saving one person and five involve no evil, so they could be realized in the Kingdom of Ends. Nevertheless, we still find a difference in what the Formula of Humanity and the Formula of Universal Law prescribe. Therefore, the Formula of Humanity is also ideal in the sense that it does not always account for

non-ideal circumstances generated by physical limitations, such as those present in the tradeoff situation considered in this paper. The Formula of Universal Law therefore seems to be the right principle to guide action in a world with both evil and unfortunate tradeoffs.

Works Cited

Herman, Barbara. "Mutual Aid and Respect for Persons." *Ethics*, vol. 94, no. 4, 1984, pp. 577–602. *JSTOR*, www.jstor.org/stable/2380319. Accessed 20 May 2021.

Kant, Immanuel. *Kant: The Metaphysics of Morals*. Edited by Lara Denis. Translated by Mary Gregor, 2nd ed., Cambridge University Press, 2017.

Kant, Immanuel, and Christine M. Korsgaard. *Kant: Groundwork of the Metaphysics of Morals*. Edited by Mary Gregor, Cambridge University Press, 1998.

Korsgaard, Christine M. "The Right to Lie: Kant on Dealing with Evil." *Philosophy & Public Affairs*, vol. 15, no. 4, 1986, pp. 325–349. *JSTOR*, www.jstor.org/stable/2265252. Accessed 20 May 2021.

Taurek, John M. "Should the Numbers Count?" *Philosophy & Public Affairs*, vol. 6, no. 4, 1977, pp. 293–316. *JSTOR*, www.jstor.org/stable/2264945. Accessed 20 May 2021.

5.22 - 5.23
2021

Philosophizing (in) Uncertainties

@ NYU
Annual Undergraduate
Philosophy Conference

Double Standards of Rape: The Paradox of She Said, He Said Cases

Charlotte Goldberger

McGill University

May 2021

Abstract

This essay uses Georgi Gardiner's definition of 'she said, he said' (SSHS) cases in "She Said, He Said: Rape Accusations and the Preponderance of the Evidence," and the two standards used with respect to Gardiner's paradox, to analyze SSHS cases. Currently, the two main standards employed are the 'preponderance of the evidence' standard (PE) and the 'clear and convincing evidence standard' (CCE). Using Gardiner, I argue, in objection to CCE, for a third standard, named here the 'defined responsibility standard' (DRS), to be used in the resolving of the paradox.

To do so, I first define Gardiner's paradox and what SSHS cases are. Next, I expound on DRS, focusing on what makes it the best option to solving the paradox. From this, I outline the argument for why CCE helps solve Gardiner's paradox better than PE by presenting the Hofstra University case as an example of false rape accusations, and why a higher standard of evidence is necessary to legitimize severe sanctions like expulsion. I later expose flaws in CCE in relation to the paradox and SSHS cases more generally, notably by bringing up statistics demonstrating that false rape accusations are minimal and that cases—such as that of Hofstra University—do not warrant a higher burden of proof on the plaintiff. Finally, I offer an argument in favor of DRS,

describing how it helps solve the paradox in a way both CCE and PE cannot do, and briefly note that regardless of the standard, rape culture is what needs to change, which is best done by combining PE and CCE.

Double Standards of Rape: The Paradox of She Said, He Said Cases

Introduction

In “She Said, He Said: Rape Accusations and the Preponderance of the Evidence,” Georgi Gardiner, in the context of American universities, discusses ‘she said, he said’ cases (SSHS), their inherent paradox, and the two existing main standards of evidence. Gardiner (2019) defines SSHS as rape accusations and “corresponding denials” that have no additional, significant “case-specific” evidence, such as third party witnesses or credible alibis (p. 2). Currently the ‘preponderance of the evidence’ standard (PE) and the ‘clear and convincing evidence’ standard (CCE) have been employed in Title IX sexual harassment and violence investigations, including those of rape (p.1). According to the Legal Information Institute, under PE, the burden of proof is met when the accusing party is able to persuade or prove to the “fact finder” that there is an over 50% chance the allegation made is true (“Preponderance of the Evidence,” n.d.). Similarly, CCE means that evidence brought forward is “highly and substantially more likely to be true than untrue,” and that the “fact finder” has to be persuaded that the allegation is “highly probable” (“Clear and Convincing Evidence,” n.d.). Gardiner argues that a paradox arises where, using PE, severe sanctions (e.g. expulsion from the university)¹ are both legitimate and illegitimate. In this paper, I argue that the best solution to this paradox is a new standard, which I will hereby refer to as the

¹ To maintain word count, I will assume that severe sanctions are the only options when it comes to SSHS cases.

‘defined responsibility standard’ (DRS). The DRS combines elements of PE and CCE, arguing for a lower standard of evidence for the plaintiff and a higher one for the defendant.

To do so, I will first define the paradox and show that combining the standards into the DRS would prove beneficial, especially to victims, and help start to solve the paradox. Next, I will consider the objection that, due to worries of false rape accusations (FRAs), CCE should be the employed standard, and use the 2009 Hofstra University case in which a group of students were falsely accused of rape as an example of the impact FRAs have. Then, I will respond to this objection, explaining that while FRAs are impactful and problematic, they rarely occur. Additionally, CCE makes it harder for victims to feel comfortable bringing forward allegations, due to the higher burden of proof placed upon them, making this standard a questionable choice when underreporting is already an established, and major, issue. Finally, I will conclude and briefly point out that (i) both the people objecting to and the people suggesting a new standard are fighting for the same thing—the end of rape culture, and (ii) standards should not be imposed, they should be modified on a case-by-case basis depending on case-specific evidence available.

Gardiner’s Paradox

Gardiner (2019) lays out six claims² to demonstrate the paradox, with her first claim being that PE should “govern Title IX investigations” (p. 1). From there, she explains claims 2-6. Claim 2 says that PE is met if the accusation is most likely true based on evidence presented and assuming an investigation was properly carried out (p. 2). Claim 3 says that in most SSHS cases, accusations are most likely true based on evidence available and claim 4 that formal verdicts of rape can cause “legitimate considerable consequences” for the accused (p. 2). Claim 5 notes that in some cases,

²She refers to them as Claims A-F, but for clarity I will use numbers.

one versus one incompatible testimony, like in SSHS cases, makes it so that severe sanctions are illegitimate, due to the case relying on competing testimonies and not “individualised evidence” (p. 2). Consequently, as claim 6 points out, a paradox arises whereby considerable sanctions are both legitimized, based on 3 but also not legitimized, based on 5 (p. 2). However, as Gardiner points out, PE is problematic, as it is “too low” of a burden of proof, meaning that a higher standard, such as that of CCE is necessary, in the governance of Title IX investigations (p. 1).

Solving the Paradox: Defined Responsibility Standard Versus Existing Standards

Combining PE and CCE into the DRS by lowering the standard of evidence for victims and increasing it for the accused would help start to solve the paradox by making severe sanctions like expulsions more substantiated. By increasing the standard of evidence for the defendant, rigorous evidence would still be provided in line with CCE, thereby substantiating severe sanctions should the defendant be unable to convincingly prove their innocence. Lowering the standard of evidence for the alleged victims would take into consideration that sometimes evidence cannot be provided by victims, perhaps due to trauma or drugs (e.g. ‘date-rape drugs’ such as gamma-Hydroxybutyric acid, otherwise known as GHB).

Further, currently and historically, the majority of rapes have gone unreported (Kimble, 2018; National Institute of Justice, 2010), due to victims fearing retaliation, not wanting others to know, feeling shame and guilt, believing that police or campus officers will not do anything about it, etcetera (Kilpatrick et al., 2007; Wolitzky-Taylor et al., 2011). Consequently, having different standards of evidence for the plaintiff and defendant would maintain the legitimacy of sanctions, as proof would still need to be provided by both accuser and accused, while more appropriately

reflecting each party's capacity to present evidence.³ Thus, the paradox would be solved insofar as sanctions would be legitimized.

The Case for CCE and Against PE

However, an objection can be made that CCE should be the employed standard in rape cases as it is a more "rigorous" standard than PE ("Clear and Convincing Evidence," n.d.), which would help solve the paradox because severe sanctions would be better legitimized compared to PE. Additionally, lowering the standard of evidence for accusers would (i) automatically place a bias against the accused, due to them needing to prove their innocence while accusers are seemingly exempted from it, and (ii) make it easier for plaintiffs to falsely accuse people of rape, such as in the Hofstra University case, as it would be easier for them to bring up rape accusations. The case was brought on by Danmell Ndonge, an 18 year old student at Hofstra, after she claimed to have been gang-raped in the bathroom of a dormitory by five men (The Associated Press, 2009). After four out of the five were arrested and the fifth, through an attorney, offered video evidence proving the defendants' innocence, Ndonge admitted to making up the accusation due to fearing her boyfriend would view her as a "slut" (Crowley, 2009).

³ Lowering the standard would also mean that less evidence would need to be provided by victims, which would, in turn, allow them to more easily bring forward rape allegations, both case-wise and emotionally, as they would no longer need to shoulder the burden of proof.

Anna Rittgers (2011) states that cases like SSHS, when using PE, “[force] colleges and universities to abandon traditional notions of due process.” According to the Fifth Amendment of the U.S. Constitution, liberty cannot be denied without “due process of law” (U.S. Const. amend. V). It could be argued that severe sanctions such as expulsion or requiring the defendant to serve jail time count as infringing upon one’s liberty. Abandoning due process renders aforementioned sanctions illegitimate due to the defendant not being given due process. Using CCE would better solve the paradox as it would mitigate the legitimacy of evidence issue raised in the paradox and reduce FRAs, by ensuring that evidence brought forward is more rigorously analyzed.

Objecting to the Objection: FRAs’ Lack of Statistical Importance

Although theoretically more likely to have accurate results, compared to PE, ensuring that the threshold for significance of evidence is higher does not outweigh the cost of having schools not taking on cases—such as rape that happens between faculty and/or student members of a university outside of campus (e.g. a private apartment)—meaning that students in these instances would likely be left without resources. As well, the fear of rampant false rape accusations is disproportionate with their frequency.

Based on reports and studies conducted, research shows that only approximately 2-10% of rape cases were found to be false (Lisak et al., 2010; Lonsway et al., 2018; National Sexual Violence Resource Center, 2012). Importantly, this percentage includes cases deemed false because they were inconclusive or evidence was considered insufficient and, further, differing perspectives on what ‘false’ means results in the variability of FRA statistics (Rumney, 2006, pp. 130-132). According to a report published by the National Academies of Sciences, Engineering,

and Medicine, there is “ample evidence that the crimes of rape and sexual assault are substantially undercounted through police reports,” and that reasons as to why rapes and assaults are undercounted is partly due to victims not reporting these crimes and partly from the way law enforcement “handles both the victims and the police reports of those crimes” (Kruttschnitt et al., 2014, p. 36). While seemingly promising, CCE is a flawed solution to Gardiner’s paradox as it fails to consider that FRAs are not common enough to require a higher standard of evidence. Rather, CCE likely dissuades victims from coming forward, as it makes it more difficult for their assaulters to be brought to justice. Severe sanctions like expulsion therefore remain justified when combining CCE with PE into the DRS, as a high standard of evidence would remain, legitimizing the sanctions, without deterring victims.

Conclusion

In this paper I have argued that CCE is not a sufficient answer to Gardiner’s paradox, and a better way of approaching it would be to combine PE and CCE into DRS by lowering the standard of evidence for accusers and increasing it for the accused. Doing so would enable victims to bring in as much evidence as they can without penalizing them for having endured trauma, and place the burden of proof upon the accused. This ensures that severe sanctions like expulsion are legitimized based on the amount and convincingness of evidence brought forward while also mitigating underreporting. Though not perfect, combining PE and CCE into the DRS would enable victims to feel more comfortable bringing forward rape allegations and ensure that accuser and accused are granted the opportunity to state their case. However, as Amanda Hess (2009) notes, regardless of the standard used, in the end, rape culture is what needs to be rectified, something both PE and CCE strive to achieve. Thus, rather than making PE and CCE compete, they must be reconciled, as suggested in this paper.

Works Cited

- Crowley, K. (2009). Twisted Motive Behind Rape Story. Retrieved from <https://nypost.com/2009/09/18/twisted-motive-behind-rape-story/>
- Gardiner, G. (2019). *She Said, He Said: Rape Accusations and the Preponderance of the Evidence*. Unpublished manuscript. Philosophy Faculty, University of Tennessee.
- Hess, A. (2009). False Rape Accusations and Rape Culture. Retrieved from <https://washingtoncitypaper.com/article/394081/false-rape-accusations-and-rape-culture/>
- Kilpatrick, D. G., Resnick, H. S., Ruggiero, K. J., Conoscenti, L. M., & McCauley, J. (2007). Drug-facilitated, Incapacitated, and Forcible Rape: A National Study. Retrieved from <https://www.ojp.gov/pdffiles1/nij/grants/219181.pdf>

Kimble, C. (2018). Sexual Assault Remains Dramatically Underreported. Retrieved from <https://www.brennancenter.org/our-work/analysis-opinion/sexual-assault-remains-dramatically-underreported>

Kruttschnitt, C., Kalsbeek, W. D., & House, C. C. (Eds.). (2014). Estimating the Incidence of Rape and Sexual Assault. pp. 35-48. The National Academies Press.

doi:10.17226/18605. Retrieved from <https://www.nap.edu/read/18605/chapter/1>

Legal Information Institute. (n.d.). Clear and Convincing Evidence. Retrieved from https://www.law.cornell.edu/wex/clear_and_convincing_evidence

Legal Information Institute. (n.d.). Preponderance of the Evidence. Retrieved from https://www.law.cornell.edu/wex/preponderance_of_the_evidence

Lisak, D., Gardinier, L., Nicksa, S. C., & Cote, A. M. (2010). False Allegations of Sexual Assault: An Analysis of Ten Years of Reported Cases. *Violence Against Women, 16*(12), 1318–1334. <https://doi.org/10.1177/1077801210387747>

Lonsway, K. A., Archambault, J., & Lisak, D. (2018). False Reports: Moving Beyond the Issue to Successfully Investigate and Prosecute Non-Stranger Sexual Assault. Retrieved from <https://www.nsvrc.org/sites/default/files/publications/2018-10/Lisak-False-Reports-Moving-beyond.pdf>

National Institute of Justice. (2010). Reporting of Sexual Violence Incidents. Retrieved from <https://nij.ojp.gov/topics/articles/reporting-sexual-violence-incidents>

National Sexual Violence Resource Center. (2012). False Reporting - Overview. Retrieved from

https://www.nsvrc.org/sites/default/files/2012-03/Publications_NSVRC_Overview_False-Reporting.pdf

Rittgers, A. (2011, September 07). RITTGERS: Sometimes, Women Lie About Rape.

Retrieved from

<https://www.washingtontimes.com/news/2011/sep/7/sometimes-women-lie-about-rape/>

Rumney, P. N.S. (2006). False Allegations of Rape. *The Cambridge Law Journal*, 65(1),

128-158. Retrieved from <http://www.jstor.org/stable/4509177>

The Associated Press. (2009). No Charges Planned in Hofstra Rape Hoax. Retrieved

from <https://www.nytimes.com/2009/09/26/nyregion/26hofstra.html>

U.S. Const. amend. V.

Wolitzky-Taylor, K. B., Resnick, H. S., McCauley, J. L., Amstadter, A. B., Kilpatrick, D.

G., & Ruggiero, K. J. (2011). Is reporting of rape on the rise? A comparison of women

with reported versus unreported rape experiences in the National Women's

Study-Replication. *Journal of interpersonal violence*, 26(4), 807–832.

<https://doi.org/10.1177/0886260510365869>

5.22 - 5.23
2021

Philosophizing (in) Uncertainties

 @ NYU
Annual Undergraduate
Philosophy Conference

Carnap and Impredicative Definitions

Junyan Jiang

New York University

May 2021

Abstract

This essay is an evaluation of Rudolf Carnap's 1931 work "The Logicist Foundations of Mathematics," presented in a symposium on the foundations of mathematics. The article consists of four sections. It first offers an overview of the logicist project — i.e., the reduction of all of mathematics to logic — and is further divided into two interrelated components: the reduction of mathematical concepts to logic & the reduction of mathematical theorems to logic, treated respectively in sections 1 and 2 of the article. Then Carnap brings in the problem of impredicative definitions; he sees it as the most destructive challenge to logicism and argues against two existing solutions by Russell and Ramsey. In section 4 he proposes his own attempt at salvaging logicism, claiming that "if the theory just sketched proves feasible, logicism will have been helped over its greatest difficulty."

After providing an outline of Carnap's arguments, I contend that the dilemma posed by impredicative definitions is in fact ontological. By drawing on his own analysis of logicism's connections to intuitionism and formalism, I supply a semantic perspective as opposed to Carnap's "formal" point of view, revealing precisely where Carnap's suggestion surpasses "traditional"

logicism. This will allow us to conclude that despite his ingenuity, Carnap avoids the ontological problem altogether, making his proposal a less than satisfactory alternative for any “hard-core” logicist.

Carnap and Impredicative Definitions

In an essay titled “The Logician Foundations of Mathematics”¹—originally part of a 1931 symposium on the foundations of mathematics—Carnap quotes an exemplary impredicative definition to illustrate its serious threat to logicism. In particular, a property is said to be “hereditary” if: once it belongs to the number n , it necessarily belongs to the number $n + 1$. “Inductive numbers” are then defined as those numbers with all hereditary properties of 0, i.e.:

$$\text{Ind}x := \forall f (\text{Her}f \wedge f(0) \rightarrow f(x)).$$

The danger of the above definition is that the universal quantifier on the right-hand side quantifies over *all properties* f . That is, even Ind itself could be substituted for f in the definition, causing an implicit circularity that allows the concept being defined to be potentially contained in the logical structure used to define it. So the definition is obviously useless, because we are defining something in terms of itself—and this is why Russell calls such impredicative definitions “vicious circles.” Carnap gives a concrete manifestation of such circularity. According to him, “it is sometimes claimed” that the absurdity of impredicative definitions could be demonstrated by simply plugging in a random number. For example, suppose we want to confirm that 2 is inductive, then we need to check whether all properties f satisfy the formula within the scope of the universal quantifier.

¹ Carnap, Rudolf, translated by Erna Putnam & Gerald R. Massey. “The Logician Foundations of Mathematics.” In *Philosophy of Mathematics: Selected Readings (Second Edition)*, edited by Paul Benacerraf & Hilary Putnam. Cambridge University Press. 1984. pp. 41-52.

Since Ind is unquestionably a particular instance of f , we are sure to encounter it at some point during this process. Thus in order to check whether Ind2 holds, sooner or later we need to substitute Ind into the expression on the right-hand side, checking one by one whether Ind is hereditary, whether it belongs to 0, and most importantly: whether or not Ind2 holds.

This last point, no doubt, posed insurmountable difficulties. Russell was eventually forced to circumvent impredicative definitions altogether by ramifying the already complex type theory, which nearly cost him the entire kingdom of real numbers: most of them cannot even be expressed impredicatively. Such painful losses in turn led Ramsey, Russell's pupil, onto a path that attempted to shake off all philosophical dilemmas with a quasi-Platonist "theological mathematics" that directly allows impredicative expressions. The underlying rationale proceeds like this: because the existence of mathematical properties is irrelevant to our perceptions, the way we define them does not affect these concepts themselves, and so on. Carnap does not empathize with either approach. He wants to somehow "have Ramsey's results without retaining his absolutist conceptions." To this end, he revisits the verification procedure for Ind_x given x : the problem arises when one checks "every single property." Only when we consider this step necessary do we hit the "unbreakable circle" and "run headlong against the property 'inductive.'" From Carnap's perspective, there is really nothing inevitable about this belief:

[T]he verification of a universal logical or mathematical sentence does not consist in running through a series of individual cases, for impredicative definitions usually refer to infinite totalities. The belief that we must run through all the

individual cases rests on a confusion of “numerical” generality, which refers to objects already given, with “specific” generality. We do not establish specific generality by running through individual cases but by logically deriving certain properties from certain others. (§4)

By interpreting the universal quantifier in the definition of inductive numbers as referring to “particular generalities,” Carnap actually *syntacticized* the verification of Ind_x. For instance, if we still want to confirm the truth of Ind₂, i.e., “whether 2 is inductive,” all we need to do is to prove

$$\vdash \text{Herf} \wedge f(0) \rightarrow f(2).$$

That is: f is reduced to a pure symbol, and its function-meaning here is not that of the variable x in an *equation* (“numerical” generality), but the variable x in *logic* as a well-formed formula (“specific” generality). The verification of Ind₂ is thus transformed into a simple exercise in syntax. First, Ind₀ clearly holds: $\text{Herf} \wedge f(0)$ certainly implies $f(0)$. We then note the equivalence between Herf and $\forall n (f(n) \rightarrow f(n + 1))$, which, combined with

$$\forall n (f(n) \rightarrow f(n + 1)) \vdash f(0) \rightarrow f(1),$$

gives

$$\text{Herf} \wedge f(0) \vdash f(1).$$

Repeat the same steps and we arrive at Ind₂ as required.

Is Carnap’s “attempt at a solution”² satisfactory? Not necessarily. The typical logicist’s antipathy toward the definition of “inductive numbers” stems, after all, from the way it is defined; the prime motivations behind their concern are in fact ontological. As Carnap himself outlines in the beginning paragraphs of his essay, the goal of logicism can be split into two components: first, to

derive all mathematical concepts from logic; second, to derive all mathematical theorems from logical axioms in a purely logical way. Although together these two goals point to the ultimate task of “reducing all mathematics to logic,” they are nonetheless two different sentences. Such a distinction itself implicates the following logicist perspective on being: that mathematical concepts and mathematical theorems are different types of objects, the former being logical constructions, while the latter, at least intuitively understood, are statements *about* the former.³ Moreover, for mathematical concepts, all but the most fundamental ones are allowed to have no explicit logical construction; everything else is mandated to be explicitly defined in terms of logical constructions whose existence has been firmly established. That is: the whole point of the defining *equation* is to use the ontologically unambiguous object on the right-hand side to transfer its ontological clarity to the ontologically still-dubious object on the left-hand side through the “divine action” of the equal sign.⁴ As Carnap rightly observes, such a “constructivistic tendency” brings the logicists — as far as their attitude toward mathematical objects is concerned—closer to the intuitionists. It is also this quest for existential clarity that made Russell reject the vicious circle of impredicative definitions: once we allow a concept to define itself, the right-hand side of the defining formula irrevocably loses its full determinacy as a logical being: it is contaminated by the ontologically still-dubious object on

² The title of section 4 in his essay.

³ “Ordinary mathematics” certainly endorses this dichotomy, perhaps to a greater extent than the logicists; one

could argue that it was “elevated to canonicity” when the distinction between “terms” and “well-formed formulas” as separate classes (more accurately, sets) of objects became commonplace in introductory textbooks of logic. And this isn’t the end: the latter is placed higher in the conceptual-logical hierarchy because terms are literally the building blocks of well-formed formulas.

⁴Here “being” should not be taken to mean the “being” that logicism explicitly rejects; Russell, for example, famously remarked that mathematics deals only with the “possibility” of being. In our context, it is not so much “being” *per se* that logicism opposes, but rather the Platonic sense of the term. After all, logical structures can be seen as one particular species of being-object, even if its specific status in mathematical ontology is not yet clear. It is in this sense that we employ the term “being.”

the left-hand side of the equation; alternatively: the sanctity of the equals sign is itself violated.

On the other hand, apropos of mathematical theorems, logicians have instead a stronger “methodological affinity” with formalism:

Logicism proposes to construct the logical-mathematical system in such a way that, although the axioms and rules of inference are chosen with an interpretation of the primitive symbols in mind, nevertheless, *inside the system* the chains of deductions and of definitions are carried through formally as in a pure calculus, i.e., without references to the meaning of the primitive symbols. (§4)

We should note that Carnap extends the aforementioned affinity even to the level of definitions. This apparently justifies his syntacticized interpretation of the notion of inductive numbers, since according to him, formalization is inherently part of the orthodox methodology of logicism; hence it should be so when it comes to definitions as well. His previous defense of impredicative definitions, indeed, follows the same lines: although Russell argues that allowing

impredicative definitions necessarily leads to antinomies (in *practice*) — which, in the original context, *might* be a syntactic concept—this is not necessarily the case. We cannot, at least, derive the inevitability of antinomies from any fundamental principles. In fact, under most circumstances, there's nothing logically contradictory about impredicative definitions; the real numbers serve as an excellent case: everyone loves them. Another example is the inductive numbers mentioned earlier, which, even if it drove Russell crazy, still managed to achieve its original purpose; i.e., all natural numbers can indeed be *confirmed* to be inductive. Carnap himself puts it as follows:

We see then that the definition of inductiveness, although impredicative, does not hinder its utility. That proofs that the defined property obtains (or does not obtain) in individual cases can be given shows that the definition is meaningful. If we reject the belief that it is necessary to run through individual cases and rather make it clear to ourselves that the complete verification of a statement about an arbitrary property means nothing more than its logical (more exactly, tautological) validity for an arbitrary property, we will come to the conclusion that impredicative definitions are logically admissible. (§4)

This passage spells out Carnap's true position. As he admits earlier (it is necessary that we "have Ramsey's result"), his first and foremost objective is to testify to the practicality of impredicative definitions from a pragmatic point of view. Moreover, he makes a conceptual shift from discussing "definitions" to speaking of "confirmations," thereby reducing the "meaning problem" of the former to that of the latter. In this way, Carnap successfully avoids the ontological issues of logicism through a "turn to formalism."

Yet the ontological problem remains, and the gap leading up to it is opened up precisely by Carnap himself in a pair of parentheses (“more exactly, *tautological*”). Regardless of how he avoids metaphysics by resorting to some kind of “formalism,” as soon as we re-examine both sides of the argument from a semantic point of view, we will find that even purely formal operations have ontological consequences due to their shadowy semantic counterparts. This is a fateful fact that follows from the reliability theorem, a property any meaningful deduction system should satisfy.

More specifically, for any x , determining the *truth value* of Ind_x by running through every property qualifies as a paradigmatic application of model theory’s definition of truth. $\forall f$ $(\text{Her}f \wedge f(0) \rightarrow f(x))$ is true if and only if $\text{Her}f \wedge f(0) \rightarrow f(x)$ holds for all f . Traditional logicians stop right here because for them not every f in the given domain has a clear ontological status, which is synonymous to admitting that there are instances of f such that the truth value of $\text{Her}f \wedge f(0) \rightarrow f(x)$ *cannot be determined*. Ind itself is a perfect example. Therefore they claim that the truth value of Ind_x is also uncertain; as a result the existing definition of Ind must be meaningless. The hesitation that the logicians exhibit while endeavoring to assign a truth value to Ind helpfully reflects their metaphysical and epistemological position. We see that, at least with regards to their “stubborn” insistence on the ontological clarity of defined objects *and* our knowability of it, they resemble the intuitionists more than the formalists—which is consistent with our previous observations.

Through avoiding the above issues and syntacticizing them, Carnap indeed surpasses traditional logicians. He suggests that so long as we could prove $\text{Her}f \wedge f(0) \rightarrow f(x)$ —a more illuminating formulation would be: so long as we could prove $f(x)$ *given* $\text{Her}f \wedge f(0)$ —we are safe to declare that Ind_x is true. Semantically, the equivalent statement is: for any property f , Ind_x is true if $\text{Her}f \wedge f(0) \models f(x)$.

Now Carnap's contribution lies in the remarkable fact that the two-layer hypothetical structure here (both within the formal language and within the meta-language) is actually able to eliminate every problem on the surface. If the truth values of HerInd, Ind0 and Indx do exist, then the whole expression would turn out true given any possible T-F configuration of these three terms, permitting us to arrive at the truth of Indx unimpeded. Even if their truth values do not exist, we can still obtain "vacuous truth" from the meta-structure "A if B."

Unfortunately, Carnap's ingenious maneuver fails to resolve the real problem. If we do not know for sure whether the truth values of those terms truly exist, we cannot ascertain the truth value of "the truth values of those terms exist" either. There's nothing ontologically refreshing. The only difference is, we have "upgraded" the central difficulty involved by a single "order" through Carnap's turn to formalism. This reminds one of Russell and Whitehead's reluctant response when they were confronted with set theory's existential axioms: since we don't know whether some desirable set S exists or not, anything derived from the existence of S could only be formulated in the form "if S exists, then" We could thus characterize Carnap's proposal as a more subtle and esoteric compromise that was nevertheless in line with the logicist tradition.

5.22 - 5.23
2021

Philosophizing (in) Uncertainties

@ NYU
Annual Undergraduate
Philosophy Conference

Which Came First, the Camera or the Photograph?

Aashutosh Mukerji

Boston University

May 2021

Abstract

In his book, *Four Arts in Photography: An Essay in Philosophy*, philosopher Dominic Lopes proposes a new definition of photography in response to the controversial argument that photography is not art. Perhaps most famously championed by photographic skeptic Roger Scruton, the claim is that the purported art of photography is nothing more than the mind-independent tracking of the objective features of a scene, with no room for personal artistic influence. Lopes' new theory of photography demonstrates that Scruton's understanding of photography is underdeveloped and misguided. The new definition characterizes photography as a process with several stages, rather than an instantaneous event that occurs when a shutter button is pressed. All of these stages seem to provide ample opportunity for the artist to display their intentions and influence in. However, though this new definition furnishes more room for the existence of photographic art, it is open to a host of significant objections. In this paper, I consider the objection that it is overinclusive, which I found to be the most weighty critique of Lopes' work. I ultimately conclude that though the overinclusive objection leads us to believe that the new definition is counterintuitive, in reality, Lopes' theory is not harmed by it in any meaningful way.

Which Came First, the Camera or the Photograph?

In the vibrant field of aesthetics, there are many conflicting accounts of what constitutes photographs and the (contested) art of photography. Some, like skeptic Roger Scruton, have boldly asserted that photographs cannot be art. This charge is usually grounded in the claim that photography by itself cannot do anything except track the relevant features of a scene in a mind-independent fashion. Unlike other art, the claim goes, photographs cannot reflect the attitudes, thoughts, feelings, etc., of the photographer. While Monet's paintings might easily lead us to the inference that he found water lilies remarkably serene, a photograph of that same scene might reveal nothing other than the way in which the place objectively appeared at the precise instant the photograph was captured.

This argument may seem vaguely plausible on the surface, but seems to fall apart quite easily when our understanding of photography is examined. As it turns out, skeptics like Scruton harbor misleading illusions about what exactly is involved in the process of photography. Most notably, as Dominic Lopes points out, it is exactly that: a process. Scruton seems to operate under the mistaken assumption that the production of the photograph is an instantaneous and purely mechanical event that is unaffected by the intentions of the photographer in question. In reality, photography has several distinct stages, each of which allow for a great breadth of intentional choices (usually on the part of the photographer) that drastically affect the outcome of the image. I will briefly summarize these stages, which Lopes details fully in his book, *Four Arts in Photography: An Essay in Philosophy*.

First, we have what is called the “pro-photographic scene,” which refers to the actual state of affairs of a particular place in the world that stands before a photographic apparatus. Next, light from the scene is (usually) directed through an aperture and lens. This is what Lopes refers to as the “light image,” which is different from the final photograph in that it has not yet been recorded and it can change over time as the inputted light shifts. Thirdly, we have the recording of this light image, occurring as a raw file on a digital camera, an undeveloped photograph, or something of this nature. Lastly, there is the creation of what we know as the photograph. The raw data file is processed by software to create a visual image, or the photograph is developed and can be viewed (Lopes 79-80). As a result of these steps, Lopes proposes a new way in which we might define photography that places special emphasis on it not being a single event: “A photograph is an image output by a mark-making process taking input from an electro-chemical event that records information from a light image of a pro-photographic scene” (Lopes 81).

The principal point from this more accurate detailing of the photographic process is that the photographer can intervene at any of these steps and exercise artistic agency that is very much like that which a painter would. To name just a few examples, the aperture size, shutter speed, and recording mechanism are all places where photographers can affect photographic outcomes in such a way as to represent their beliefs or intentions in their work.

So now we know photographs can be art. We’ve contrived to create a new definition that leads us back to what anyone with a basic understanding of art and photography knows. For the remainder of this paper, however, I’d like to take a closer look at Lopes’ new and improved definition and consider an objection that it might be too inclusive. To be clear, though I favor

the new Lopes definition over the standard (mind-independent feature tracking) definition that Scruton presupposes, the purpose of this paper is not to advocate for either. Rather, this paper simply aims to analyze a particular objection to the new theory and explain why it ultimately fails. My reader can draw their own conclusions about the overall stability of the new and old theories from this objection if they so desire, but that is beyond the scope of this paper.

A common formulation of the over-inclusive objection is considered by Dawn Phillips when she lays out her own definition of photography, which bears many similarities to Lopes'. She asks us to imagine the following: If a patterned sheet of fabric is laid over a piece of driftwood, and exposed to strong sunlight for an extended time, this process might result in the same pattern now becoming visible on the wood. By taking this just a tad further, there seems to be no great logical leap in temporally placing the creation of this particular piece of driftwood before the invention of the camera. The obvious counterintuitive result is that this would allow photographs to exist before cameras. Undaunted, Phillips is all too eager to bite the bullet on this; she readily accepts that this surely is a photograph without much additional explanation, and moves on (Phillips 339). Perhaps she simply does not care much for our intuitions, or perhaps she feels they are due for some restructuring anyway.

Though Lopes does not consider this sort of objection to his argument, such a piece of wood seems equally plausible to be photographic under his definition. A mark-making process is evident, and it certainly takes input from an electro-chemical event. All of this culminates in the recording of information from a light image. Lopes might deflect this objection by specifying the types of electro-chemical reactions that must be present in order for a photograph to be created, thereby allowing him to exclude this type of reaction. This route is a problematic one to take, however, as it forces Lopes to detail the exact set of reactions that constitute photography, which opens his new definition up to further counterexamples that do not clearly fit the boundaries Lopes may set. There is also the concern that it is overly ad hoc, and thus the significance of the category itself

may be diminished. Instead, I would invite Lopes to follow Phillips' lead and accept the piece of wood as a photograph. I argue that the existence of pre-camera photographs is less contrary to our intuitions than one may think.

First things first, the practice of photography must be distinguished from its product, the photograph. The practice involves an intentional act on the part of a human (or human-like creature), who consciously engages in the taking of pictures and understands the outcomes of his or her actions. The product, of course, is just the result of these intentional actions. Though one might think it is impossible to create the product without the practice, this is not the case. The proverbial monkeys might mash away at their keys and eventually produce the complete works of Shakespeare, but they are not writing, at least not in the way we conceive of that practice. Likewise, there is a nominal sense in which a dog who accidentally steps in ink and walks across a canvas is painting, but we understand that she is not really engaging in the practice we know to be painting. Nonetheless, a painting is produced and so is *Hamlet*. There seems to be no reason why this cannot be extended to our piece of driftwood. Of course, none of the objects or forces involved in the creation of the photograph are engaged in photography as a practice, but the photograph can exist independently of that.

One objection to this that might be posed is that we're allowing our definition of art to be undermined if we allow a piece of driftwood to be a photograph. After all, if the pattern on the wood is especially aesthetically successful, it may very well rival the aesthetic properties of photographs produced in standard fashion (i.e. through the use of a camera). If photographs are allowed to be created without human intervention, then we seem to quickly be sliding down the slippery slope of allowing the creation of all art without human input as well. What my objector might have confused, however, is the difference between photographs and art. We can require, as most do, that something must involve an intentional human action to be art; the mere

presence of aesthetic properties does not make something art. Sunsets, forests, and oceans may be beautiful, but none of them can be art in the way a photograph of them can. Likewise, the photographic driftwood with extraordinary aesthetic properties is not a work of art until a person sees it, deems it worthy of artistic status, and perhaps hangs it in a gallery next to other works of art. The slippery slope is thus levelled, our artistic pretensions are preserved, and we can still call the piece of driftwood a photograph without labelling it art.

Lopes' new definition of photography is useful in that it allows us to reaffirm that photography is indeed art, but opens itself up to the criticism of being too inclusive in the process. I propose a way in which some of those objections can be accepted without causing damage to the overall theory. Though this outcome may seem to be in conflict with our intuitions at first, proper consideration shows that this case is no different to those present in painting, writing, and numerous other art forms. Whether this allows the new theory to ultimately prevail over the old, or whether there are other objections to Lopes' theory that pose greater concerns, is something I leave to my reader to decide.

Works Cited

McIver, Dominic Lopes. *Four Arts of Photography: An Essay in Philosophy*. Wiley, 2016

Phillips, D. M. "Photography and Causation: Responding to Scruton's Scepticism." *The British Journal of Aesthetics*, vol. 49, no. 4, 2009, pp. 327–340., doi:10.1093/aesthj/ayp036



nyuaupc.org

The Annual Undergraduate Philosophy Conference is organized by **NYU Philosophy Forum**.

Visit us at

- nyuphilosophyforum.org
- as.nyu.edu/philosophy/undergraduate/philosophyforum

All articles published in the AUPC Journal are fully open access: immediately freely available to read, download and share. All articles are published under the terms of the Creative Commons Attribution License. The Creative Commons Attribution License permits use, distribution and reproduction in any medium, provided the original work is properly cited and allows the commercial use of published articles.

