THE UBC JOURNAL
OF PHILOSOPHICAL ENQUIRIES

2022-2023
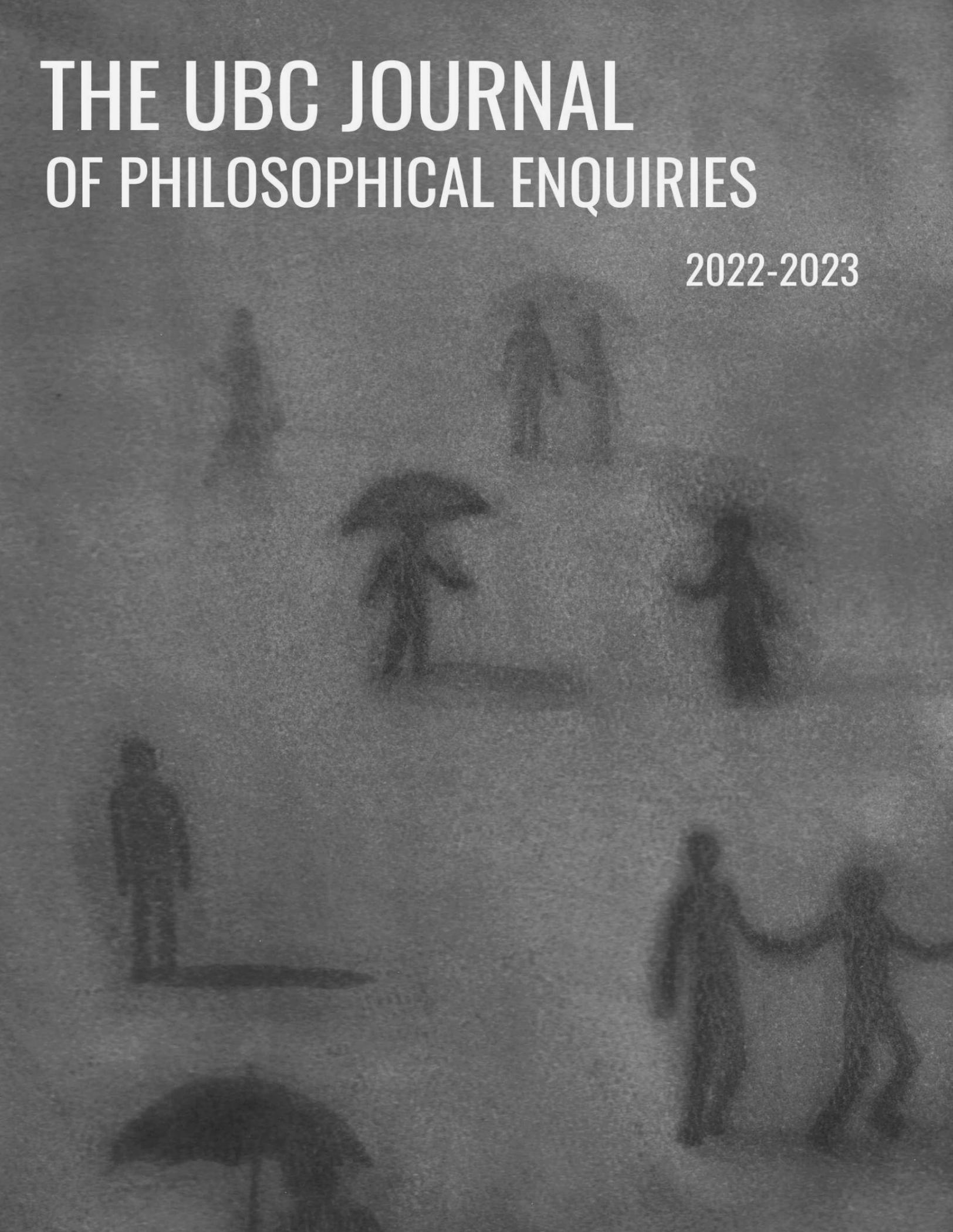
The UBC Journal of Philosophical Enquiries

Volume 3: 2022-2023

This publication is produced in association with the Philosophy Students'
Association and the Department of Philosophy at the University of British Columbia.

Editorial Note

Writing good academic essays is no easy task. Anyone who has ever attempted can testify to the many hours spent reading, thinking, thinking twice, drafting and redrafting, all to produce a few polished pages. And then there is the doubting, removing or re-working of paragraphs and sometimes entire arguments and sections. But it is such an essential task, the value of which one appreciates in retrospect. For we think through the medium of language, and the more supple, intricate, powerful and dynamic this medium is, the better we are not only at sharing our thoughts with others, but with conceiving and shaping the thoughts themselves.

To endeavour to write well is thus to endeavour to think more lucidly and profoundly. The following essays written by undergraduate students are no small achievement, rather, the fruits of an extended process of selection and editing vis-a-vis our team of student editors. These talented authors were not only given the opportunity to enliven the undergraduate academic culture in the humanities at the University of British Columbia, but also to develop themselves as scholars by committing to the publication of their work. This latter step entails a crucial leap from writing for oneself and one's professor to writing for a general audience of many different knowledge backgrounds, assumptions, interests, and purposes.

With all the more enthusiasm, we would like to present the second issue of the UBC Journal of Philosophical Enquiries since its revival in 2017. Despite its title changes throughout the years, the journal's mission has remained essentially the same: to provide a venue for showcasing exceptional and promising undergraduate academic work in philosophy and adjacent disciplines. We would like to thank our editors, without whose talent and dedication the journal would not be possible. Their names are listed on the previous page. Additionally, we thank the UBC Department of Philosophy for their consistent and faithful support, as well as the UBC Philosophy Students' Association for their financial and administrative assistance.

Sincerely,

Oliver Bontkes and Brandon Bach
Co-Editors-in-Chief

Table of Contents

The UBC Journal of Philosophical Enquiries

A Kantian Justification for Revolution: Applying Kant's Concept of "Private Use of Reason" and Just War Theory to Fanon's Argument for "Counter-Violence"

Mark Prestash

Argue ... but obey. – Immanuel Kant[1]

Decolonization, which sets out to change the order of the world, is clearly an agenda for total disorder. – Frantz Fanon[2]

In *What is Enlightenment*, Kant advocates a balance between thinking for oneself and the maintenance of social stability. Independent thought leads to independent action and should be curtailed insofar as the result of too much "civic freedom" might be social discord.[3] This view of Kant's appears to be in tension with Fanon's position in *The Wretched of the Earth*, where Fanon argues that, due to the inherently oppressive nature of colonial societies, violence is a legitimate form of political participation in which colonized subjects may participate in order to secure their freedom from their colonial masters. Despite the fact that these two theories seem to be in opposition, I will argue that they can in fact be reconciled. In part 1 of this paper, I will expand on the primary point of divergence between Kant and Fanon: the legitimacy of unilateral action. In part 2, I will argue that despite their apparent differences, Kant and Fanon both begin their analysis from a similar position: the idea that people should reject their "guardians" and become enlightened. Finally, in part 3, I will first examine Kant's just war theory, in which he argues that people are justified in using violence to resist invaders, and then I will examine Kant's theory of legitimate peace, in which he argues that unless the underlying issues of a conflict are resolved, a peace that results from one side being too exhausted to continue fighting is illegitimate. Then, I

---

[1] Immanuel Kant, "An Answer to the Question: 'What Is Enlightenment?,'" in *Kant: Political Writings*, ed. Hans Siegbert Reiss (Cambridge: Cambridge University Press, 1991), 59.

[2] Frantz Fanon, *The Wretched of the Earth* (New York: Grove Press, 2021), 2.

[3] Immanuel Kant, "An Answer to the Question: What Is Enlightenment?," in *What Is Enlightenment?: Eighteenth-Century Answers and Twentieth-Century Questions*, ed. James Schmidt (Berkeley: University of California Press, 2004), 7:63.

will argue that the decolonial revolution for which Fanon advocates can be conceived as the continuation of a just war of the sort sanctioned by Kant.

Part 1: The Use of Public and Private Reason

What is Enlightenment?

If Kant believed that social stability was of the utmost importance, why did he not simply advocate for the unquestioning obedience of individuals to the state? The reason is this: Kant imagined that there was another aspect of human life which was of great value: enlightenment, being the ability "to make use of one's own understanding without the guidance of another."[4] To become enlightened, people should free themselves from their "self-incurred immaturity," a condition in which they are afraid to form their own opinions.[5] When people are in this condition, they will defer to the judgment of self-appointed "guardians," people who will make decisions for them.[6] Enlightenment can only spread through a society once a few enlightened individuals begin to think for themselves, and once their peers notice and begin to think likewise. However, the concept of enlightenment seems to be in tension with the maintenance of social stability: the former seems to degrade the latter, and this leads Kant to develop his theory of the public use of reason.

The Public Use of Reason

Public use of reason is when we submit our ideas to "the entire public of the reading world,"[7] in order to persuade them to adopt our position.[8] Kant argues that once people achieve enlightenment, or at the very least, become more enlightened in certain fields, they should only exert their "public use of reason" to the degree that public order is maintained.[9] For example: it would be legitimate for an enlightened teacher to write a letter to a newspaper critiquing a new curriculum, but illegitimate for the teacher to refuse to teach the curriculum. If someone found

---

[4] Kant, trans. Schmidt, 58.
[5] Kant, 58.
[6] Kant, 58.
[7] Which I interpret to mean anyone who might consider these ideas.
[8] Kant, 59.
[9] Kant, 59.

themselves in a position where they "could not in conscience conduct his office," they should resign from it, rather than engage in what Kant called the "private use" of their reason.[10]

The Private Use of Reason

The private use of reason is when one has a duty to their society to obey the orders of another. This is exemplified for Kant by the case of the employee working in a "civic post or office" who acts according to the judgment of their employer: the public servant obeys in order to serve their society, and abstains from acting unilaterally, even when they disagree with their orders. In the case of the citizen, one has an obligation to, for example, pay one's taxes—and this is the exercise of private reason. Like the public servant, the citizen is obligated to obey the dictates of the wider community, and it is in this context that unilateral action by the individual is deemed by Kant to be inappropriate. To return to the example of the teacher, they would be acting inappropriately if, in their capacity as a public servant (which obligates them to the private use of reason), they refused to teach the official curriculum. If people were permitted by society to be governed solely by their own reason, and then to act unilaterally while remaining obligated to exercise only their private use of reason, this would cause massive social disruption: people would be able to break any laws they disagreed with, and any enterprise which required cooperation would quickly fall apart.

Objections to Kant

However, one might object that Kant's endorsement of public use of reason is due simply to a desire for knowledge for knowledge's sake: perhaps the public use of reason is an exclusively intellectual tool, unable to facilitate the kind of social change required to address unjust social hierarchies. Even if we grant that the general population can become enlightened, and so develop better and better political ideas, it seems that without having a mechanism in place in order to put such ideas into practice, such ideas would not be very useful. For example, if one were to develop a great economic theory which could end poverty, but the idea was never implemented, then it would accomplish little. I disagree with this view. Kant stated that a monarch's "legislative authority depends precisely upon his uniting the collective will of the

---

[10] Kant, 60-61.

people in his own."[11] Therefore, if the "collective will of the people" were to change through the public use of reason, the will of the government could and should change alongside it.[12] That is, a government only deserves to be obeyed if it is seen to be legitimate, and a government is only seen to be legitimate if it facilitates the change that its citizens decide should take place through the public use of reason. This is especially true for democracies, as people can use their public use of reason to shape the "collective will of the people," and the people can use elections to implement their will. The sort of social change facilitated by the public use of reason would for Kant be gradual, and would avoid the chaos that would occur if everyone were to act unilaterally while under obligation to society to make exclusive use of their private use of their reason. However, this gradual social change is only possible in societies where the public use of reason shaping the "collective will of the people" can result in political change, and in the colonial societies Fanon wrote about, this is not the case.

Fanon on the Private Use of Reason

Fanon argues that colonized people are justified in using violence in order to overthrow a colonial government, and that the successful use of violence in such a situation will result in empowering the colonized population. Unlike Kant, then, Fanon advocates for disobedience by the subjects of a colonial government in situations where that government expects them to obey. This is because a colonial regime interacts with its colonized subjects in primarily violent ways, and thus lacks the legitimacy which obligates the Kantian citizen to curtail their use of public reason. The colonized subject is for Fanon justified in engaging in "counterviolence," a type of violence that is equivalent to self-defense.[13] The colonial government can only come to power through violence, since people will not willingly part with their land, which provides "bread and, naturally, dignity," without fair compensation.[14] Then, once the colonial government seizes power, the colonizer sees plainly that "there is not one colonized subject who at least once a day does not dream of taking the place of the colonist."[15] Furthermore, unlike less segregated capitalist societies—in which the proletariat can theoretically become part of the bourgeoisie

---

[11] Kant, trans. Reiss, 58.
[12] Kant, 58.
[13] Fanon, 46.
[14] Fanon, 9.
[15] Fanon, 5.

through means that society deems legitimate—as subjects of a colonial regime, the colonized cannot become part of the colonizer class. Colonizers are rich because they are white, and are white because they are rich.[16] This inability of the colonized to integrate into colonial society further contributes to the culture of violence in a colonial society, since the colonizers know that the colonized want to take their place, and can only do so through massive and non-gradual social change, which entails the use of violence. Therefore, the colonizer seeks to create a state so repressive and efficient in its application of violence, that to challenge it with violence in turn is perceived by the colonized subject to be suicidal. Since the colonial regime arranges society in such a way that large-scale social change cannot occur peacefully, both the colonized subject and the colonizer understands that if the colonized subject were ever to act on their desire for liberation, such action would be violence.

The Futility of the Public Use of Reason in Colonial Societies

Although the colonizer may permit the intelligentsia among their colonized subjects to form and participate in political parties, and to attempt to enact peaceful change through their public use of reason, this will only solidify the colonizer's control, and will allow them to "wage a rearguard campaign in the fields of culture, values, and technology, etc."[17] That is, the colonial regime purports to permit the public use of reason in the context which Kant finds appropriate—after having done one's obligatory duties to the state—but in fact the colonial regime ensures that this public use of reason is ignored, heard but not listened to, and so is therefore rendered illusory. The colonizer uses the limited participation of the colonial intelligentsia to create the illusion that peaceful change is a viable alternative to violence, and thus the colonial regime can placate its colonized subjects, and maintain control. No real change, such as the regaining of land, power and dignity by colonized subjects, can occur through the colonized intelligentsia's "public use" of reason, since accommodating the desires of the colonized subject is inherently threatening to the continued existence of the colonial state.[18]

Therefore, any attempt by colonized subjects to engage in public use of reason in Kant's approved-of context will be doomed to fail, because the colonial state will never engage with

---

[16] Fanon, 5.
[17] Fanon, 9.
[18] Kant, 55.

them in good faith. Thus, the colonial government "is not a machine capable of thinking, a body endowed with reason. It is naked violence and only gives in when confronted with greater violence."[19] The colonial government cannot allow the colonized subject to effectively engage in the political process, since their demands for the return of their land and political power are incompatible with the continued existence of the colonial state. Since the colonial state only "speaks the language of violence," the colonized have no other choice but to interact with the state in the same way.[20] Therefore, violent acts committed by the colonized are to be regarded not as violence per se but as counterviolence.

Conclusion

Therefore, Kant and Fanon seem to disagree about how one should act once one leaves one's state of immaturity. Kant was attempting to write a universal text, and presupposed that maintaining social order was always the correct course. Kant "insistently abstracts ... issues in order to see them in their widest sense," which prevents him from addressing the specific issues facing certain societies.[21] However, Fanon was writing specifically about colonial societies, and began his analysis by arguing that these societies were illegitimate and resulted in such catastrophic harm to their victims that chaos is preferable to the continued existence of an oppressive colonial regime. Kant believed that once people become enlightened, they should only engage in the public use of reason in the appropriate context, otherwise obeying the state and adhering to the private use of reason. But Fanon believed that, in colonial societies, once people become enlightened, they are justified in engaging in disobedience in situations in which the colonial regime demands that they relegate themselves to the private use of reason, in order to reclaim their land and dignity: for colonial regimes are illegitimate, and cannot rightly demand the obedience of the colonized subject.

Part II: Agreement Between Kant and Fanon

Before continuing it will be necessary to define two Kantian concepts: that of the guardian and of the state of immaturity.

---

[19] Fanon, 23.

[20] Jean-Paul Sartre, preface to *Wretched of the Earth*, by Frantz Fanon (New York, NY: Grove Press, 2021), l.

[21] Geoffrey Galt Harpham, "So... What 'is' Enlightenment?," 594.

The Concept of Guardians

"Guardians" are the elites within a society who dictate what it is that other people in the society should think. For example, a guardian might be the "pastor who has a conscience for me, a doctor who judges my diet, and so forth."[22] Once the guardians establish control over their society, they then strive to keep the people whose thoughts they direct in a state of immaturity.

The Concept of the State of Immaturity

People are in a state of immaturity when they allow guardians to think for them. Furthermore, this state is often "self-incurred."[23] For example, it is difficult to "make use of" one's "own understanding" when doing so might set you against the prevailing beliefs of society.[24] However, once people reject the guardians, thus leaving their state of immaturity, they become enlightened and are able to make their own decisions. That is, one becomes enlightened only insofar as one thinks for oneself—so long as one is subordinate to a class of guardians, one cannot be considered enlightened.

Kant and Fanon on Enlightenment

Although *The Wretched of the Earth* seems to come to a different conclusion than *What is Enlightenment*, both Fanon and Kant begin their analysis from a similar, although not identical, idea: people should not allow guardians to think for them. Where they differ is that Kant believes there to be a binary opposition: people are either in a state of immaturity, where they are unable to "make use of" their "own understanding without guidance from another," or they are enlightened, in that they think for themselves.[25] However, Fanon rejects this binary and argues that there are other ways of thinking, such as communal thinking, where people express the views of their community and ancestors.[26] Therefore, for the purposes of this paper, I will use "state of immaturity" to mean "deferring to a guardian's way of thinking," instead of in the sense of "failing to make use of one's own understanding without direction from another." In the following paragraphs, I will justify both my claim that the concept of "guardians" can be

---

[22] Kant, trans. Schmidt, 58.
[23] Kant, 58.
[24] Kant, 58.
[25] Kant, 58.
[26] Fanon, 33.

intelligibly applied to Fanon's thinking, and that Fanon can be shown to agree with Kant that people should grow out of their state of immaturity.

Guardians in Fanon's Thinking

If we apply Kant's idea of guardians to Fanon's thinking, we can see that Fanon describes the colonial regime in a way that is similar to the way in which Kant describes guardians: both the class of guardians and the colonial regime are comprised of social elites who work to keep their subjects in a state of immaturity. As Fanon notes, "in colonies, the official, legitimate agent, the spokesperson for the colonizer and the regime of oppression, is the police officer or the soldier."[27] These colonial "guardians" strive to keep colonized subjects in a state of immaturity, so that they do not challenge the colonial system. However, unlike in the European societies upon which Kant based his theory, populated by citizens, in colonial societies which rule over colonized subjects, guardians do not simply use their powers of persuasion to make the people "stupid."[28] Instead, the colonial guardians are the police officers or the soldiers who use violence in an attempt to keep the colonized in a state of immaturity. This, along with the colonizer's claim that the colonized are sub-humans in need of civilizing, reifies the colonized into "things," who are thus almost completely excluded from political life. However, such brutal and dehumanizing treatment can also encourage colonized subjects to reject their guardians.

Unlike in Kant's society, where he argued that "it is so easy to be immature!," living in a state of immaturity for the colonized subject results in great suffering, both materially and spiritually, and creates a tension within society that must be resolved.[29] While the colonizer lives in a world of "stone and steel," and "lights and paved roads,"[30] the colonized subject lives in a "world with no space," where they are "hungry for bread, meat, shoes, coal, and light."[31] The colonial subject is dehumanized, labeled as animalistic, and denied social recognition. This suffering can uniquely position the colonized subject to reject their guardians, since they are extremely motivated to change the system that causes their suffering. As Fanon puts it, "the last shall be first [to revolt], the first last," meaning that the least privileged people in a society will

---

[27] Fanon, 3.
[28] Kant, 58.
[29] Kant, 58.
[30] Fanon, 4.
[31] Fanon, 4.

be the most motivated to revolt, and the most privileged people will be the least motivated to revolt.[32] Even if "the colonized subject begins to doze off or forget" the reason for their suffering, "the colonist's arrogance and preoccupation with testing the solidity of the colonial system will remind him on so many occasions that the great showdown cannot be postponed indefinitely."[33] In Fanon's time, this showdown was beginning, and this showdown, it seems to me, followed a pattern very similar to that of Kant's theory of enlightenment.

Enlightenment in Fanon and Kant: Similarities

Fanon notes that entire groups of people in countries such as Algeria were, in his time, becoming empowered to engage in their own ways of thinking and thereby to participate in decolonial action, which inspired other groups of people to similarly rise up against their colonial oppressors. This idea of Fanon's is very similar to Kant's idea that certain individuals among the population would first become enlightened, thereby inspiring their peers to become enlightened as well. Therefore, in the "international and African context [of 1961], the poverty-stricken and independent population achieves a social consciousness at a rapidly accelerating pace."[34] However, this raises the question: "what would colonized people do with this new social consciousness?"

Private and Public Use of Reason After Enlightenment

This question led Fanon to argue that once colonized people begin to leave their state of immaturity, they will rationalize the use of violence against the colonial regime, acknowledging the illegitimacy of the regime and acting upon their reason when the regime demands obedience. It is this violence which empowers the colonized subject to fully leave their state of immaturity. However, the colonial system is designed to keep the colonized subject "in their place," and is always ready to inflict violence upon them should they attempt to disobey the regime. Thus, since the colonized people have an inherent and immediate need to change the system that oppresses them, but have no peaceful way to do this in a system designed to keep them "in their place," "the dreams of the colonial subject" have to be "muscular dreams," meaning the

---

[32] Fanon, 2.
[33] Fanon, 16.
[34] Fanon, 12.

colonized people have a subconscious need for violent action.[35] This subconscious aggression results in the colonized subject existing in a state of "muscular tension" that "periodically erupts into bloody fighting between tribes, clans, and individuals."[36] The colonizer eagerly uses this violence to justify their claim that the colonized are savage and animalistic, in desperate need of the "civilizing" guidance of the guardians, reinforcing the dehumanization and lack of recognition of the colonized subject. The colonized subject retreats into the realm of myth, which allows them to imagine the colonizer's power over them as being diminished in strength when contrasted with supernatural horrors, rendering their state of immaturity slightly more comfortable. This state of "muscular tension" can only begin to change once the colonized subject ceases their "'head-in-the-sand behavior,'" and emerges from their state of immaturity into a state of enlightenment, redirecting their aggression towards the colonizer.[37]

When the colonizer detects this emergence of the colonized subject out of the state of immaturity, the instinctual reaction of the colonial regime is to respond with violence. However this "repression intensifies the progress made by the national consciousness," accelerating the willingness of the colonized to use violence against the colonizer.[38] Then, "by regaining control over his body," through properly directed violence, "the colonized subject becomes aware of his own position in colonized society," and feels authorized to make changes to said society.[39] Through this process of regaining control over themselves, the colonized have violent encounters with the colonizer as "an other," fusing the "body and mind" of the colonized subject.[40] For Ty, Fanon shows that "such a deployment of force not only reverses the medusan hold of colonization so as to create new, disalienated subjects but transforms violence itself—from an affliction to a mode of self-making."[41] Therefore, the colonized subject realizes that they have an agency of their own and that the colonist is an equal of theirs; they are no longer "things" but by thinking for themselves become people. However, when the colonial regime realizes that the colonized are becoming so empowered, they assign "a series of good souls" to the colonized

---

[35] Fanon, 15.
[36] Fanon, 17.
[37] Fanon, 17.
[38] Fanon, 32.
[39] Stefan Bird-Pollan, "Fanon," *Critical Horizons* 13, no. 3 (2012): 377–99, https://doi.org/https://doi.org/10.1558/crit.v13i3.377, 394.
[40] Bird-Pollan, 393.
[41] M. Ty, "The Resistance to Receptivity: Or, Spontaneity from Fanon to Kant," *Cultural Critique* 113, no. 113 (2021): 1–27. https://doi.org/https://doi.org/10.1353/cul.2021.a803752, 1.

"who in the 'Symposiums on Culture' spell out the specificity and richness of Western values," in a desperate attempt to return them to their state of immaturity.[42] It is then the responsibility of the colonized to keep themselves on the path of empowerment by rejecting these guardians.

Therefore, it seems that both Kant and Fanon in fact agree that people should "have the courage" to make use of their own understanding and that they only disagree on the form that this understanding should take, since Kant imagined that the attainment of enlightenment should be accomplished individually whereas Fanon thought that, for the colonized, enlightenment can only be achieved by the community as a whole.[43]

Part III, Kant's Just War Theory Applied to Decolonial Revolutions

Reconciliation of Kant and Fanon's Arguments

Although it may seem at first that Fanon's argument that the use of violence by the colonized subject against the illegitimate rule of a colonial regime is justified in order to achieve decolonization and Kant's argument that people should only make public use of their reason after fulfilling one's obligations to a legitimate state are incompatible, I believe that they can be reconciled. Kant was attempting to write a universal text, but seemingly assumed "that European existence is, properly speaking, true human existence per se."[44] In this section I will first show that Fanon's argument, which was based on a sociological and psychological examination of particular societies, when contrasted with Kant's "universalist" theory, shows that while Kant's theory holds true for the European societies he was familiar with, it does not apply to societies such as those Fanon described. But this does not mean that Kant must be abandoned in favour of Fanon—rather I will show that a reading of Kant's theory of just war can accommodate and incorporate Fanon's account of colonial society with Kant's universalist ambitions. Kant's theories of just war and legitimate peace suggest that for a society which is established through an unjust war and which is never rendered legitimate by a just peace might be rightly made the subject of the type of revolution which Fanon advocated for. On such a reading of Kant, a

---

[42] Fanon, 8.
[43] Kant, 58.
[44] Tsenay Serequeberhan, "The Critique of Eurocentrism and the Practice of African Philosophy," in *Postcolonial African Philosophy: A Critical Reader*, ed. Emmanuel Chukwudi Eze (Cambridge, MA: Blackwell, 1997), 144.

revolution in this context might be seen to be continuous with the waging of a just war, such that the conduct of such a revolution might be sanctioned by Kant's just war theory.

Kant's Just War Theory

Kant believed that there were certain situations in which violence was justified. In *Kant's Ethics of War and Peace*, Orend argues that Kant held a "just war theory," meaning that Kant takes the view that if certain conditions are met, such as "rights-violating aggression" which cannot be defeated except through violence, then a violent response by those whose rights are being violated is justified.[45] This view is compatible with Kant's deontological thinking in general, since we can universalize the following maxim in the spirit of the categorical imperative: "when faced with rights-violating aggression, I am always morally justified, but not required, to defend myself and my nation, including by using armed force, necessary for self-defense."[46] This view is also compatible with Kant's assertion that, within a state, violence ("coercion") is justifiable when it is "a hindrance to a hindrance of freedom."[47] Kant further states that anything that is "contrary to right is a hindrance to freedom," and therefore it is right to resist such a hindrance.[48] That is, it is right to apply violent means in order to resist those who have applied violent means with the aim of restricting freedom. Concerning violence between states, Kant also reserves the right to self-defense—citizens may rightly "undertake voluntary military training from time to time" in order to defend themselves and their state from aggressors."[49] Therefore I concur with Orend insofar as "Kant is a defender of a version of the just war theory."[50]

---

[45] Brian Orend, "Kant's Ethics of War and Peace," *Journal of Military Ethics* 3, no. 2 (2004): 161–77, https://doi.org/https://doi.org/10.1080/15027570410006507, 168.

[46] Deven Anand, "Kantian Just War Theory," 28.

[47] Immanuel Kant, "The Metaphysics of Morals," in *Kant: Political Writings*, ed. Hans Siegbert Reiss (Cambridge: Cambridge University Press, 1991), 134.

[48] Kant, "The Metaphysics of Morals," 134.

[49] Immanuel Kant, "Perpetual Peace: A Philosophical Sketch," in *Kant: Political Writings*, ed. Hans Siegbert Reiss (Cambridge: Cambridge University Press, 1991), 95.

[50] Andrew Fiala, "Pacifism," Stanford Encyclopedia of Philosophy, May 8, 2023, https://plato.stanford.edu/entries/pacifism/#DeonPaci, para 2.

The Continuation of a Kantian Just War After a "Peace"

If it is justifiable to defend one's country with violence in a just war, then I see no reason why it would be unjustifiable to continue that violent struggle once part, or all, of said country is occupied by the invader, and then falls under the rule of the illegitimate government of the invader. I believe this view is supported by Kant. Kant argues that a "suspension of hostilities" which occurs when "both parties may simply be too exhausted to continue the war," is not seen to be a legitimate peace, since the underlying issues which caused the conflict are not seen to be resolved.[51] Kant's theory of just war, and what constitutes a legitimate peace, explains why it seems reasonable, for example, that the French people did not lose the right to resist the Nazis once the swastika was raised in Paris and the Vichy and then Nazi regimes established governments—such governments are illegitimate. Likewise, this theory can be applied to the colonial states which Fanon was concerned with.

Kant's Theory of Just War Applied to Colonial States

If we then apply our maxim "when faced with rights-violating aggression, I am always morally justified, but not required, to defend myself and my nation, including by using armed force, necessary for self-defense" to the European colonization of Africa, I believe that the colonized peoples of Africa are seen to be justified in resisting their colonizers.[52] National defense is a *jus ad bellum*, a just cause for war. Kant argues that "'no state shall forcibly interfere in the constitution and government of another state,'"[53] and defines a state as "a union of an aggregate of men under rightful laws."[54] Although undertaking an anthropological examination of every colonized people to determine if they fit Kant's criteria of a state is beyond the scope of this paper, I believe that it is reasonable to assume that some of them will be found to meet Kant's criteria for the conduct of a just war. Therefore, since colonization amounts to a state forcibly interfering with the constitution or government of another state, thereby violating the rights of the people of that state, it seems to me that Kant's just war theory permits colonized

---

[51] Kant, "Perpetual Peace: A Philosophical Sketch," 93-94.
[52] Anand, 28.
[53] Kant, "Perpetual Peace: A Philosophical Sketch," 96.
[54] Kant, "The Metaphysics of Morals," 138.

peoples to defend themselves and their nation "by using armed force, necessary for self-defense."[55]

Kant's Theory of Peace Applied to Colonial States

Since the underlying issues of these wars—land and dignity—were never resolved, the "peace" that takes hold once the colonizer establishes a government is merely a temporary "suspension of hostilities,"[56] and the establishment of an illegitimate government. Therefore, since Kant's just war theory, as I have depicted it here, permits colonized peoples to engage in a just war against their colonizers, I believe that violent decolonization can be rightly described not as potentially illegitimate revolutionary violence, but rather as the ongoing conduct of a just war of the sort sanctioned by Kant's theory of just war.[57]

Conclusion

In conclusion, I believe that while Kant's *What is Enlightenment* and Fanon's *Wretched of the Earth* initially seem to be irreconcilable, they can actually be synthesized into one coherent theory concerning how it is that people should behave towards their government once they have left their state of immaturity. Both Kant and Fanon agree that people should not remain in a state of immaturity, and should attain enlightenment by thinking for themselves rather than deferring to the decisions of guardians. In societies with legitimate governments, however, Kant's point holds that people have a duty to only engage in the public use of their reason after fulfilling their obligations to the state, during which they agree to practice the private use of their reason. However, in societies where the government has been established through an unjust war, such as the colonial societies with which Fanon is concerned, people are justified in disobeying the state in the most extreme form possible: revolution.[58] In such cases of illegitimate government, Fanon

---

[55] Anand, 28.

[56] Kant, "Perpetual Peace: A Philosophical Sketch," 93.

[57] Kant writes in a note that "any *legal* constitution, even if it is only in small measure *lawful*, is better than none at all, and the fate of a premature reform would be anarchy." (Kant, "Perpetual Peace: A Philosophical Sketch," 96). However this does not address the challenging a foreign invader, which I have argued is precisely what decolonization is. Furthermore, since Kant justifies submitting to an illegitimate constitution on consequentialist grounds, I believe that Fanon's argument that the continued existence of an oppressive colonial state results in greater harm to the colonized than the "total disorder" of decolonization, can be made to fit within Kant's framework. Fanon, 2.

[58] Or Europe, which "Nazism transformed ... into a genuine colony." Fanon, 57.

and Kant can be seen to agree that the use of violence is permitted in order that an illegitimate government can be overthrown and replaced by a legitimate government, and that such a violent process as this can be defended as the continuation of a just war.

Bibliography

Anand, Deven. "Kantian Just War Theory." Senior Thesis, Trinity College, 2021. https://digitalrepository.trincoll.edu/theses/934.

Bird-Pollan, Stefan. "Fanon and the Critical Ideals of German Idealism." *Critical Horizons* 13, no. 3 (2012): 377–99. https://doi.org/10.1558/crit.v13i3.377.

Byrd, B. Sharon, and Joachim Hruschka. *Kant's Doctrine of Right*. Cambridge, NY: Cambridge University Press, 2010.

Fanon, Frantz. *The Wretched of the Earth*. New York, NY: Grove Press, 2021.

Fiala, Andrew. "Pacifism." Stanford Encyclopedia of Philosophy. Stanford University, September 15, 2018. https://plato.stanford.edu/entries/pacifism/#DeonPaci.

Harpham, Geoffrey Galt. "So... What 'is' Enlightenment? An Inquisition into Modernity." *Critical Inquiry* 20, no. 3 (1994): 524-556. http://www.jstor.org/stable/1343868.

Jacobsen, Mogens Chrom. "Kant and the Modern State System." *Jahrbuch Für Recht Und Ethik / Annual Review of Law and Ethics* 17 (2009): 85–107. http://www.jstor.org/stable/43593971.

Kainz, Howard. "Is 'Just War' Theory Justifiable?" *Journal of Social Philosophy* 27, no. 2 (1996): 158–67. https://doi.org/10.1111/j.1467-9833.1996.tb00243.x.

Kant, Immanuel. "An Answer to the Question: 'What Is Enlightenment?'" In *Kant: Political Writings*, edited by Hans Siegbert Reiss, 54–60. Cambridge: Cambridge University Press, 1991.

Kant, Immanuel. "An Answer to the Question: What Is Enlightenment?" In *What Is Enlightenment?: Eighteenth-Century Answers and Twentieth-Century Questions*, edited by James Schmidt, 58–64. Vol. 7 of *Philosophical Traditions*. Berkeley: University of California Press, 2004.

Kant, Immanuel. "Perpetual Peace: A Philosophical Sketch." In *Kant: Political Writings*, edited by Hans Siegbert Reiss, 93–130. Cambridge: Cambridge University Press, 1991.

Kant, Immanuel. "The Metaphysics of Morals." In *Kant: Political Writings*, edited by Hans Siegbert Reiss, 131–75. Cambridge: Cambridge University Press, 1991.

Masek, Lawrence. "All's Not Fair in War: How Kant's just War Theory Refutes War Realism." *Public Affairs Quarterly* 16, no. 2 (2002): 143-154.

Orend, Brian. "Kant's Ethics of War and Peace." *Journal of Military Ethics* 3, no. 2 (2004): 161–77. https://doi.org/10.1080/15027570410006507.

Sartre, Jean-Paul. Preface to *Wretched of the Earth*, by Frantz Fanon, xliii-lxii. New York: Grove Press, 2021.

Serequeberhan, Tsenay. "The Critique of Eurocentrism and the Practice of African Philosophy." In *Postcolonial African Philosophy: A Critical Reader*, edited by Emmanuel Chukwudi Eze, 141–61. Cambridge, MA: Blackwell, 1997.

Starke, Steven Charles. "Kant's Just War Theory." PhD diss., University of South Florida, 2016. https://digitalcommons.usf.edu/etd/6398.

Ty, M. "The Resistance to Receptivity: Or, Spontaneity from Fanon to Kant." *Cultural Critique* 113 (2021): 1–27. https://doi.org/10.1353/cul.2021.a803752.

Augenblick and Nietzsche's Eternal Return: Movements Towards a New Reading

Dax Hamouth

Introduction

Quite a lot of philosophical reflection has been directed towards the temporal aspect of Nietzsche's Eternal Recurrence or Return[59]; and though this is by no means misplaced, I cannot help but fear in the pursuit of the future and the past, there had been an un-earthful neglect of the present, of the 'Moment' (German: *Augenblick*, lit. glance/blink of an eye)—and Zarathustra's wisdom of is the wisdom of the earth[60]. Nietzsche puts an extremely strong emphasis on Augenblick, both in the language and playing out of his introduction to Eternal Recurrence, "This long lane back here: it goes on for an eternity. And that long lane out there—that is another eternity…this gateway is where they come together. The name of the gateway is inscribed above it: 'Moment'"[61]. It is so important that I suggest that Nietzsche's implementation of Augenblick is an essential consideration to understand Eternal Recurrence; and that a focus on Augenblick allows for new understandings of Eternal Return to be generated.

The Eternal Return has been interpreted and over interpreted. Readings vary from a kind of cosmological metaphysics of time, to an ontology of philosophical selection that overcomes the limits and will-to-nothingness of the dialectical form[62]. Heidegger famously accuses Nietzsche (*qua* return of the same) as inverting but not overcoming metaphysics—insofar as Heidegger understands Return in connection with will to power as an attempt to understand the Being of being[63]. These assumptions lie however in the false assumption that the Eternal Return must be a kind of *metaphysica* or otherwise have a direct reality bearing onto-status. Rather, there is a more thorough investigation to be done vis the understand of Eternal Return as a kind

---

[59] 'Return' and 'Recurrence' are used entirely interchangeably in this essay.
[60] Eg, Nietzsche Z.I.3
[61] Nietzsche Z.III.2
[62] This is the quintessential Deleuzian reading. See Deleuze's *Nietzsche & Philosophy*; as well as *Difference and Repetition.*
[63] Heidegger 164;see more generally Heidgger's *Nietzsche Vol. II: The Eternal Recurrence of the Same and the Will to Power.*

of existential myth; a self-making for taking control of the becoming of self and the rooting of identity not in the looping willing of becoming, such that the essence of Being is that of becoming; but instead that the process of becoming is itself Being. For his part, the understanding of the Eternal Return in the poetic sense is directly opposed by Heidegger as "[dragging] the thinker down to the flatland of current opinion"[64]. I disagree with this. At the very least, we must acknowledge and never forget that no small part of what elevates Nietzsche is his near-infinite interpretability; there are only degrees and ranks of truth, and thus if we are to honour his own philosophy, multiple readings of everything must be put up for proper competition.

In light of the Augenblick of gateway, the return then is not just the matter of a life to be relieved, it is also the proclamation of the eternal moment up and against eternity; that is, the return is the rebellion of the moment of becoming against Platonic eternities[65]. It is a myth about the lived reality of the 'self' that exists necessarily in the Moment of Augenblick–; a present now into which past and future are blended to form a melange. Thus, Nietzsche's Augenblick in the Eternal Return finds in itself the convalescence or knotting together of time and a life, a Nietzschean rejection of the duality of being and time. This is all to say that what I suspect Nietzsche is really trying to accomplish with the concept of Eternal Recurrence is to trigger a moment of ἔκστᾰσῐς, ecstasy, in which the individual can stand outside their self and observe their 'identity' as merged with time in the present moment; this ecstasy Nietzsche seeks to trigger, for the aim of therapeutic and spiritualization of the self, in line with his desire to have us affirm Life.

To begin, I will attempt to sketch out the character of Augenblick in the Eternal Return. In broader literature, Augenblick refers to the motif of a philosophical Moment of the present-now in which there is the ability for transformation or action[66]. In Nietzsche, Augenblick appears in *Thus Spoke Zarathustra* as the gateway of Moment, of Augenblick, which is the moment of present-now between the eternities of the Eternal Recurrence. Identity is enraptured into Augenblick by the subjective experience of it: it contains both the selfhood of the person, say Zarathustra, and all the time of his life. The Augenblick as possessing the highest level of

---

[64] Ibid 166.

[65] Nietzsche seems to indicate such in section III.13.2, as I will discuss later.

[66] See Koral Ward's *Augenblick: The Concept of the 'Decisive Moment' in 19th- and 20th-century Western Philosophy*

reality of any discrete division of time, supersaturates our experience of being; or rather, being exists in the immanency of Augenblick. This is the way in which the moment 'draws after it all things that are to come'[67], namely that Augenblick contains in it the intersection and totality of a life. In this way Eternal Return is in part a movement away from the Platonic dichotomy of being and time as separate. Augenblick *qua* Eternal Return, is furthermore characterized by a self-recursion in which, conceptually, the self is able to watch the self-watching the self, ad infinitum; and this self-watching, I will argue, is a phenomenon which causes not just a return to eternal recursion, but a recursive-self-watching where something is added and changed to the self. It should be noted here this recursiveness does presuppose a mandatory return to Augenblick over a return to 'history', which is to say on this particular point I will be reading with Deleuze, who I will attempt to draw insight from and critique. I will also engage with Catherine Malabou's own critique of Deleuze and Derrida to consider the ways in which the Eternal Return attempts to surmount a biographical understanding of self in favour of an existentialism of the present.

The function of Eternal Return then, is a therapeutic myth to bring on this ecstatic state of self-watching and merger of being and time, as both Lawrence Hatab and Robin Small suggests, calling Eternal Recurrence a "tragic-myth-poetic concept"[68]: a thought-experiment that, in its moment of consideration, before we judge it to be true or false, has the same power as if it was true[69]. Importantly, being a myth does not render it untrue; instead it exists as a truth in the mythic sphere, which though it may or may not overlap with factuality, and has no less importance in its operative means: it is true in its proper mythic context[70] and it is a *perspective* on truth which can be useful and as important as factuality.

---

[67] Nietzsche Z.III.2
[68] Hatab 98: "…I want to argue that eternal recurrence can be understood as a tragic-myth-poetic concept, a formation meant to engender a "virtual reality"—a literal immediate disclosure that is yet not constructed as a cosmological fact."
See also Hatab 99: "Eternal recurrence can be taken as a conceptual myth of life affirmation intended to operate with immediate (literal) disclosive force".
And more generally, see Hatab "Chapter 5: Making Belief" for Eternal Recurrence as myth-making.
[69] Smalls 38
[70] Hatab 95: "Mimetic literality can be called "make believe" in a positive sense, as making belief in the milieu of poetic performance and reception".

Character of Augenblick: the Eternal Present

> "This long lane back here: it goes on for an eternity. And that long lane out there—that is another eternity. They contradict themselves, these ways; they confront one another head on, and here, at this gateway is where they come together. The name of the gateway is inscribed above it: 'Moment'."[71]

Thus spoke Zarathustra in "On the Vision and Riddle". One can consider the past and future to be important affects, but the reality is that all of time is contained in the present now, and so the gateway called Augenblick[72] enraptures all of time into itself; the state of all things is raptured into the present-moment, which becomes the moment in which we can observe a semi-static stillness of becoming existence. Augenblick means 'blink or glance of an eye', and its quality of eternal present finds part of itself in observation. Robin Smalls identifies the Moment as an agonistic struggle between the past's 'I was' and future's 'I will be', in which the past always threatens to take over the future because the moment passes away into it[73]. This however does not consider that the past is contained in the present and exists only as the quality of memory and things-that-once-were. This is to say, the present passes away into itself; Augenblick contains both the baggage of itself (past) and the possibilities made to itself (future). The present itself is ultimately characterized by the experience of the past and future simultaneously. One lives at the nexus between that which has happened and that which could come to pass, and that nexus is the present in which the 'happened' part conflates with the openness of the future 'could/will happen'. These three registers exist together in the sense a physical spatial manifold—whereby the events of the past have brought about the layout of the present, which itself will inevitably collapse and pass into the future set of spatial contingencies. Smalls' identification of this agonistic struggle between poles of temporality is apt; however, the Augenblick is not just the resulting tension between past-future, it is also the concept of the container holding them, their battleground; the past and future run contest, but the contest can only take place in the present-now of Augenblick. Augenblick therefore exists as both the resultant of the struggle, but also exists 'before/outside it'; every moment passes away into itself, and in the 'next' moment past-future fight on its corpse—'next' to be taken lightly, as there is

---

[71] Nietzsche Z.III.2
[72] Smalls 96-97; Ward 43
[73] Smalls 98-100

only one moment-now. Augenblick can be confusing, because it is the stillness-which-moves, or the present-instant-where-things-still-unfold time and untime together. Standing at the edge of the Moment we can observe the effects of temporality's contest. The gateway of Augenblick then is the subject-watching of the past and future in the present.

Deleuze in his *Nietzsche and Philosophy*, says "the moment must be simultaneously present and past present and yet to come, in order for it to pass"[74]. Here we can see traces of the influence of Bergson and Spinoza on Deleuze. To unpack Deleuze, he means that for time to pass, there must be a simultaneous co-existence of the three registers, otherwise time would function on the necessity of the existence of a singular discrete framing of the present. Future would never arrive and past would never be, for a new present could never come to replace the current present to allow for passage. To quote, "If the present did not pass of its own accord, if it had to wait for a new present in order to become past, the past in general would never be constituted in time…We cannot wait, the moment must be simultaneously present and past, present and yet come…"[75]: a new present cannot come to replace the current present, as the present constitutes itself and nowness in-itself, such that the new present would be identical to the current present, such that no new present could exist. Both 'new' and 'current' are in actuality negated by the term 'present', for all presents are the one present of present-now. Deleuze uses this point then to criticizes cosmological-interpretation and mechanical readings of Eternal Recurrence, showing that the cosmological- reading implies necessarily a final end-state in which becoming is ended before a 'reboot', and consequently that an end to becoming is incompatible with Nietzsche[76]. The Eternal Return for Deleuze then is an answer to Plato—and later, to Hegel[77]. Deleuze's references Plato as saying "if everything that becomes can never avoid the present then as soon as it is there, it ceases to become and is then what it was in the process of becoming"[78]. Plato here holds the opposite view, namely that becoming is rendered into being by the nature of a universal singular present which is supplanted by a new present—i.e., time as snapshots of existence. Deleuze's echoing of Bergson in his interpretation of Nietzsche is meant to frame Eternal Return as a solution to the Platonic problemata of

---

[74] Deleuze 48-49
[75] Ibid
[76] Ibid
[77] See 147-189; see also Deleuze's *Difference and Repetition*; as well as Deleuze and Guattari's duology of *Capitalism and Schizophrenia*
[78] Ibid, 47

becoming; Nietzsche—and Deleuze—want to preserve true becoming in their thought by interpreting time as a process rather than snapshots.

It remains a question whether Deleuze is entirely correct in his reading of Nietzsche, as he is known for creative misreading; however, the above is an important point regardless of Deleuze's' faithfulness to Nietzsche[79]: that if Nietzsche's eternal becoming is most important, Eternal Return must always return or recur into Augenblick, and moreover, the Return constitutes an infinite series of Augenblicks unfolding and co-existing—but also because of the inherent immanence of Augenblick, only one Augenblick. Deleuze finds 'return' to be the answer to the problem of the passage by means of a self-referential state to becoming that allows for the present to pass[80]. This is the insight that I would like to highlight: that Eternal Recurrence is a return-of-moment, which discretely cannot deal with other moments, in one instance; or in other words, the Eternal Return can only ever take into itself a particular Augenblick in a particular instance of its invocation, because the recurrence does not entail a recurrence of all of history—for this may platonically affix 'History' in a way possibly unacceptable to Nietzsche[81]—but rather that it forms a fold back into the becoming of the present-now, which is Augenblick. The important contradiction however is that Augenblick is both the plural and singular: each moment is an especial instance of it and entirely separate from each, but simultaneous every Augenblick is the one Augenblick, as there can only ever be the one moment-now—consisting of the co-existing three registers of past, present, and future—held going forward in becoming forever. So, observing Augenblick is to watch this moment, but also every moment. Augenblick, when under observation by a second-order-self-awareness of a self-contained within it, then enfolds all past and future into present, but only in that exact instance the moment is glanced at.

Nietzsche's Augenblick is characterized by a self-recursive watching, as return is to the present-now as well as where Being 'starts'. One enters and looks upon the gateway of Augenblick, and then is able to see the two lanes of past and future[82]; this is to say that, when one enters the gate, they have left behind the normal passage of time and entered into a conceptual

---

[79] 'Regardless' since this must be true if we take Nietzsche as favouring eternal-becoming (eg, Z.III.12.8). The question is not 'if he favours eternal-becoming, can the E.R. return through all time', but rather 'does he favour *eternal* becoming in the first place'. This reading of Augenblick presupposes he does—though it does not settle the question at all.
[80] Ibid. See above.
[81] Ie, Deleuze's reading. See Deleuze 48-49.
[82] Z.III.2

space in which they observe themselves in the return to now, because of the nature of Eternal Recurrence recurring towards Augenblick. The eternality of this creates an infinity of depth in which the watcher watches themselves, watching themselves on the next or previous 'recurrence' or 'return'. It is to be thought of, by analogy, as if one were in a room with no doors or windows and only a camera on the wall behind them, and a screen in front, such that you are able to watch yourself watching yourself. This watching of the self is an engagement of a second order self-watching and judgment. It is an experience of Ancient Greek ἔκστᾰσῐς: literally 'to stand outside the self', which comes into English as 'ecstasy', and of which Nietzsche's favourite, Dionysus, is god. The watching of oneself creates a conceptual infinite recursion: 'you watch yourself, watching yourself, watching yourself, ad infinitum'. This return to Augenblick creates an infinite depth in the present-now—and Augenblick is a phenomenon-in-itself which contains this infinite depth of identity. It can be contrasted with the infinite duration of identity—i.e., afterlife—of Christianity, which also creates infinity of self. However, Nietzsche's Eternal Return rather than creating infinity of duration, which by consequence must extend out of earthly life, creates a moment of infinite depth of the self in the present, grounding one in the earth by means of Augenblick. Indeed, when Zarathustra first relates Eternal Recurrence, he is describing not a firsthand account of the Moment or his journey up the mountain, but instead relates it to sailors[83]. Zarathustra is returning to Augenblick in his vision, but what returning to Augenblick really means, is the act of watching or remembering yourself in Augenblick—which because of the return, becomes recursive. Zarathustra's vision is perhaps better understood as a memory than a dream—though because of the conceptual space of return-to-Augenblick, a metaphorical 'memory-of-future'. But the nature of Augenblick is that the vision is also Augenblick. Zarathustra returns to Augenblick in another Augenblick, and this can be drawn out further, eternally. Important here is that Zarathustra, despite being in a loop of recursive self-observation, is changed by this watching of the self; that is, in the return to Augenblick, the self is given over and dissolved into eternity, but then something is added or changed. It becomes a kind of existential reckoning of selfhood to engage the second-order self-awareness, to allow for an evaluation of self by the self. It is a way of taking a life into account without falling to the trap of will-to-death which other forms of autobiographical evaluation do; to take one's life in its totality as a story and to evaluate one's becoming along lines of biography necessarily implies one's own

---

[83] Z.III.2

death as it makes us consider our life from the perspective of the moment of death, of legacy. The Eternal Return however as a thought-methodology allows for the consideration of one's own becoming without involving the moment of death.

In these ways, Eternal Recurrence escapes the other-worldliness eternity of afterlife, and instead allows for life-affirmation and grounding in the earth itself; and thus, Augenblick is a moment of infinite depth that creates an infinite present to be appreciated, rather than an infinite future, which denies the living-now and earthly life. The Eternal Recurrence then finds its purpose in its ability to trigger the gateway of Augenblick by means of the intellect and passion, in the same way the Ancient Greeks would use intoxication to bring on ἔκστᾰσῐς. The question is left open however, if Eternal Return is one instance of an ἔκστᾰσῐς device, or is a type of form to describe the/all way(s) in which Becoming can be evaluated in-itself without reference to Being.

Existential Myth

If the Eternal Return is to be taken as a kind of existential myth of narrative, the question of utility must be explored—what does Nietzsche seek to accomplish with it and what was his insight at Sils-Maria? The essential proposal is that as a poem-myth, the Eternal Return seeks to provide an existential 'imaginary'[84] narrative of understanding the self without referent to biography. Biography for its part always engenders a kind of death; it is an invocation of an after—after my biography: after my death. The Eternal Return however, by rooting the self in the Augenblick avoids the logic of towards-death by involving the infinite depth of the present; but to expand on this we must examine Malabou's critique of Derrida and Deleuze.

Catherine Malabou wrote a critique of both Deleuze's and Derrida's understandings of Eternal Return. Her critique of Deleuze leaves something wanting, however her critique of Derrida is quite useful for our purposes here. In her reading of Derrida, it is the motif of anniversary and signature that becomes critiqued. Derrida's own contribution to Eternal Return is that the biography of the one who thinks the eternal return is inherently bound up with their return; that, "the thematic of differance is precisely introduced in the analysis of the duality of

---

[84] Imaginary in the sense of not bearing high ontological proximity to reality; imaginary in the same way that cultural functions are imaginary but still have bearing on life in situ: language is imaginary insofar as words provide some sort of referent to something outside of language, and only have full existent status when they are deployed by a user. This view generally holds more or less true almost regardless of what one takes language exactly to be—the exceptions are few and include only views which hold the sounds of language to be drawn out of some sort of reality or transcendent matrix.

the circles, the first of which can be called *ontological*, the second *autobiographical* [her italics]"[85], and so for Derrida *Augenblick* becomes the anniversary recalling and enacting the marriage of the two returns. Ontology thus becomes wedded to Autobiography in a particularly hetrosexual marriage; that is, Ontology takes on the name of Autobiography: "The philosopher from then on no longer speaks of life in general, but always of his life, in its name"[86] until death do them part(?). Certainly, even Derrida sees this, as he says the autobiographical name is fundamentally a "dead-man's"[87] name. The Eternal Return becomes a mediation, a selector, between Life and Death[88]. I question here if Nietzsche who presents opposed so harshly the logic of death, insofar as he insists on living in *this* world, for *this* world, not for the other world; and indeed, his passage "On Free Death"[89] which on the topic of dying correctly, itself seems to propose a conception of death rooted in life. In other words, to die correctly means to die at the right time and to will one's death. This should not be taken to mean suicide, or at least only suicide; rather Nietzsche here, I propose, is speaking of the way in which death should happen from the point of view of the living. To will death supposes the ability to will, which intern supposes that even grappling with death must be approached from a fundamentally lived perspective; to relinquish life at the right time does not mean necessarily to kill oneself, but rather it is dealing with a living for this world, and thus a timely exit from the stage when our part is due. It is a looking-forward or looking-out at death, which does not imply death so much as end-of-life. The logic of death is bound up with a looking at things from the point of view of stand-still; but to approach death as being something willed is to look at death from the point of view of becoming. Thus, even in his dealings with death Nietzsche maintains *joie de vivre* which is not reducible to the logic of death, even if it deals with death. In any event, for Derrida to say the eternal return produces a death—a spectre, as Malabou says[90]—is unfair to Nietzsche's philosophical project. Even the selection between Life or Death is alien to Nietzsche's thought, as we have seen, since in the end for him there is only Life, the Eternity Zarathustra dances with; to be a mediation between life and death, the Return would have to itself constitute an invocation

---

[85] Malabou 26
[86] Ibid
[87] Derrida in Malabou 26: 'The name, Derrida writes, "is always and a priori a dead man's name, a name of death. What returns to the name never returns to the living. Nothing ever comes back to the living."'
[88] Malabou 27
[89] Z.I.21
[90] Malabou 27

of the two. Nietzsche, I argue, only ever wishes to invoke life, and his invocation of death are themselves understood only from the unitary lens of living.

Moving on, Deleuze finds of the return that Eternal Recurrence's power lies in its ability to dissolve reactive forces, for the thought of an eternalized return transmutes the reactive force of nihilism into an active force and destruction of the weak by its willing. The return then functions as a groundless wheel which serves to test these forces in his equation of "willing=creating"[91]. Malabou is quite right when she criticizes Deleuze on the accounts that no where in Nietzsche implies that, one, negative dialectics do not also recur, two, that the selection is automatic, and three that Deleuze is "violent"[92] in his anti-Hegelianism. Certainly, it seems less likely in Nietzsche that the reactive forces themselves would not also return—even if with difference. The very statement of eternal recurrence seems to imply immediately the possibility of negative recurrence— "[italics are all mine]...every *pain* and every joy and every thought and *sigh* and everything unutterably *small* or great in your life will have to return to you"[93]: the small alongside the great comes back. Malabou proposes that the principle of selection being automatic also carries no textual evidence and in fact seems less likely considering that the Eternal Return can also destroy the one who entertains it. It is a process or tool which can be failed or used improperly. I will refrain from commenting on the last charge of a violent anti-Hegelianism, aside from the fact this prejudice, even if it is a prejudice, is perhaps more acceptable than Malabou gives it credit for, or at least worth consideration.

Where Malabou's critique begins to fall short, I would say, is the charge that "Difference is something that is imposed upon [Nietzsche]"[94]. Deleuze, I agree with Malabou, construes the Eternal Return incorrectly, because he approaches it as a kind of ontological selection principle. However, Deleuze's great insight is to see Nietzsche as a philosopher of difference, and it is too far to deny this part of his analysis in denying the more questionable aspects of Deleuze's Nietzsche. In *Difference and Repetition*, Deleuze has the insight that repetition always carries with it a kind of pre-identical difference. Reading Nietzsche through Deleuze would find that even if Deleuze is wrong as to how to approach the return, the fundamental quality of new difference in return must be preserved. Malabou does not offer sufficient grounds to eliminate

---

[91] Deleuze "II: Active and Reactive". See especially, 68-71.
[92] Malabou 25
[93] GS 341
[94] Malabou 25

the idea of philosophy of difference as a feature of Nietzsche's oeuvre. Even if one was to say nothing else returned with difference, at the absolute minimum in the Eternal Return *you* return with difference. Malabou's desire to move away from Nietzsche as a philosopher of difference, I think, is one which is either unnecessary or hasty. She criticizes Deleuze's lack of textual evidence, that there is no focus on 'difference' in Nietzsche, but does not proceed to offer substantial evidence contra Deleuze. Certainly, the term difference does not have to appear frequently to be able to identify it as a principle of Nietzsche, as Deleuze does. There is more to say on this point, however we must move on, as it is not our focus[95].

Returning then to Nietzsche, his Augenblick is expressly the moment in which being and time are infolded, and the identity is pressed into the skin of the world, such that it becomes both a part of Life but also synonymous with it; and so is both the plural and the singular. Augenblick-as-recurrence is a summoning of the primordial Dionysian figure, the generic all-human of whom every identity is a mask on[96] ; in the moment of Augenblick, the self is so far recursively watched, it is lost. Identity is dissolved as a face in the crowd: the self-dissolves, so only the Dionysiac-flux, the will, remains. This is to say that, in the moment of consideration of the Eternal Recurrence, we give ourselves over to non-identity insofar as the idea of us as defined by our posthumous considerations—i.e., of legacy—are rendered void, because of its conceptual infinite of present. Some, maybe all sometimes, foolishly define ourselves by referent to posthumous legacy, and in doing so invoke being untowards death as a corner of our identity[97]. Nietzsche seems to suggest this dissolving of posthumous-self in several passages, such as 30 of "On Old and New Tablets', in which it seems to be indicated that the will remains even when the self dissolves. Note Nietzsche's italics in, "O you, my Will! You turning of all need, *my own* necessity!"[98]. Here and in this passage, the will is presented as something that exists as a necessity, even as far as to be a 'necessity' of having a self, such that below or 'after-undressing-of' self, there is will, the Dionysian. This return to naked-will by means of the loss of identity however is transient in the moment of Augenblick; what is given is returned at the same time it is given: identity is handed over to Augenblick, but it is also always defined by it, living within it, and draws its instantiation out of it, so it is returned with the passing of the

---

[95] I am indebted here to my professor Willow Verkerk who discussed this paper with me, and whose ideas on this topic have certainly inspired mine.
[96] Deleuze X-XI; see also 6-8 & 10-12
[97] See also Adriana Caverreoro's *Relating Narratives* for a more in depth critique of the logic of death
[98] (Z.III.12.30)

moment. Human-selfhood as a quality emerges out of our past as the 'inheritances' we take from our lives and our milieu, and out of the future as promises we make and the current-direction of our becoming. Nonetheless, identity, that which gives us our own particularity, exists only in the present—since it is only the present that 'exists'. Will then is like a naked body that wears clothes of selfhood. Consequently, we are given an 'tailoring' or 'transmutation' back from Augenblick. This is the new understanding, inspiration, or affirmation of our selfhood and fate, which the second-order self-awareness learns by watching itself undress in the mirror.

Before a philosopher, Nietzsche was a philologist; and thus here at the end, it bears returning to the classics. In Book XIII of Homer's *Odyssey,* Odysseus returns home without knowing it[99]. However, shortly thereafter Athena, in the form of a young shepherd boy tells Odysseus that he has in fact arrived home[100]. In this we see Wisdom as a Shepherd tells Odysseus upon reaching an unfamiliar shore that he has in fact arrived home; and that the home Odysseus sacrificed for the glory of contest in the Trojan War, has now been returned altered and ready for Odysseus to reclaim. This rapture of νόστος (*nostos*; 'return home'), is itself a moment of Augenblick and opportune return. By allegory, the Eternal Return has us submit the self into the struggle of Augenblick, where we are forced to relive our heaviest traumas, struggles, and darkness, before it is returned to us by our Wisdom—i.e, the intellectual conscience—who watched the unfolding drama—the Dionyastic—transformed, with new being, new meaning, and new opportunity. Thus, we are confronted by our personal identity, the human condition, and the "River of Becoming"[101] that is the world, all enraptured and interdependently connected; we are shown to be *of* Life, *of* the earth.

Conclusion

Augenblick is then an essential aspect for reading the Eternal Return. There has been a large focus on the nature of the return as a return from a certain past to a certain future, where, what I propose, Nietzsche is really trying to do, is create a mythic tool, to overcome nihilism and affirm Life—for if Augenblick contains all of Life. Augenblick is an important motif to consider if we want to pull out the therapeutic and spiritual functions of Eternal Return. Eternal

---

[99] Homer XIII. 213-221
[100] ibid 252-285
[101] Z.II.12

Recurrence rather than being a return to *chronos* is a return to *kairos*. It is a return to the opportune moment of ecstatic Augenblick in which one's own identity is given over to a recursive watching and then returned differently such that we can overcome our traumas, just as Zarathustra is able to will compassion for the dog that terrified him as a child[102]. Thus, Nietzsche's Augenblick is the moment of transvaluation of values and the transvaluation of one's own life and Life. As a therapeutic myth contra nihilism, the Eternal Recurrence brings us to an opportunity of *amor fati* because every instance of joy and suffering are muddled together to form the idea of Life itself, presented without specific content, and only with supreme total form. Augenblick breaks down being and time—including personal being and personal time—in their massive plurality, just as from a distance many people at a train station muddle together to become only an ocean of bodies. What is left then to affirm is not a life, or even our life exactly—though it forms a piece—but rather the glory of raw existence, of all Life itself: *amor fati.* And so, Zarathustra's meaning of the earth, then, is that the human-being is inseparably a part of Life itself.

Bibliography

Nietzsche, Friedrich. *Thus Spoke Zarathustra*. Translated by Graham Parkes. Oxford, 2008.
Nietzsche, Friedrich. *Beyond Good and Evil.* Translated by Marion Faber. Oxford, 2008.
Nietzsche, Friedrich. *Genealogy of Morality.* Translated by Maudemarie Clark and Alan J.
         Swensen. Hackett Publishing, 1998.
Deleuze, Gilles. *Nietzsche and Philosophy*. Translated by Hugh Tomlinson. Athlone Press, 1983.
         This ed. printed by Columbia University Press.
Hatab, Lawrence J. *Nietzsche's Life Sentence: Coming to Terms with Eternal Recurrence*.
         Routledge, 2005.
Heidegger, Martin. *Nietzsche Vol. II: The Eternal Recurrence of the Same and the Will to Power*.
         Translated by David Farrell Krell. San Francisco: Harper, 1991.
Homer. *The Odyssey*. Trans. Robert Fagles. Penguin Classics, 1996.
Malabou, Catherine. "The Eternal Return and the Phantom of Difference." *Parrhesia* 10. 2010.
Small, Robin. "Time and Becoming in Nietzsche's Thought." *Bloomsbury Studies in Continental
         Philosophy*. London: Continuum, 2010.
Ward, Koral. *Augenblick: The Concept of the 'Decisive Moment' in 19th- and 20th-century
         Western Philosophy (1st ed.)*. Routledge, 2008

---

[102] Z.III.2

What's So Bad About Lying?

Sophie Macdonald

Introduction

In *Lying, Misleading, and What Is Said,* Jennifer Saul offers a nuanced definition of *lying*. In philosophy of language, lying is often thought of as 'saying false things,' however, Saul believes that we lose moral repercussions when we restrict ourselves to this definition. Saul outlines an example of a politician named 'Tony', who believes that there are no weapons of mass destruction in Iraq and then utters "there are weapons of mass destruction in Iraq."[103] In this scenario, Iraq *does* have weapons of mass destruction. Despite this, Saul argues that Tony should not be let off the moral hook for accidentally saying a true thing. Under the broad definition of *lying*, Tony would incur no moral repercussions. Saul uses moral repercussions to determine what utterances we should consider *lies*. I will use Saul as a springboard for my own project: the moral and epistemic consequences of lying make it so linguistic practices *require* a warranting context. Thus, a linguistic convention is to *assume,* unless otherwise agreed upon, that we are *always* in a warranting context.

Groundwork

Saul defines lying as: "If the speaker is not the victim of linguistic error/malapropism or using metaphor, hyperbole, or irony, then they lie iff (1) they say that P; (2) they believe P to be false; (3) they take themselves to be in a warranting context."[104] The preliminary supposition is straightforward and intuitive. If someone mistakenly calls a donkey a horse, we are not prone to label them a liar. In these preliminary cases, most people can discern what is occurring and receive the communicative message. Many listeners are quick to pick out instances of error, metaphor, hyperbole, and irony to retrieve the speaker's intended message. This is a process we are so adept at, that we cannot fathom calling someone a liar when they engage in one or more of these processes. Something to note, misleading often relies upon exploiting these concepts. Bill

---

[103] Jennifer Saul, *Lying, Misleading, and What Is Said* (Oxford University Press, 2012), 4.
[104] Saul, *Lying,* 19.

Clinton may not have *lied* when addressing his relationship with Monica Lewinsky, but he certainly did mislead by exploiting linguistic processes. Before we continue unpacking Saul's definition of lying, let us spend a few moments establishing what *misleading* is. Two major components constitute misleading for Saul: it is a success term and it offers deniability.[105] Meaning, a misleader must successfully mislead their audience while maintaining deniability that they misled their audience. Common cases of misleading exhibit misleaders manipulating pragmatic functions of conversation. To return to Bill Clinton, he did not outrightly say a false proposition, he said: "there is no improper relationship."[106] By employing a present tense verb, Clinton led his audience to believe there was *never* an improper relationship, when in truth, there was no improper relationship *at the moment.* Clinton exploited the pragmatic function of implicature. His audience was clearly concerned with the existence of a relationship *at any point in time,* an intention Clinton was aware of. Yet, by not asking for clarification if he did not understand the question, and directly answering the question, his audience could reasonably believe he understood what information they desired to know. Misleading is not central to our debate. However, it is important to note that other abuses of linguistic practices, like misleading, can yield similar consequences as lying.

To continue with lying, premise one is concerned with something that actually appears on the discourse record. Often, propositions are considered the only vessels capable of truth conditions. Thus, it is much easier to determine instances of lies when (1) we have a proposition with truth conditions, and (2) we are able to isolate and examine the utterance. In premise two, the liar realizes they are saying something false. Which refers back to the preliminary supposition, the liar knows they are not making an error or using a linguistic tool, they *intentionally* utter something they believe to be false. By using the word *believe* instead of *know,* Saul introduces an additional point: The liar does not need to know what they are saying is actually false, they simply need to *believe* it is false. For Saul, lying is morally reprehensible because of the *intention* to deceive. The focus of premise three is whether we are in a situation where the truth is expected to maintain the quality of the conversations – a warranting context.

---

[105] Saul, *Lying*, 71 and 78.
[106] Ibid, 2.

Contexts of Assertions

The bedrock of this paper is the reasonable assumption that we are nearly always in a warranting context. A warranting context is one where "an audience has a *right to expect* that the speaker believes what is said to be true."[107] Why should we expect that speakers make true utterances? These roots trace to H.P. Grice and his maxims. Specifically, the supermaxim of Quality which dictates: "try to make your contribution one that is true."[108] As Grice outlines, dealing exclusively with the truth is not an absurd notion – it is a basic maxim we are expected to uphold. However, what would a warranting context look like? Often in philosophy of language, the courtroom is invoked. If you are about to take the stand, you swear an oath to tell the truth, undeniably placing you in a warranting context where you are *expected* to tell the truth. If we generally think of ourselves in a warranting context, when did we swear to tell the truth? Often, there is no particular instance. Rather, it is an adopted convention convention and one that we take ourselves to be in most of the time, if not all of the time. Only those with a love of lexical warfare or deception will try to refute this assumption. Though we are conventionally in a warranting context, there are instances where you could be in a non-warranting context. We can imagine scenarios where people have agreed that lying is permissible, such as in games. Consider a game that involves bluffing, such as Poker or Truc. In such games, a player is permitted to *bluff,* which under Saul's definition of lying, does technically indict these players. However, it is unlikely for us to charge these players with the moral consequences that other instantiations of lies would incur. Afterall, by joining the game players have agreed that tactics such as bluffing are permissible and necessary to gameplay. There is a more challenging example I will raise. Imagine the Gestapo abducts you and places you in an interrogation room, demanding the truth of an event. Whether you take yourself to be in a warranting or non-warranting context varies in this example. On one hand, you may question whether the Gestapo has the authority to demand a non-warranting context. While on the other hand, you may believe that there is no such thing as a justified non-warranting context, no matter how unique the situation is. We will turn to Kant and Constant's disagreement on this issue to further explore the minutiae details of warranting and non-warranting contexts.

---

[107] Saul, *Lying,* 75.

[108] H.P. Grice, "Logic and Conversation," in *Syntax and Semantics*, ed. Peter Cole and Jerry L. Morgan (New York: Academic Press, 1975), 46.

Obligated to The Truth

In the Gestapo example, there are two major replies: (1) you are morally justified to lie to the Gestapo, and (2) you are *not* morally justified to lie to the Gestapo. Constant aligns with response (1) and Kant aligns with response (2). Constant, and those who agree with response (1), do so because "to tell the truth is a duty but only to someone who *has a right to the truth.*"[109] We only owe the truth to those who do not trample upon our rights. The move Constant makes is that normally we are in a warranting context, but by infringing upon your rights, the Gestapo have changed the context to a non-warranting one. Since the Gestapo have altered the linguistic playing field, you are under no obligation to tell the truth *and* there will be no repercussions if you lie. Kant heartily disagrees with response (1). For Kant, "truthfulness (if he must speak) is an unconditional duty."[110] No case exists where it is morally permissible to lie, even if the murderer is at your door or the Gestapo have locked you in an interrogation chamber. Under Kant's philosophy, we are *always* in a warranting context. This discussion illuminates that there are cases where it is unclear if you are in a warranting or non-warranting context, however, it is only in exceptional cases where this vagueness arises. We would only take ourselves to be in a non-warranting context if: (1) we explicitly agreed to it, or (2) we are examining an outlier case, such as the Gestapo example. Yet, even with these outlier examples, there is no consensus that it *is* morally permissible to lie, only that it *might* be.

Obligations to Whom?

The general consensus is that we should strive for truth in communications. But to whom do we owe the truth to? In "Dealing with Overhearers," Edward Schaefer and Herbert Clark argue we have obligations only to our direct addressees. Therefore, we have no real obligations towards overhearers – those who we are not directly communicating with. Furthermore, overhearers "don't share in or benefit from conversational responsibilities."[111] While speakers have no obligations to overhearers, overhearers also have no obligations to the conversation. I

---

[109] Immanuel Kant, "On a Supposed Right to Lie from Philanthropy (1797)," in *Practical Philosophy*, ed. Mary Gregor (Cambridge: Cambridge University Press, 1996), 614.

[110] Ibid.

[111] Herbert Clark and Edward Schaefer, "Dealing with Overhearers," in *Arenas of Language Use* (University of Chicago Press, 1992), 3.

disagree with Schaefer and Clark's claim that obligations flow from the speaker to the direct audience, as they are only really considering conversations between two individuals, when we know that there are a myriad of permutations of speaker(s) and audience linguistic practices that can occur. Before I express my worries in more detail, we will turn to an example:

> *Cafe Dilemma:* Jerry is standing in line at a cafe and sees his coworkers – Lisa and Alma, at a table nearby. Jerry works in the human resource department at their company and Alma is a recent hire. Jerry overhears Lisa say to Alma: "I overheard the big boss mention your name Alma, she isn't impressed with you – she's going to fire you." Jerry knows that their boss was not discussing firing Alma, Lisa is just trying to haze Alma.

Under Schaefer and Clark's theory, Lisa has no conversational obligation to Jerry and Jerry has no obligations to that conversation. However, does Jerry not have an obligation to monitor the quality of the conversation if he hears things he plainly knows to be false? Say Alma sleeps at the office to ensure she is never late. Is Lisa the only one responsible for Alma's subsequent actions based on this information? No, I believe Jerry also holds responsibility, because Jerry knew the truth but did nothing. No matter one's supposed speaker and audience role, there are obligations to truth. Monitoring the quality of conversations you are able to, where you know what is true and false, you are obligated to step in if speakers begin to lie. Obligations of truth are at play even if you are not the speaker.

Semantic Truth vs. Pragmatic Truth

If we have such strong obligations to truth, what does it mean for something to be true? The discussion in this arena is vast, though truth is often thought of in semantic terms. However, there are philosophers like Wilson and Sperber, who advocate for a pragmatic formulation of truth. I will outline both the semantic and pragmatic conception of truth to demonstrate that we need aspects of both theories.

In semantic theories of truth, a basic assumption is that "propositions are the primary truth-bearers."[112] Explicature carries the truth conditions of a linguistic action. This seems reasonable *prima facie* – what is said can either be true or false. Philosophers like Donaldson and

---

[112] Matti Eklund, "Theories of Truth," in *The Routledge Companion to Philosophy of Language*, ed. Gillian Russell and Delia Graff Fara (Routledge, 2011), 199.

Lewis claim: "an utterance can be said to have a literal meaning which is capable of being either true or false when the result of combining its linguistic sense with its reference is a proposition."[113] Essentially, an utterance is true when its proposition property refers to its reference. If Cleo says: "Hey, stranger" to a person she has never met, she is saying something true. Her proposition includes reference to a person she has never met before, which she says to a person she has never met before. If Cleo says this to someone she *has* met before, say her mother, her utterance would be false. Semantic truth is primarily concerned with using language in a consistent and comprehensible manner. If you properly use language, you should not be susceptible to saying false things. We can see how the semanticist formulation of truth required Saul to create a broader definition of lying than simply 'saying false things.' Because, as Saul accounts for, when Cleo says 'Hey, stranger' to her mother we are not likely to call her a liar, even though Cleo did utter something false – we understand this to be a teasing utterance when addressed to someone we are acquainted with. Although the semanticist's picture is unable to capture more complex practices, such as lying, it does give us part of our obligation to truth – we should use language, to the best of our ability, correctly and use our propositions to properly refer to our intended reference. If we do so, we are placing a lower cognitive load on our audience and making our communicative message much easier to retrieve. We now have one premise in our obligation to truth.

Sperber and Wilson argue that semantic truth misses aspects that we want our formulation of truth to have. They write: "in actual English [...there are] literal meanings plus other figurative meanings,"[114] where the semantic theory does not properly account for anything beyond literal meaning. The semanticist may reply that an instance of figurative meaning signals that a Gricean maxim has been broken, so implicature needs to be recovered. Sperber and Wilson deny this reply, because (specifically in instances of trope) recovering implicature will not *justify* breaking the maxim in the first place.[115] Sperber and Wilson claim that "semantic vagueness clearly exists,"[116] in certain cases, propositions have no clear reference because the reference itself has no single literal meaning. Sperber and Wilson are subtly attacking a deeper point of contention between the semanticist and pragmatist – that semantic meaning is highly specific and

---

[113] Deirdre Wilson and Dan Sperber, "Truthfulness and Relevance," *Mind* 111, no. 443 (2002): 586.
[114] Sperber and Wilson, "Truthfulness," 587.
[115] Ibid, 589.
[116] Ibid, 595.

traceable, while pragmatic meaning is loose and determined from instance to instance. Sperber and Wilson are not necessarily denying this implicit belief concerning pragmatics. Rather, they highlight that semantics is not as neat and tidy as semanticists often present it to be. A result of this argument is the claim that: "linguistically-encoded meaning is far too schematic and fragmentary to be capable of being true or false; it is just an input to further processing."[117] Meaning, there are no clear criteria we can follow to determine if an utterance is true or false. They are shifting emphasis away from truth conditions and towards *relevance,* where "an utterance is *relevant* when the hearer, given his cognitive dispositions and the context, is likely to derive some genuine knowledge from it."[118] Sperber and Wilson are working under a more collaborative framework of communication – successful communication allows hearers to derive genuine knowledge from the speaker's utterance. This leads us to the second premise in our obligation for truth, that linguistic practices require speakers *not* to break the trust that hearers have placed in them. Thus, the speaker's obligations are: (1) to format propositions to the best of their ability, (2) avoid deceiving our audience, and (3) maintain the trust the audience has placed in them. If a speaker adheres these three obligations, the audience has a better chance retrieving the communicative message and deriving genuine knowledge.

Epistemic Norms

We have three obligations of truth for speakers, but how do they work in practice? Epistemic norms create specific criteria for assertions and acquiring knowledge, so we will turn to them to answer this question. Rachel McKinnon proposes the Supportive Reason Norm (SRN) as an epistemic norm.[119] SRN is defined as: "(i) One has supportive reasons for *p,* (ii) The relevant conventional and pragmatic elements of the context are present, and (iii) One asserts that *p* at least in part because the assertion that *p* satisfies (i) and (ii)."[120] SRN is much more flexible and adaptable than the more common truth and knowledge based norms. Evidently, Kant leans heavily onto the knowledge based norm. Where you assert things because you take them as knowledge, where knowledge for him is facts. In our discussion, we want to integrate SRN because truth and knowledge norms are problematic. Truth norms are problematic because there

---

[117] Ibid, 628.
[118] Ibid, 599.
[119] Rachel McKinnon, *The Norms of Assertion: Truth, Lies, and Warrant*, (Palgrave Macmillan UK, 2015), 4.
[120] Ibid.

is a simple question that is incredibly difficult to answer: how do I know p is true? Similarly, to see knowledge norms fall apart we need only ask: How do I know p is knowledge? We can answer these questions with varying degrees of certainty. However, in each answer we will likely appeal to context. Thus, it is best that we have a norm that already incorporates context into its definition. However, in all three of these norms, lying violates and weakens their integrity. Which is the point Kant drives at, bit by bit, lying rots our norms until nothing remains.

Case Study

The discussion thus far has been rather theoretical. To ground us, I will outline a case study. In recent news, Mary Turpel-Lafond is being examined for claiming false indigenous ancestry. Turpel-Lafond is a prominent legal figure and advocate for children's rights. "[Winona Wheeler] said Turpel-Lafond owes the public some answers about her ancestry because she has had a high-profile career while representing herself as Indigenous. For example, she has claimed to be the first treaty Indian appointed to the Saskatchewan court."[121] Wheeler highlights public figures in the political sphere are always in a warranting context. Figures, such as Turpel-Lafond *owe* the people the truth, by skirting this responsibility, Turpel-Lafond becomes suspect in the public's eye. The issue is centred around semantic truth – "in other words, in claiming to be a treaty Indian, Turpel-Lafond was claiming to be on the federal government's registry of status Indians."[122] Consider: "Turpel-Lafond has for decades publicly said she is a treaty Indian. For example, when CBC's Pamela Wallin asked her about her Indigenous status in a nationally televised interview in 1999, Turpel-Lafond said, 'I'm a treaty Indian.'[123] When Turpel-Lafond uttered 'I am a treaty Indian,' her proposition was directly referring to being on the government's registry of status Indigenous peoples. Her proposition can only be true if her name is actually on that registry. Evidence points that Turpel-Lafond is not on this registry and that she is aware of that fact. Perhaps Turpel-Lafond's belief was mistaken and her utterances were false simply through error. If that is the case, she has not taken any steps to set the record straight. When CBC reached out to her, "she didn't reply," and continuously avoided directly answering questions

---

[121] Geoff Leo, "Disputed History," CBC News, October 12, 2022, https://www.cbc.ca/newsinteractives/features/mary-ellen-turpel-lafond-indigenous-cree-claims.
[122] Ibid.
[123] Ibid.

concerning her heritage[124]. Thus, we must consider this case with the evidence we have, which indicates that she has intentionally lied about her heritage. A further problem within this case-study is: "non-Indigenous people taking away opportunities from First Nations, Métis and Inuit people by improperly claiming Indigenous ancestry."[125] By claiming indigenous ancestry, Turpel-Lafond claimed an opportunity that perhaps, otherwise may not have been available to her. Since, the nature of her legal career was very much tied to the rights of indigenous children and representing indigenous peoples, importance was placed on heritage. Clearly, there are several morally suspect things occuring in this case study: avoiding her obligation to truth (especially as a public figure), capitalizing on a marginalized group, and creating distrust. In the following, we will examine the consequences of creating distrust through lying.

Consequences

I briefly mentioned that lying rots our epistemic norms. A multitude of reasons explain this: (1) conversations require trust that speakers will respect the warranting context; (2) lying, while not always saying false things, does not adhere to our obligations of truth; (3) hearers are unable to derive genuine knowledge from a speaker's lies; and (4) without the ability to derive genuine knowledge our epistemic norms are useless. You may wonder why we care about preserving epistemic norms. Consider a world with no epistemic norms: we would have no consistent way to formulate knowledge. Perhaps you could still have knowledge without epistemic norms, but what you consider *knowledge* would be a constantly moving field post. Since the definition is always in flux, there would be no certainty if something was knowledge or not. While this is just a plainly horrible consequence, there is an additional consequence. When people lie in a warranting context, they break the trust of their audience. Sometimes, a speaker's lies remain undiscovered, but consider lying analogous to smoking. One cigarette may not kill you, but if you keep smoking, you increase your likelihood of consequences such as lung cancer. If a single person lied one time, without getting caught, perhaps that is not enough to create distrust between speakers and audiences. However, the more instances where people lie, the higher the likelihood that one lie *will* be caught. The more people lie, the more instances of lies we will catch, even if some lies go unnoticed. Thus, with each lie we catch, our distrust of

---

[124] Geoff Leo, "Disputed History."
[125] Ibid.

speakers rises. Recall, the Turpel-Lafond case, are we surprised that a public figure got caught in a lie? Unlikely, we are accustomed to public figures lying, and thus, our trust in the honesty of public figures is likely quite low. Nearly everyone realizes that it is not ideal to have no trust in public figures, especially legal or political ones, yet, that is the situation we live in. For instance, the OECD found that disadvantaged groups and younger people (ages 18-29) have especially low trust in the government.[126] What then are the consequences of the dissolution of trust? Several things occur in the absence of trust, the most salient to our discussion being the loss of reliable testimony. Arindam Chakrabarti writes: "I think every competent user of language can on many occasions pass on knowledge and receive genuine knowledge through telling and being told."[127] Chakrabarti refers to this as "knowledge by testimony,"[128] since it is through the testimony that you tell or are told, that knowledge is gained. A large portion of our knowledge comes from testimony, the language you speak is the result of testimony. Indeed, nearly everything you learned as a child was through the testimony of adults. The role of testimony does not dissolve when you are no longer a child, even as an adult there are many things you rely on testimony for. You take the testimony of the weather channel to determine if you should pack an umbrella. You take the testimony of health officials to determine if you should get vaccinated. Here arise the consequences of no trust. If you do not trust your health officials, perhaps you will not get vaccinated. If you do not trust the honesty of political figures, perhaps you will not vote. Trust in testimony is the heart of knowledge, when this trust is broken it is no small matter.

Conclusion

There are serious epistemic and moral consequences to lying. The liar breaks their obligation to truth and thus becomes morally reprehensible. Recall, speakers are not the only ones with obligations to the truth, audience and overhearers have an obligation to step in if they know that a speaker is lying. If they do not, they are morally reprehensible as well, though perhaps not to the degree the liar is. Demonstrated in the Gestapo example, there can be conflict if a lie is considered *always* morally reprehensible. Though this is an outlier case, most other instances of lies will incur moral consequences. The moral consequences stem from the liar's

---

[126] "Trust in Government."
[127] Arindam Chakrabarti, "Telling as Letting Know," in *Philosophy of Language: The Big Questions* (Blackwell Publishers, 1998), 138.
[128] Ibid, 136.

demolition of trust that the audience has placed in them. Without trust, we cannot formulate knowledge by testimony which is an integral part of our epistemic system. A question to consider as we close: can trust be regained? Perhaps trust could be recovered and perhaps not. If trust *could* be recovered, it would be through the truth. The liar(s) would have to adhere to the obligations of truth. We return once again to the undeniable duty to truth, which is a combination of semantic and pragmatic formulations. To properly communicate with others, we have a duty to uphold a warranting context, breaking this trust incurs too great of moral and epistemic consequences.

Bibliography

Chakrabarti, Arindam. "Telling as Letting Know." In *Philosophy of Language: The Big Questions*, 136-44. Blackwell Publishers, 1998.

Clark, Herbert and Edward Schaefer. "Dealing with Overhearers." In *Arenas of Language Use.* University of Chicago Press, 1992.

Donaldson, Tom and Ernie Lepore. "Context-Sensitivity." In *The Routledge Companion to Philosophy of Language,* edited by Gillian Russell and Delia Graff Fara, 116-131. Routledge, 2011.

Eklund, Matti. "Theories of Truth." In *The Routledge Companion to Philosophy of Language*, edited by Gillian Russell and Delia Graff Fara, 199-208. Routledge, 2011.

Grice, H. P.. "Logic and Conversation." In *Syntax and Semantics,* edited by Peter Cole, & Jerry L. Morgan, 41-58. Vol. 3 of *Speech Acts.* New York: Academic Press*,* 1975. https://doi.org/10.1163/9789004368811_003.

Kant, Immanuel. "On a Supposed Right to Lie from Philanthropy (1797)." In *Practical Philosophy*, edited by Mary Gregor, 605–616. Cambridge: Cambridge University Press, 1996.

Kupfer, Joseph. "The Moral Presumption Against Lying." *Review of Metaphysics* 36, no. 1 (1982): 103. http://www.jstor.org/stable/20127795.

Leo, Geoff. "Disputed History." CBC News, October 12, 2022. https://www.cbc.ca/newsinteractives/features/mary-ellen-turpel-lafond-indigenous-cree-claims.

McKinnon, Rachel. *The Norms of Assertion: Truth, Lies, and Warrant.* Palgrave Macmillan UK, 2015.

Saul, Jennifer. *Lying, Misleading, and What Is Said.* Oxford University Press, 2012.

Unknown. "Trust in Government." *OECD*. https://www.oecd.org/governance/trust-in-government/.

Williams, J. Robert. "Vagueness." In *The Routledge Companion to Philosophy of Language,* edited by Gillian Russell and Delia Graff Fara, 143-152. Routledge, 2011.

Wilson, Deirdre and Dan Sperber. "Truthfulness and Relevance." *Mind* 111, no. 443 (2002): 583–632. http://www.jstor.org/stable/3093623.

Intersectionality as a Flexible Framework:
How Intersectionality Should Be Applied to Climate Justice

Elia Demers

Kimberlé Crenshaw created the concept of intersectionality as a legal tool for considering the systemic discrimination Black women and other women of color face. Intersectionality is therefore a structural concept that does not seek to address personal subjective experience. I will argue that intersectionality can be extended to other structural identities beyond race and gender and is a necessary framework for evaluating other social justice issues, such as climate change. However, in this process of adopting intersectionality in other contexts, it is essential that its origins are acknowledged and learned in order to keep the voices of women of color present in the conversation. Understanding intersectionality in reference to Crenshaw's work is important for appropriately expanding the concept. I will first discuss Crenshaw's characterization of intersectionality and the importance of considering structural relationships rather than subjective ones. I will then analyze a particular case where the origins of an activism strategy were not acknowledged, and how this process of utilizing another group's methods of activism without respecting the origins of those methods leads to the further marginalization of that community. Next, I will provide a parallel example of how frameworks of activism can be extended beyond their original purpose successfully by highlighting Hannah-Jones' work in the 1619 project. I will conclude the essay with a key discussion of climate justice, and how it is necessary to broaden the definition of intersectionality to include other issues such as climate change in order to consider marginalized voices and produce equitable policies.

The concept of intersectionality derives from Crenshaw's social justice work pertaining to women of color. Women of color are often left out of both feminist and anti-racist discourse and this exclusion contributes to structural and individual violence.[129] Structural violence can consist of the institutional processes which contribute to harming people and groups through unjust or discriminatory practices that prevent access to necessary resources, rights, or outcomes.

---

[129] Kimberlé Crenshaw, "Mapping the Margins: Intersectionality, Identity Politics, and Violence against Women of Color," *Stanford Law Review* 43, no. 6 (July 1991): 1244.

In other words, when an institution, such as a government, restricts a population's ability to have opportunities or receive important aids or benefits due to discrimination, that population faces structural violence. Individual violence can be understood as targeted harm towards specific people. An example I will later discuss is sexual and domestic violence towards women of color. At the individual level, women are harmed by an abuser. At the structural level, legal cases surrounding this abuse are dismissed, and these women are not adequately protected by the law or justice system.

Intersectionality is a tool that can give visibility to Black women and other women of color in legal and political settings, leading to increased protection from these forms of violence. The discrimination women of color face is often located at a unique position of intersecting discriminations against their race and gender simultaneously.[130] By focusing on the way race and gender intersect, intersectionality provides a framework which recognizes the particular kinds of marginalization women of color face. For Crenshaw, this framework of intersectionality aims to create terminology, discourse, and strategies for understanding the ways in which women of color are situated within two groups that have potentially differing political agendas.[131] In feminism and anti-racism, the full picture of each movement is limited by the dominant perspectives and agendas that have driven the current narratives surrounding political emancipation. More specifically, the way women of color experience sexism is not the same as white women, and the way women of color experience racism is not the same as men of color.[132] Feminism and anti-racism both fail to integrate women of color's voices by ignoring the issues relating to either movement.[133] Feminism does not examine race enough, and anti-racism does not consider gender enough. These shortcomings reproduce and continue the subordination of women of color.[134] A new analysis of race and gender through an intersectional lens can enable women of color to find representation within these legal and political contexts.

The limited scope of both feminist and anti-racism movements can be better articulated through a discussion of violence against women of color. Crenshaw examines cases of domestic abuse and rape against women of color from a legal viewpoint to reveal the inability of non-intersectional laws to protect these groups of women. In legal contexts, the sexism and

---

[130] Crenshaw, 1244.
[131] Crenshaw, 1252.
[132] Crenshaw, 1252.
[133] Crenshaw, 1252.
[134] Crenshaw, 1252.

racism faced by women of color who are victims of rape and domestic violence cannot be separated and understood in isolation from one another.[135] Oftentimes the legal cases of women of color are dismissed, and their abusers are not convicted or face only minor repercussions. These legal consequences are not a matter of discrimination against gender or discrimination against race alone but are a matter of the intersecting forms of discrimination in which women of color are oppressed and harmed. In the existing movements of feminism and anti-racism, addressing the specific forms of violence that impact women of color is not always a central part of the agenda. Feminism that does not consider race and anti-racism that does not consider gender will overlook this form of violence, at both the individual and structural level. Therefore, the application of intersectionality is crucial in combating this violence.

It is important to clarify exactly how intersectionality is defined, insofar as intersectionality is not a theory of identity and is not limited to only race and gender.[136] Intersectionality is not about individual identity and subjective experiences but rather the larger positions of power and discrimination within a system in which individuals find themselves. As Crenshaw writes, "the organized identity groups in which we find ourselves in are in fact coalitions."[137] In other words, our structural identities relate to the coalitions and groups we are in that are influenced by intersecting forms of marginalization or discrimination. Intersectionality does not relate to personal experiences and presentations of identity but to the larger context in which structural identities are influenced. A person cannot be intersectional, and the elements of their identity do not intersect within themselves. Instead, a person's circumstances and the ways in which they are impacted by discrimination can be intersectional. People are then positioned in society at these intersections of discrimination. As Crenshaw discusses in an interview, identity is not an individually contained concept but a relationship between people and social systems, history, and institutions.[138] Ultimately, intersectionality allows for coalitions between different structural identity groups that face intersecting forms of discrimination but is not a term to describe the subjective experience of identity.

If intersectionality is best understood as a tool that evaluates structural relationships of discrimination, then its application to complex social, political, and legal issues could contribute

---

[135] Crenshaw, 1277-1278.
[136] Crenshaw, 1244.
[137] Crenshaw, 1299.
[138] Kimberlé Crenshaw, "What is Intersectionality?" Interview by National Association of Independent Schools. YouTube, June 22, 2018. Video, 1:11-1:20. https://www.youtube.com/watch?v=ViDtnfQ9FHc.

to the creation of more effective policies. Climate change is one such issue that impacts populations at differing levels based on preexisting social systems, including forms of discrimination and injustice. For instance, one consequence of global warming is increased 'urban heat islands' in cities.[139] When a city has little vegetation and tree cover, the concrete buildings and streets will trap and reflect heat in the city, especially during a heat wave, making it more dangerous for people living there.[140] However, not everyone lives in the same location and environment. Those most susceptible to the heat island and heat waves are those living in lower income areas of densely populated cities that have insufficiently funded the incorporation of green spaces. Understanding the demographics of these lower income neighborhoods requires an examination of institutional racism. Discrimination against race at a structural level has resulted in higher concentrations of minority groups living in lower income cities. Discrimination against lower income populations has resulted in a lack of resources and support for establishing green spaces in cities, considering the fact that low income neighborhoods have less green space and parks.[141] Both forms of discrimination intersect to place certain communities at a disproportionately high level of risk to heat exposure. Through this application of an intersectional analysis to the problem of climate change, it becomes apparent that achieving climate justice requires social reform.

It has already been discussed by Crenshaw that creating legal and political change surrounding anything unrecognized in the current discourse requires intersectionality. Climate change will contribute to a variety of inequities that will not be entirely addressed within the current, non-intersectional legal and political paradigms, particularly since many of these injustices will be multifaceted. Therefore, intersectionality can be utilized to reframe our understanding of climate injustice, which can result in more comprehensive policies and movements.

While the scope of intersectionality can be expanded to incorporate more forms of structural discrimination, within broader contexts such as climate change, this process of expansion should be accompanied by an acknowledgment of intersectionality's origins. If the

---

[139] Dan Li and Eli Bou-Zeid, "Synergistic interactions between urban heat islands and heat waves: the impact in cities is larger than the sum of its parts," *Journal of Applied Meteorology and Climatology* 52, no. 9 (2013): 2051-2052.
[140] Li and Bou-Zeid, "Synergistic interactions," 2052-2053.
[141] Salvatore Saporito and Daniel Casey. "Are there Relationships among Racial Segregation, Economic Isolation, and Proximity to Green Space?" *Human Ecology Review* 21, no. 2 (2015): 123, 126, 128.

foundation of intersectionality is ignored or unlearned, it would harm those the concept was initially intended to support. Intersectionality can be extended to reach other marginalized groups and intersecting forms of discrimination beyond race and gender, as Crenshaw herself touches upon,[142] but this process of extension should continue to uplift the origins of intersectionality as a way to keep women of color present in the conversation out of which they have often been left. Some ways of appropriately utilizing intersectionality are citing Crenshaw when defining intersectionality, referencing the history of the term and who originally established it, and ensuring intersectionality is correctly understood prior to extending this concept beyond race and gender. The purpose of such actions is not simply to reiterate the origins of intersectionality rhetorically, but to ensure the goal of intersectionality itself is not forgotten as intersectional analyses expand in size and complexity. Misrepresenting intersectionality as a concept of identity and subjectivity is one such way of diminishing its strength, credibility, and foundation. Intersectionality relates to specific systemic discrimination at intersecting points, not the specific categories of each person's identity.

In order to better articulate the risks of applying a tool without acknowledging its foundation, a separate example about the framework of activism can be examined. Carbado discusses how the use of activism strategies without a recognition of their origins results in a naturalization of whiteness and an erasure of Black history. Carbado uses the example of the Don't Ask, Don't Tell policy which prevented homosexual people from joining the military. Gay rights activists protested this policy by comparing it to the exclusion of African Americans from the military in the past.[143] This analogy between gay people and Black people also creates an equivalency between race-based and sexual orientation-based discrimination in the military, which diminishes civil rights history, leaves out Black gay and lesbian people, and naturalizes whiteness as the default race of the gay liberation movement.[144] By aiming to replace the previous discussion of Black men being excluded from the military with homosexual men being excluded, Black civil rights history is displaced, and the intersections between homosexuality and Blackness are erased from the conversation.[145]

---

[142] Crenshaw, "Mapping the Margins," 1299.
[143] Devon W. Carbado, "Colorblind Intersectionality," *Signs* 38, no. 4 (2013): 827.
[144] Carbado, 827.
[145] Carbado, 829.

Further, Carbado explains that gay rights activists in this case did not entirely ignore Black civil rights history but leveraged elements of that history to their advantage, because they utilized the "moral authority of the civil rights movement."[146] These gay rights activists created arguments to support their case of discrimination that derived from the original arguments made by Black civil rights activists regarding the moral implications of exclusion based on race. However, these activists did not engage with the details of civil rights history and therefore displaced Black people from the conversation entirely.[147] There was no acknowledgement of the particular struggles of Black civil rights activists or even of the intersecting struggles faced by gay Black people in the context of the military. Gay rights activists used the framework of activism developed by civil rights activists without acknowledging civil rights history, which further marginalized the Black community. This case reveals the dangers of misrepresenting the origins of a framework. It can serve as an analogy to the expansion of intersectionality in that it reveals what would result from disregarding how Black women and other women of color paved the way for this tool. Ignoring the foundation of intersectionality when applying it in other contexts is detrimental to the original purpose of intersectionality and results in the perpetuated subordination of those who established the framework.

While extending the definition of intersectionality incorrectly can be highly problematic, expanding its scope in an inclusive and considerate manner can be incredibly beneficial. Nikole Hannah-Jones created the 1619 Project as a way to reimagine the foundation of the United States of America and the meaning of democracy in relation to the fight for liberation by Black Americans throughout history. Because Crenshaw's intersectionality also stems from the judicial, political, and social systems within the US, it is appropriate to examine another parallel example from within the US context. Hannah-Jones addresses the long history of Black Americans fighting for their freedom and rights. She discusses how Black Americans are "foundational to the idea of American freedom [and]...have been the perfecters of this democracy."[148] Since the beginning of the country's development, Black Americans have held the country accountable to the ideals written on paper and have shaped the true meaning of democracy and freedom. Even if those who created the Constitution did not believe the words they wrote about freedom and

---

[146] Carbado, 831.
[147] Carbado, 831.
[148] Nikole Hannah-Jones, "Our democracy's founding ideals were false when they were written. Black Americans have fought to make them true," *New York Times*, para 10, August 14, 2019. https://www.nytimes.com/interactive/2019/08/14/magazine/black-history-american-democracy.html.

equality, Black Americans have continued to believe in those words and have fought to make them true in reality, for everyone.

This ongoing fight for freedom has opened many pathways for other movements and groups to follow. The political involvement of Black Americans helped to create change, such as the 14th Amendment which grants citizenship to anyone born in the US. Since this initial fight for rights by Black people in America, "nearly all other marginalized groups have used the 14th Amendment in their fights for equality."[149] Beyond this single amendment, many other legal, political, and social changes have occurred due to the liberation movements of Black Americans in their ongoing struggle for freedom.[150] Remarkably, Black Americans have been fighting for far more than their own rights. Abolitionist movements, civil rights movements, and modern day activism have paved the way for all other current liberation movements involving many other minority groups in America.[151] Black resistance has shaped modern democracy and has created a framework for other minorities and marginalized groups in fighting for their rights.[152] The purpose of Hannah-Jones' work is to recenter Black Americans in the conversation of American history in order to understand this activist framework, its origins, and its evolution.

The 1619 Project has proven the success and benefits of using specific frameworks in broader contexts to support and protect the rights of many different people and groups. The crucial detail is how this process of adopting frameworks is done while maintaining the integrity of the framework. In other words, one cannot forget the origins of this foundation for activism and the contributions Black Americans have made to open pathways for so many other people who face discrimination. Carbado explains the dangers of utilizing strategies and work done by Black people without acknowledging the importance of that history. If we as a society attempt to apply theories or frameworks, such as intersectionality and activism, without accepting, admitting, learning, and appreciating how these theories and frameworks were developed and by whom, those who fought to create these useful, empowering tools will be erased from the conversation. Therefore, intersectionality certainly can be expanded in its definition to reach other marginalized groups who face different forms of intersecting discrimination beyond race and gender. Black women and other women of color must be involved in the conversation and

---

[149] Hannah-Jones, para 33.
[150] Hannah-Jones, para 31-35.
[151] Hannah-Jones, para 31-35.
[152] Hannah-Jones, para 31-35; 45-46.

recognized as the group who established intersectionality as a legal tool in order to expand the scope of intersectionality's definition. Once the scope is properly expanded, intersectionality can become a very powerful tool for achieving justice in any context that is not adequately recognized in existing activist movements.

It can be argued that intersectionality is not only beneficial in other contexts if applied appropriately, but intersectionality is actually *necessary* for understanding and improving our ability to combat discrimination in these other contexts. Without the application of intersectionality to contexts beyond race and gender discrimination, strategies for legally and politically addressing these dynamic problems will fail to be comprehensive and will remain incomplete. For example, climate change is an intersectional issue. Some people and groups around the world are impacted by climate change both directly and indirectly at disproportionate levels than others. Climate change is obviously a global event that does and will impact humans across the world; however, the extent to which some countries, communities, identity groups, and individuals will be impacted negatively by climate change is not equal. There are many social dimensions that intersect in relation to climate change, such as gender, race, location, and economic status, that are important to consider when addressing and contextualizing climate change.[153]

Kaijser and Kronsell make use of an intersectional framework in order to expand the understanding of climate justice, with reference to Crenshaw. They acknowledge Crenshaw's foundational definition of intersectionality as a tool for incorporating the voices of women of color, but add that this framework is successful in exposing different power structures and discriminatory institutions pertaining to climate change.[154] For example, understanding how class and economic status merge with several other factors reveals how those who are the largest contributors to climate change and those who are the largest sufferers of climate change are often different groups.[155] There is no simple way to understand climate injustice through single factors of marginalization; the issue is far too complex and integrates a variety of discriminatory factors and power struggles.[156]

---

[153] Anna Kaijser and Annica Kronsell, "Climate change through the lens of intersectionality," *Environmental Politics* 23, no. 3 (2013): 418-420.
[154] Kaijser and Kronsell, 419.
[155] Kaijser and Kronsell, 421.
[156] Kaijser and Kronsell, 419-422.

Consider a particular scenario relating to climate change. A single mother who is both racialized and lives in a low income neighborhood will experience significantly more difficulty adjusting to climate disasters. Her ability to support herself and her child financially while protecting their health from exposure to climate disasters is severely impacted by discrimination against her race, gender, and income. These three factors of discrimination intersect in the context of climate injustice. This particular example is still closely connected to the original application of intersectionality by including women of color in the conversation.

There are many additional examples of climate injustice beyond race and gender occurring on an intersectional level. A homeless person during an extreme hot or cold weather event is more at risk for health consequences and death than someone with a roof over their head. Those who have the financial means to rebuild after climate related weather events (i.e. flooding, hurricanes, fires, etc.) are able to recover from a disaster more easily than those who do not have these means. These latter two examples describing socioeconomic resources depend on a variety of other factors impacted by racism, sexism, and so forth. Understanding which populations are more likely to be homeless or have a lower income is an intersectional framing of climate injustice. These disparities can play out at an individual level, community level, or even within an entire country. These examples of inequalities experienced during climate change are the results of preexisting unjust and discriminatory systems that impact several intersecting structural identities.

The purpose of intersectionality is not merely to put a name to an issue and create terminology. Having a new vocabulary to express certain experiences is certainly important for establishing recognition of individuals and groups; having the words to describe an experience is empowering. However, one should not stop there. Crenshaw created intersectionality as a legal tool. While terminology is necessary in order to foster recognition, intersectionality is also a tool that can be used to develop change through legal means. Acknowledging the intersecting discriminatory structures groups face is crucial, but this acknowledgement will have been for nothing if it does not also lead to social, political, and legal change.

In the context of climate justice, using the framework of intersectionality can allow for positive change that includes many different perspectives. Adapting climate policies that consider the intersecting factors of injustice are important for addressing the disproportionate

effects of climate change.[157] Using this framework to understand climate change is relevant in all types of climate related policies. More specifically, governments seek to create legal and political preventative actions by implementing policies that aim to mitigate climate change. Lowering carbon emissions, adjusting to clean/green energy, building sustainable infrastructure and green cities, etc. are all examples of preventative measures that institutions can take to reduce the negative effects of climate change. If climate change is understood to be an intersectional issue, then it is also clear that reducing these negative effects is an ethical, social justice, and human rights issue. Governments can also take measures to respond to disasters and protect people and communities from current and future climate events. Creating infrastructure to support victims of climate disasters, providing medical relief and aid to help those who face these disasters, and other such programs are ways in which governments can support the safety of people around the world. If intersectionality is not taken into account when implementing these responsive policies, certain forms of discrimination and inequality will be overlooked.

In the current discourse around climate change, Pacific Islanders are often cited as an example of 'climate refugees' who must be relocated and rehoused in mainland countries.[158] Those living on such islands contribute very little to the pollution that drives global warming but suffer some of the worst repercussions, including severe weather events, rising sea levels, and sometimes forced relocation. As a result, climate activists and governments from wealthy Western countries believe they must have a role in this relocation of Pacific Islanders, in order to protect these 'climate refugees.' The political and legal changes implemented by Western nations attempt to accomplish climate justice by targeting the issue of disparity in climate consequences. However, these policies and approaches are not intersectional; they only consider the one dimensional problem of unequal impact on more vulnerable populations. While some island residents will have no choice but to relocate, others are situated in island nations that could survive the initial climate disasters, such as Vanuatu.[159] The appropriate solution for these nations would be to ensure access to infrastructure that can support the island and help residents move inland, to higher ground. However, Western nations overlook these more appropriate solutions and still classify islands like Vanuatu as unlivable, with residents requiring relocation.[160] Policies

---

[157] Kaijser and Kronsell, 427.
[158] Nikita Perumal, "'The place I live is where I belong': community perspectives on climate change and climate-related migration in the Pacific island nation of Vanuatu," *Island Studies Journal* 13, no.1 (2018): 46-48.
[159] Perumal, 49.
[160] Perumal, 49.

that aim to relocate and provide safe haven for refugees are being made without the input of native populations. This attempted solution does not use the framework of intersectionality, and the result is a deficient climate justice policy.

If intersectionality were to be applied in this scenario, it would become apparent that Pacific Islanders not only face injustice through the unequal burdens of climate consequences but also face discrimination through the legacy of colonization, and that these two forms of oppression intersect. If intersectionality was used in this context, the political solution would consider the islanders' voices and would create change that did not forcefully displace native populations. Instead, it would help them rebuild, allowing them to remain in their own homes. The non-intersectional approach of many Western countries perpetuates discrimination against Pacific Islanders by forcing relocation; this lack of intersectionality results in further oppression and fails to provide climate justice. An intersectional framework is therefore necessary to create equitable change for all types of policies that must be developed to protect all populations, and vulnerable ones in particular. If intersectionality is not adopted in these other areas, whose voices will be lost? Who will be overlooked?

Intersectionality should be defined with respect to its foundation and origins, but can be expanded to include other forms of discrimination in other contexts. The framework and tool intersectionality provides is highly beneficial and applicable to many social issues and structures. Intersectionality is not only useful in contexts beyond legal discussions of race and gender, but is actually essential to incorporate in these other contexts. Climate justice is an excellent example of the importance of intersectionality in understanding how the issue of climate change impacts people disproportionately around the world, and how climate related legal and political policies are only successful if they create equitable change. However, there are potential dangers to consider when extending intersectionality beyond race and gender. If the origins of intersectionality from Crenshaw's legal work relating to women of color are not acknowledged, intersectionality will be exploited at the expense of those who established it. Women of color must be kept in the conversation of intersectionality and must not be overlooked when addressing other areas of intersecting discrimination. Crenshaw and women of color paved the way for other voices to be uplifted, and this powerful work must not be forgotten. However, once its origins have been fully embraced, intersectionality can become a valuable and beautiful way to advocate social justice in many ways and across many spaces.

Bibliography

Bou-Zeid, Eli and Dan Li, "Synergistic Interactions Between Urban Heat Islands and Heat
        Waves: The Impact in Cities is Larger than the Sum of its Parts." *Journal of Applied
        Meteorology and Climatology* 52, no. 9 (2013): 2051-2064.
        https://www.proquest.com/scholarly-journals/synergistic-interactions-between-urban-heat
        /docview/1445365331/se-2.

Carbado, Devon W. "Colorblind Intersectionality." *Signs* 38, no. 4 (2013): 811-845.
        https://doi.org/10.1086/669666.

Crenshaw, Kimberlé. "Mapping the Margins: Intersectionality, Identity Politics, and Violence
        against Women of Color." *Stanford Law Review* 43, no. 6 (1991): 1241-1299.
        https://www.jstor.org/stable/1229039.

Crenshaw, Kimberlé. "What is Intersectionality?" Interview by National Association of
        Independent Schools. YouTube, June 22, 2018. Video, 1:54.
        https://www.youtube.com/watch?v=ViDtnfQ9FHc.

Hannah-Jones, Nikole. "Our Democracy's Founding Ideals Were False When They Were
        Written. Black Americans Have Fought to Make Them True." *New York Times*, August
        14, 2019. Para 1-57.
        https://www.nytimes.com/interactive/2019/08/14/magazine/black-history-american-demo
        cracy.html

Kaijser, Anna and Kronsell, Annica. "Climate Change Through the Lens of Intersectionality."
        Environmental Politics 23, no. 3 (2013): 417-433.
        https://doi.org/10.1080/09644016.2013.835203.

Perumal, Nikita. "'The Place I Live is Where I Belong': Community Perspectives on Climate
        Change and Climate-Related Migration in the Pacific Island Nation of Vanuatu." *Island
        Studies Journal* 13, no.1 (2018): 45-64. https://doi.org/10.24043/isj.50.

Saporito, Salvatore and Daniel Casey. "Are There Relationships Among Racial Segregation,
        Economic Isolation, and Proximity to Green Space?" *Human Ecology Review* 21, no. 2
        (2015): 113-131,181,183. https://doi.org/10.22459/HER.21.02.2015.06.

The Substance of Experience

Ben Liljedahl

What is the most fundamental question in metaphysics? A few come to mind, but I think a good bet would be: "What is the substance of the universe?" It is an ancient question, but it is just as relevant today as it was thousands of years ago. It may seem like modern science has given a satisfactory answer to this question, but I beg to differ. First of all, metaphysical questions are different kinds of questions than scientific questions. Metaphysics certainly draws upon science, but not every question can be answered through the scientific method. This is especially true concerning the substance or nature of the physical world itself. So when we look to science for an answer to this question, we are really looking to the metaphysical paradigm which scientists accept.

The predominant metaphysical paradigm scientists currently subscribe to is known as 'Material Realism,' the position that all that exists is matter and energy. Essentially, matter is taken to be the substance of the universe. This paradigm, however, is crumbling, and it is crumbling on all sides. While many theorists are now using discoveries in quantum physics to disprove core Materialist concepts like 'local realism,' there is a more basic issue I would like to point out.[161] There's actually no need to bring in quantum physics, because the issue is much more fundamental than any material mechanics. By addressing the more fundamental aspect I believe the problem actually becomes more straightforward. I myself was once a Material Realist, but now have a very different perspective. In this paper I will relay my philosophical journey of discovery in investigating this problem, and the insights which followed.

As the philosopher George Berkeley so astutely pointed out way back in 1710, there is a problem surrounding our whole idea of what matter is.[162] [163] He made his claim to fame by exposing common lapses in logic which lead to a confused understanding of matter. This is a big deal, if we are taking matter to be the very substance of the universe. People tend to operate

---

[161] Howard Wiseman, "Death by Experiment for Local Realism," *Nature* 526, no. 7575 (2015): 649-650. https://doi.org/10.1038/nature15631.

[162] George Berkeley, *A Treatise Concerning the Principles of Human Knowledge* (Indianapolis, Bobbs-Merrill, 1957).

[163] George Berkeley, *Three Dialogues Between Hylas and Philonous*, 1st ed., eds. Michael B. Mathias and Daniel Kolak (New York: Routledge, 2007).

under 'Naïve Realism,' whereby our perceptions are simply taken to be the real external world, end of story. However, are the sensible objects of our perceptions really the external objects themselves? Well, through a series of clever dialogues, Berkeley proved that they can't possibly be.[164] So what we see around us is not the external world, but merely perceptual representations of it. This gets into the question of how we know anything about the world at all, let alone matter.

Forget about the universe for a minute. Let's just start with our stream of experience. If we adopt a classical empiricist stance, then we assume that all knowledge about the universe comes from our experience. (All explicit knowledge, we'll say, to appease the nativists.) Therefore, any knowledge about the substance of the universe must come through our experience. It follows, then, that we must first understand the nature of experience before we are able to understand the nature of the universe. It was from this starting point that I formulated the more specific question: "What is the substance of subjective experience?" This is the question I began my journey with, and it is the question I shall attempt to answer in this paper.

I believe consciousness is the substance of subjective experience. Consciousness? A substance? Sounds crazy, I know. Seriously though, let's consider it. Take a moment to analyze your conscious experience. There is some kind of substance to it, that much is undeniable. Just look around, pay attention to your visual perception. Certainly there's some sort of substance to the images you see, the sensations you feel, and the sounds you hear. So what, exactly, is conscious experience made out of? Maybe it's consciousness itself! I believe it is. In an important sense this is something that is directly intuited, but I will try to present some arguments for it. To clarify, 'consciousness' here simply means awareness. This metaphysical framework is called 'Idealism,' and it's more popular than you may think.

Before going over the philosophical arguments for consciousness as a substance, I want to describe the experience I had that led me to this seemingly ridiculous conclusion. As a cognitive science major, I had thought a lot about that ancient puzzle known as "the mind-body problem." The mind-body problem is the question of what the metaphysical relationship is between mind and body, or between mental properties and physical properties. It was briefly discussed during the beginning of many psychology classes I had taken, and I knew enough about it to know that every good scientist takes the position of Material Realism. That was the

---

[164] Ibid.

impression I got, at least. Other than a desire to conform, I declared myself a Materialist simply because it seemed to be the most logical stance. Usually Materialists set their metaphysics against Descartes' Dualism, a dead horse they love to keep on beating. Metaphysical paradigms like his, called 'Substance Dualism,' maintain that the universe is composed of two fundamentally distinct substances, whereas metaphysics under the alternative category, 'Substance Monism,' maintain that there is only one.

The reason that almost no one endorses Substance Dualism these days is because the paradigm comes with insoluble paradoxes.[165] The interaction problem is the main one: if mental and physical substances are fundamentally distinct, how could they possibly interact? If there is some way they interact, then there is some underlying unity of both, and there must be some laws which describe the interaction. If they never interact, then we would never even know about the other substance. On top of this there is the problem of the conservation of energy. If mental states are of a different substance than physical states, where does the energy causing them go? The brain appears to be a closed system. Energy can't just disappear from your brain, and then reappear later — it violates the laws of physics. Even if that is possible, we have no evidence for it. So Substance Dualism is ruled out by nearly everyone. Thus, most scientists nowadays turn to Material Realism, and declare the debate to be over. I recall one class in which the professor construed the only alternative monism to Materialism as being 'Solipsism.' Solipsism is a specific subtype of Idealism – an obscure, ungrounded, and unpopular one. Instead of taking mind or consciousness in general to be the substance of the universe, Solipsism asserts that your mind in particular is the whole universe. There is a logical form of Solipsism in which one admits that they can only be certain of the existence of their own mind, but when most people use the term they mean the belief that their mind alone is real. By contrasting Materialism against this straw man of a philosophy, it seemed ridiculous to dissent.

Thus, when I first learned about the modern formulation of the mind-body problem, which philosopher David Chalmers termed "The Hard Problem of Consciousness," I naturally came to align myself with the philosophy of Material Realism.[166] Subjective experience, a.k.a. consciousness, must be reducible to the material arrangements of brain activity, and nothing more. I did have some reservations on whether experience could really be identical to brain

---

[165] William G. Lycan, "Philosophy of Mind," in *The Blackwell Companion to Philosophy*, eds. Nicholas Bunnin & E. P. Tsui‑James (Cambridge: Blackwell Publishers, 1996), 173–201.
[166] David J. Chalmers, *The Conscious Mind: In Search of a Fundamental Theory* (Oxford University Press, 1996).

states; I thought that perhaps they were actually distinct and the brain was simply causing or generating the experience. So I had a bit of a Dualist leaning, but in essence I was a Materialist. How, then, did I end up jumping ship and going from Materialist to Idealist?

Well, it happened through a sudden stroke of insight. I was sitting in my desk chair one evening watching a YouTube video and contemplating the nature of the physical world. To preface, I had been exposed to Eastern philosophical ideas before, as I was also a philosophy minor at the time. (I switched to a philosophy major after my insight.) I had read about consciousness-based metaphysics, but never took them that seriously. Ideas like a "self-aware universe" seemed too mystical to be true, so I dismissed them. Nonetheless, I was aware of them. So there I was, sitting and thinking about this hard problem, when all of a sudden my view of the world changed, and I saw things in a new light. Taken by a rush of focus and awe, I felt like I was seeing the world for how it really was for the first time. I became aware that all the perceptions of my room around me were actually part of my mind! I realized that all those objects and colors and lines were literally made of experience – 'mind-stuff.' Of course, a noetic experience such as this is to some degree ineffable, but I'll try my best to put it into words. It dawned on me that the only thing we've ever known in life is our subjective experience, and our stream of experience is just our own mind. This is the whole point Solipsism drives home. It's important to note that 'mind' is being used here in the phenomenal sense, rather than the psychological sense.[167] The phenomenal concept of mind centers on experience or sentience, whereas the psychological concept of mind centers on memory, thought, belief, behavior, etc. This flash of insight included a second key realization. I realized that experience is knowledge. You know what you are experiencing, and you know what it is like to be experiencing it. Our conscious experience is immediately known to us, and is therefore something self-evident. There is no logical reasoning needed to prove that you know what you are experiencing. Later in life I found out that the philosopher Bertrand Russell had the same insight and aptly termed it 'knowledge-by-acquaintance.'[168]

Thus I began to take seriously the idea that consciousness might actually be something fundamental. Fundamental to the universe? Fundamental to reality? I couldn't tell you. But for the first time I saw the logic to Idealist systems of thought. Space-time is something that is

---

[167] Ibid.

[168] Bertrand Russell, "V.—Knowledge by Acquaintance and Knowledge by Description," *Proceedings of the Aristotelian Society* 11, no. 1 (1911): 108–128. https://doi.org/10.1093/aristotelian/11.1.108.

readily accepted as being fundamental to the universe. But consciousness? That seems absurd. Yet after careful consideration I realized that it only seems absurd because the Western tradition of thought endorses the Materialist paradigm, so it goes against most Westerners' psychological conditioning. We tend to view consciousness as being limited to living organisms, or animals, or higher mammals – or even being the sole privilege of humans. The idea that consciousness is intrinsic to the universe might bristle against this anthropocentric attitude. In some cultures around the world, however, Idealism is accepted with the same readiness with which we dismiss it. So, all that being said, let me lay out some arguments I find compelling.

The first reason I regard consciousness as the substance of experience is because all conscious experience is simply consciousness of something. After a careful analysis, I realized that any conscious experience is just consciousness of some form, or equally put, consciousness in some form. All experiences are made of the same substance of consciousness but with different content. This is a realization anyone can have.

What does it mean to be aware of something? Let's break it down. I believe the easiest and most robust examples of experience will come from examining vision, our most salient sense. So what does it mean to be aware of a visual percept? Simply put, it means that the perceptual object is in your visual field. Your visual field, in turn, is in your field of awareness, as it is part of the content of your consciousness. There could be no consciousness in the form of visual experience, if there was no field of consciousness substantiating it. Thus, I came to the understanding that in a very real sense, visual experience is "made of" consciousness. This includes the percepts themselves, meaning that everything you see is quite literally a manifestation of consciousness. So while experience can be thought of as taking place in consciousness, it can be equally conceived of as being made out of consciousness. Awareness is the background of our stream of experience, manifesting it while containing it. Therefore we can say that it is both immanent and transcendent. Perhaps a few metaphors will help. If experience is like a sculpture, consciousness is the clay. If experience is like a tapestry, consciousness is the thread. Equally, if experience is like a painting, consciousness is the paper. If experience is like a movie, consciousness is the screen. If experience is like a play, consciousness is the theater itself. My second point is that experience is dependent upon and inseparable from consciousness. For example, if consciousness were to go away, all conscious experience would go away. We represent the external world through the information taken in by our senses. Without

consciousness you wouldn't have any of these sensory experiences. If in the next second you were to "lose consciousness" as we say, the entire perception of your environment, body, and mind would disappear. There can be no experience without consciousness. This is necessarily so, whether or not there can be consciousness without any object of experience (i.e. pure consciousness events).[169] If all conscious experience is made out of consciousness, then I believe the most appropriate term for it is indeed 'substance.' Any and all experience boils down to some modification of the form of this substance.

Lastly, the concept of experience as 'knowledge-by-acquaintance' entails that there is a fundamental role played by the knower. Knowledge cannot exist without the knower, and that knowing essence is awareness. What is knowledge-by-acquaintance but the awareness of information? For example, what is an experience like sensation but the awareness of some sense-data, a type of information? All experience is knowledge-by-acquaintance, because it's all just awareness of some form. Awareness is that with which all experience is known. It may seem trivial, but if you are directly aware of some particular form, you know that.

What is special about knowledge-by-acquaintance, as opposed to its counterpart, knowledge-by-description, is that it is self-evident knowledge. As philosophers would put it, there is an 'epistemic immediacy' to this type of knowledge. Consider these questions: How do you know you exist? How do you know you are conscious? How do you know what you are experiencing? The answer to all these questions is the same: it is self-evident. There's no way we could convince ourselves that we don't exist, that we're not conscious, or that we're not having experiences. There's also no reason we have to know we are conscious. We are simply aware of it. There's no logical reasoning or sequence of arguments we employ to prove to ourselves that we are conscious. At the time of my insight I interpreted this as consciousness being something non-logical, but you could actually interpret self-evidence as being a certain form of logic.

Descartes did have many brilliant insights, one of which included the indubitability of the existence of one's own mind.[170] You cannot doubt that you exist. It's simply not possible to imagine otherwise. We could even be dreaming right now, Descartes pointed out – it's always a logical possibility.[171] But we still cannot doubt that we are having the dream! The knowledge "I

[169] Robert K. C. Forman, "Pure Consciousness Events and Mysticism," *Sophia* 25, no. 1 (1986): 49-58. https://doi.org/10.1007/BF02789849.

[170] Réné Descartes, *Meditations on First Philosophy*, trans. Michael Moriarty, Oxford World's Classics, (London: Oxford University Press, 2008).

[171] Ibid.

am" is the fundamental certainty. Experience is simply the certainty that "I am in this particular form." "I am" is awareness' knowledge of itself. It is the foundational premise upon which all knowledge is based. This is why awareness can be characterized as the knowing essence. If awareness – the knower – is the substance of knowledge-by-acquaintance, and experience just is knowledge-by-acquaintance, then lo and behold, awareness must be the substance of experience. Although you may not be familiar with it, this perspective on consciousness is actually nothing new! It's the other side of the mind-matter debate: the metaphysical philosophy known as Monistic Idealism. There have been countless Idealist philosophers throughout ancient and modern history, from Plato to Berkeley to Donald Hoffman. Hoffman is a contemporary cognitive scientist gaining attention who has recently put forth a mathematically precise Idealist paradigm called 'Conscious Realism.'[172] I think it's an excellent paradigm that merits serious consideration. But if there's one school of thought that really takes the cake for Idealist paradigms throughout human history, it's got to be the ancient Vedic philosophy found in Hinduism.

There are many sub-categories of Vedic philosophy, with lots of technical differences. The devil is in the details. What matters is that all forms of Vedic philosophy acknowledge that Brahman is the substance of reality. "What is Brahman?" you may ask. Well, "His true nature is pure consciousness," states the Aitareya Upanishad.[173] Pure consciousness is said to transcend the waking, dreaming, and sleeping states of consciousness. Everything is made out of Brahman, and Brahman, being the substance of reality itself, depends on nothing to exist. Brahman is also referred to as the true Self, because in Vedic philosophy the metaphysical Self is awareness. This is because there is no awareness without the one who is aware – they are inseparable.

Awareness is most sensibly what puts the "subject" in "subjective experience." It entails that there is one who is aware, just as experience entails an inseparable one-who-is-experiencing. If awareness itself is the subject of experience then it is this awareness/subject who is the substance of conscious/subjective experience. Indian thought and culture has embraced Idealism for thousands of years, so it's not like humankind has never known this. It's just that the modern scientific establishment has been somewhat blind to it.

---

[172] Donald Hoffman, "Conscious Realism and the Mind-Body Problem," *Mind and Matter* 6, no. 1 (2008): 87-121. https://philpapers.org/rec/HOFCRA.

[173] Swami Prabhavananda and Frederick Manchester, trans., *The Upanishads: Breath of the Eternal* (New York: New American Library, 1964).

Nonetheless, the Idealist perspective is more widespread throughout Western thought than many people realize. Take Max Planck, for instance – the founder of quantum theory. He is famously quoted as saying "I regard consciousness as fundamental. I regard matter as derivative from consciousness. We cannot get behind consciousness."[174] It would be unwise to dismiss the conclusions of such a brilliant mind, one which made so many important contributions to physics. Instead we should wonder why he believed them. The great polymath Sir Arthur Eddington is another example. He proclaimed that "The stuff of the world is mind-stuff … The mind-stuff of the world is, of course, something more general than our individual conscious minds ... The mind-stuff is not spread in space and time; these are part of the cyclic scheme ultimately derived out of it ... It is necessary to keep reminding ourselves that all knowledge of our environment from which the world of physics is constructed, has entered in the form of messages transmitted along the nerves to the seat of consciousness ... It is difficult for the matter-of-fact physicist to accept the view that the substratum of everything is of mental character. But no one can deny that mind is the first and most direct thing in our experience, and all else is remote inference."[175] I couldn't have said it better myself! The very best defense of Idealism I have come across, however, is philosopher Bernardo Kastrup's 'Analytic Idealism,' as defended in his book The Idea of the World.[176] His logical precision is unparalleled, and I suspect his popularity will only increase with time.

If we accept that consciousness is a substance, we are forced to re-evaluate paradigms such as Material Realism. If we have one substance, what grounds do we have to posit another? Occam's razor suggests that consciousness is all there is. Such a claim may be charged with denying the reality of the entities which our best scientific theories assert to exist. However, I do not see any reason that scientific realism has to be limited to Material Realism. Scientific realism is simply the belief in both the observable and theoretical aspects of the world as described by science. Since physics is the most fundamental science, this is generally equated with the metaphysical paradigm called 'Physicalism.' This stance is similar to Materialism, but is less centered on matter as substance. It merely asserts that everything in the universe is 'physical.' I actually believe that Physicalism, broadly construed, will hold true. If there is a way something

---

[174] Max Planck, interview by John W. N. Sullivan, *The Observer*, January 25, 1931: 17, col. 3.
[175] Arthur S. Eddington, *The Nature of the Physical World* (London: J. M. Dent & Sons Ltd., 1935).
[176] Bernardo Kastrup, *The Idea of the World: A Multi-Disciplinary Argument for the Mental Nature of Reality* (Winchester: Iff Books, 2019).

looks and a way that it works, it can be described, modeled, and predicted. Physics is just our way of representing the structure and laws of the universe. It is a specific kind of knowledge-by-description. Even non-deterministic phenomena like those in quantum mechanics can be described probabilistically. Every subjective experience presumably has some physical 'token' that can be represented in objective physical terms. Yet physics alone is not designed to answer such questions as: "What is the substance of the physical world?" or "What is it like to be a physical system?" Such questions are simply unable to be answered in physical terms.

Again, I'm not arguing against Physicalism per se, but rather the orthodox 'Material Realist' conceptualization of it, whereby the substance of everything is taken to be matter. I still believe that everything in the universe can be described by physics and adheres to physical laws, so I believe in the spirit of Physicalism. Everything, even matter, is said to be made of what physicists term 'energy.' So what exactly is energy? Could its intrinsic nature be phenomenal? In other words: what if the universe physics describes is one made of consciousness? I share the same sentiment towards my conclusions that David Chalmers expresses towards his own conclusions in his book *The Conscious Mind*:

> By now, I have grown almost happy with these conclusions. They do not seem to have any fearsome consequences, and they allow a way of thinking and theorizing about consciousness that seems more satisfactory in almost every way. And the expansion in the scientific worldview has had a positive effect, at least for me: it has made the universe seem a more interesting place. [177]

---

[177] Chalmers, xiv.

Bibliography

Berkeley, G. *A Treatise Concerning the Principles of Human Knowledge*. Indianapolis, Bobbs-Merrill, 1957.

Berkeley, G. *Three Dialogues Between Hylas and Philonous*. 1st ed. Edited by Michael B. Mathias and Daniel Kolak. New York: Routledge, 2007.

Chalmers, David J. *The Conscious Mind: In Search of a Fundamental Theory*. Oxford University Press, 1996.

Descartes, Réné. *Meditations on First Philosophy*. Translated by Michael Moriarty. Oxford World's Classics. London: Oxford University Press, 2008.

Eddington, Arthur S. *The Nature of the Physical World*. London: J. M. Dent & Sons Ltd., 1935.

Forman, Robert K. C. "Pure Consciousness Events and Mysticism." *Sophia* 25, no. 1 (1986): 49-58. https://doi.org/10.1007/BF02789849.

Hoffman, Donald. "Conscious Realism and the Mind-Body Problem," *Mind and Matter* 6, no. 1 (2008): 87-121. https://philpapers.org/rec/HOFCRA.

Kastrup, Bernardo. *The Idea of the World: A Multi-Disciplinary Argument for the Mental Nature of Reality*. Winchester: Iff Books, 2019.

Lycan, William G. "Philosophy of Mind." In *The Blackwell Companion to Philosophy*, edited by Nicholas Bunnin & E. P. Tsui‑James, 173–201. Cambridge: Blackwell Publishers, 1996.

Planck, Max. *The Observer*, by John W. N. Sullivan, January 25, 1931: 17.

Prabhavananda, Swami, and Frederick Manchester, translators. *The Upanishads: Breath of the Eternal*. New York: New American Library, 1964.

Russell, Bertrand. "V.—Knowledge by Acquaintance and Knowledge by Description." *Proceedings of the Aristotelian Society* 11, no. 1 (1911): 108–128. https://doi.org/10.1093/aristotelian/11.1.108.

Wiseman, H. "Death by Experiment for Local Realism." *Nature* 526, no. 7575 (2015): 649-650. https://doi.org/10.1038/nature15631.

Free Your Mind, and (Some of) the Rest Will Follow:
How to Live an Intentional Life in a Deterministic World

Anilā Lacroix

While many deeply cherish the idea of free will and think that the ability to choose what happens to one's own life is what makes life worth living, some suspect that we are simply part of a long chain of events that determined what we would be precisely at any moment long before we came to exist. I argue that there is an alternative pragmatic approach that borrows from both perspectives. Given the ongoing philosophical debate regarding whether human beings are fundamentally predetermined or free to choose their lives at will, I believe it is important to specify *why* free will matters in every person's life. Studies show that the ability to deliberate and decide one's life path is vital to mental health; individuals who believe they have no freedom to act from their intrinsic motivations fail to thrive and can become diminutive versions of themselves in all aspects of their lives.[178] For these reasons, it is essential to establish a variable kind of free will. In the following, I will demonstrate how we can achieve varying levels of free will and discuss implementing mindfulness practice to bolster free will potentials.

Determinism states that all events have prior causes and that everything in the past has fixed the causal pathway for everything in the present and the future.[179] Accordingly, determinism posits that our current behaviours and actions are the end result of all preceding moments in history. If one believes free will is the ability to live one's life with self-causation, it is commonly expected that one must have the complete freedom to generate and decide one's personal paths and desires. Most people expect to be, and believe they are, the cause of their own destiny. On face value, then, determinism denies such freedom and puts us on causally predestined paths; it follows that if all events I believe myself to be deciding and causing were actually fixed by everything that came before, then I am merely acting out my role in the cosmic

---

[178] N. Adams, T.D. Little, and R.M. Ryan, "Self-Determination Theory," in *Development of Self-Determination Through the Life-Course*, eds. M. Wehmeyer, K. Shogren, T. Little, and S. Lopez (Dordrecht: Springer, 2017).
[179] M. McKenna and J. Coates, "Compatibilism," in *The Stanford Encyclopedia of Philosophy* (The Metaphysics Research Lab, Stanford University, 2021).

deterministic play without any say in the matter.[180] Depressing days indeed. It is prudent, then, to better define determinism in order to demonstrate that it is possible to have some self-determination and free will within a causally determined world.

Daniel Dennett is a renowned cognitive scientist and philosopher of mind who offers a middle ground approach to the free will vs. determinism debate which he calls compatibilism. Dennett maintains that despite events being deterministically set in motion (by other events that were set in motion, and so on), we can observe and change some courses of action. While he agrees that events are causally determined, Dennett argues that causation and control are two different concepts. This kind of compatibilism proposes that despite living in a causal world and experiencing causally generated events, the existence of control and individual agency is still possible through limited free will. Limited free will, henceforth designated as "free will", is the ability to be a self-directed agent within the bounds of one's abilities and the environment they are in at that given moment.

As a free will agent, one can be aware of one's circumstances and deliberately change some of them. Such free will is made possible by our capacity to live in various states of awareness combined with self-control and the ability to move between these variable states according to our desires.[181] Since one understands the difference between being half-aware and being highly alert in their levels of cognitive awareness, one is likely able to identify other variances in levels of awareness. Most of us also understand that with some self-control, we can force ourselves to pay attention or ignore something when we decide to. Using this foundational skill, we can be alert to our environment and sometimes deliberately make changes therein. Therefore, I argue that we can accept Dennett's compatibilism because it offers a pragmatic assessment of our world, parts of which science has shown to be absolutely causally deterministic. But at the same time, we can recognize that many of us have *self*-determined some of our thoughts and behaviours in order to change the paths we were on.

The difference between causation and control is expanded by knowing what we can and cannot deliberately change in a deterministic world. This knowledge creates varying spaces for free will meaning that our actions and the factors of our lives may be determined but not inevitable. Some of us will go our entire lives without questioning the path their birth put us on.

---

[180] R. Repetti, *Buddhism, Meditation, and Free Will: A Theory of Mental Freedom* (Taylor & Francis Group, 2019).

[181] D. Dennett, *Elbow Room: The Varieties of Free Will Worth Wanting* (MIT Press, 2015).

Others will evaluate every aspect of their lives to decide if this is the way they want to live or if they want to change something, unknowingly demonstrating the knowledge-seeking needed to act with some free will. Dennett argues that if we are interested in having the kind of free will agency where we can live with some level of autonomy, we must become more aware and deliberate self-controllers. Self-control may therefore be defined as an individual's capacity for deliberately having an effect on deterministic external states according to one's internally motivated decisions.

Dennett recognizes that the two greatest limiting factors of self-control and self-determination are the inability to anticipate the occurrence of a variable and the inability to respond as desired within our range of capabilities.[182] Most of us are already familiar with this in some way, namely, in how we prepare for the day ahead, for where we are going and for what we expect will happen. A simple example: when we have prepared poorly for the day, we have put ourselves onto a path with fewer options than if we had planned for more possibilities. One usually does not get far without planning for transportation and for currency to pay for their plans or any unforeseen events. Dennett's approach to expanding against the limitations on our free will highlights the value and importance of planning ahead in as many aspects of our lives as possible.[183]

As agents wanting to remain in self-control with the ability to decide our lives freely, we ought to endeavour to our best abilities to forecast our yet unknown environment and the events that could transpire. Free will expands with our ability to anticipate, perceive and remain aware of the variables in a given situation. And the greater our awareness is, the greater our capacity to see probable futures in time to wilfully take steps to adapt ourselves to the event or change the variables of the event into something else if we desire to. Commonly, a self-sufficient person likes to "know what they are walking into". Limited free will uses this kind of awareness and shows that with better forecasting and planning, one has more opportunities to behave as one wants, regardless of the familiarity of the situation.

Further, the amount of free will one possesses is correlated with their knowledge of their present and future environment and their capacity to wilfully act armed with such knowledge. In order to expand their free will, one must know their abilities to respond as well as know their

---

[182] Dennett, *Elbow Room.*
[183] Dennett, *Elbow Room.*

environment in order to ascertain factually what aspects of these are changeable. Knowledge of one's abilities comes from testing themselves in various environments and consciously learning from these experiences to better refine their self-control. The correlation between possible free will and knowledge-informed actions indicates that the more awareness and ability we have, then the more free will is available and the more adaptable we can be in any given moment.

Earlier in this paper, it was stated that free will expands with our ability to perceive and remain conscious of all variables in a given situation. In this awareness, we find the imperative to evaluate ourselves more thoroughly as active factors. Where we are aware of our environments, abilities and skills, I suggest expanding and improving our ability to assess our internalised way of evaluating. I propose that the practice of mindfulness can increase free will possibilities by enhancing self-control. For the purposes of this paper, mindfulness practice is defined as intentional self-awareness and conscious self-refinement through analytical and compassionate[184] observation of one's present mind state. Dennett's compatibilism and this understanding of mindfulness are in alignment: they both propose that some things are unchangeable while some are not and our task is to discern for ourselves which is which. Consider the following example as a starting point for understanding compatibilism from the mindfulness perspective. We all must breathe; this is a causal condition. We all can choose to hold our breath or change our breathing to affect the nervous system; this is a compatibilist/mindful condition.

Mindfulness teachings also align with compatibilism in acknowledging that we cannot escape natural causal processes nor patterns we have instigated which affect our lives in direct ways, but we can act to gradually change or escape certain aspects of these patterns.[185] It is simply a fact that we cannot change the effect of the Earth's gravity on an object. But with knowledge and mindfulness, we can manipulate other variables to accommodate this fact. This means that despite the invariable effect of gravity, we adapt ourselves to situations based on our knowledge of how gravity is expected to work. To make this clearer with an example, if you set a process in motion, for instance, cliff jumping with your friends, once you jump, you are likely not escaping the water you are falling into. You can, however, pause to deliberate whether or not

---

[184] Having compassion within mindfulness is supportive of acting within reason. Without it, one could easily slip into obsessive-compulsive thoughts that increase limitations to free will. Mindfulness teachers often remind the student to be nurturing and self-kind to the human within the practice and to remember that we seek progress, not perfection. Compassion within mindfulness extends outwards so that evaluations of others are factual, but also more kind and generous.

[185] N. Latham, "Meditation and Self-Control," *Philosophical Studies* 173, no. 7 (2016): 1779–1798.

to jump or to jump with a parachute. Mindfulness gives the opportunity to examine why and how you are acting on those choices both in relation to your inner world and to your perception of the external world. The unfolding of a process that one decides to put in motion (or not) is less deterministically fixed when subject to self-control.[186]

Previously, I defined self-control as an individual's capacity for having an effect on external states according to internally motivated decisions. This definition should now be revised to be an individual's capacity for mindfully having a causal effect on external states according to one's internally motivated decisions. A mindful, self-controlled agent can identify and challenge their personal, genetic and cultural givens by testing different behaviours against their habits, "going against the grain," and investigating other ways of approaching a situation. To give a personal example,[187] I observed long ago that meditating in the morning, as prescribed by all the different schools I had learned from, was not having the intended effect. I persevered, stubbornly forcing the traditional morning practice because I thoroughly believed that was the only way to a clear mind. Despite my sincere belief that I was no longer practising actual meditation, I set out to find a different way and found a text that said, "Or, by meditation as desired".[188] I started testing different times of day and empirically observing my physiological responses. From this testing, I found more profound control and the results I sought by meditating at night, thus changing to a new pattern. Sharing this with a master teacher, they said the key was that our actions must arise from self-evaluation so we can choose the correct course of action and continuously repeat that cycle of self-evaluation and resulting self-action. Doing so makes us more responsive to the moment and less deterministically stuck in an unchosen behavioural pattern. Thus, an important step towards increasing mindfulness is to combine self-evaluation with challenging more significant concepts of what is known and testing different approaches.

One can successfully build mindfulness skills through intellectual curiosity and an openness to examining all presumed knowledge. If one is willing to consider that any preconception might be a misconception, one becomes more open to truths one would not have

---

[186] Dennett, *Elbow Room*.

[187] It is important to note that although I am using my meditation practice as an example of mindfully and wilfully challenging what one believes to be fundamental and commonly accepted truths, I am not arguing that it is necessary to meditate to be mindful. I use this example because it demonstrates how I had to force myself to challenge what I thought was the status quo of proper behaviour. By challenging presumed truths, I found a correct and reliable perception that I could use for better self-control and free will.

[188] Patañjali, "Sutra 39," in *The Yoga Sutras of Patanjali. A Study Guide for Book I,* ed. and trans. B. H. Dass (Sri Ram Publishing, 1999), 99-100.

otherwise known. Through such dispassionate mindfulness, one can best gain knowledge of their mind state which amplifies their ability to critically evaluate, identify and change ingrained or habitual thought and behaviours.[189] To refer back to the cliff jumping example, you could use mindfulness to examine the root of the decision to jump. If your impulsive response is that people always expect you to do what scares you, the opportunity for more free will is found in mindfully evaluating whether that is a true statement, or whether it is a conditioned response which then influences your presumed correct assessment of the situation. A sustained mindfulness practice can weed out many impulsive or not deliberately chosen thought patterns and their resulting perspectives and behaviours.[190]

Recall the earlier statement about the positive feedback loop wherein the more one can expand their knowledge-informed actions in the world, the more they can expand their ability to act freely. Similarly, the more one's knowledge of their mind increases through self-analysis and refinement, the more one can gain true self-control. Through a compatibilism that is amplified by mindfulness, each of us has the potential to better know ourselves and better assess our many environments. If we can genuinely discern truth from misconception, then our ability to assess and prepare ourselves for any environment is thoroughly enhanced. Thus, one is best prepared to act with greater freedom within a causal world through mindful self-knowledge and self-control.

As stated, mindfulness can inform and expand free will in spite of a deterministic world. Therefore, we should exercise our capacity to be mindful and examine the challenges we must overcome in order to realistically assess our environment and act with deliberation. I stated earlier that potential challenges to one's personally chosen action arise if one's cognitions and behaviours are habitually enacted without mindful self-reflection or analysis. Take, for example, someone unconsciously biting their nails or impulsively saying something without being aware of what they are saying or why. While these are simple psychological examples of habituated actions, they can be used to direct our awareness to the need to question whether our other thoughts and actions are habitually predetermined. Whether we are conditioned by our families

---

[189] J. H. Austin, *Zen and the Brain: Toward an Understanding of Meditation and Consciousness* (MIT Press, 1998), 125-129.

[190] It is notable, but not under the scope of this paper, that studies have demonstrated that the more one practises mindfulness, the more their neuronal capacity and activity increases to the chosen task. A brain's grey matter expands in the area of use, so a mindful person's self-control efforts grow the areas involved with mindfulness. This indicates that through self-control, one can change some aspects of their brain matter's causal state (Hölzel et al., 2011; Repetti, 2019).

or cultures to think or behave in particular ways or are genetically predisposed to specific character traits, most of us exhibit beliefs and actions that are not necessarily of our choosing. I believe people spend much of their days in some version of this conditioned, "auto-pilot" existence. Although some of our habituation may be evolutionarily and neurologically necessary,[191] we need not spend our entire lives in such automated, predictable and highly limited ways.[192]

It is easy to imagine that anyone who goes through their days on "auto-pilot" is relatively predictable. Any task done on an unchanging schedule, any opinion or perspective that remains internally unchallenged and any habit that continues unchecked can reflect a habitually determined causal pattern in one's life. It is common to hear someone ask a friend, "Why do you do that?" to which the answer is, "I don't know, I've just always done it this way." We each have automated, patterned reactions to situations, "pre-thoughts" that come from a "pre-conscious" part of the mind.[193] Any of these unchallenged causal behaviours result in our being predictable beings with little access to free will. If one lives their daily life without reflecting on any deeper considerations of why they think or act in any of the ways they do, they are unable to identify any area that is truly chosen by their free will (and likely would not think to try). Without mindfully challenging how one moves through life, that person's life is a reflection of the conditions that formed it and an echo of the perspectives that informed it. If one's life is only an environmental reflection and an echo of one's influence, then there is no personally gained insight and no free will.

Recall Dennett's argument that limited free will is made possible by our capacity to live in various states of awareness combined with self-control and the ability to move between these variable states according to our desires.[194] The key to unlocking more free will in our lives is to evaluate which thoughts and behaviours we want to experience deliberately and with full or fuller awareness. Mindfulness teachings state that one must challenge even the most mundane

---

[191] Although not in the scope of this paper, I felt it was important to acknowledge that the brain's ability to learn and process is highly reliant on habituation, for good reason. Through habituation, the brain systems are not overloaded with information. A brain with functional habituation deficits begins to have cognitive function failures throughout, potentially affecting any of the central nervous system. Nonetheless, we need not be entirely automated.

[192] L. E. R. Blok et al., "Genetics, Molecular Control and Clinical Relevance of Habituation Learning," *Neuroscience & Biobehavioral Reviews,* 143.

[193] Repetti, *Buddhism, Meditation, and Free Will,* 2019.

[194] Dennett, *Elbow Room.*

behaviours to awaken the mind to higher levels of awareness.[195] Common mindfulness questions meant to unlock more conscious cognition include: "When was the last time you brushed your teeth differently and with awareness?" or "When you came to class, did you notice the features of any of the people around you?"[196] In creating more awareness, one can be more dynamic in their choices. Upon finding other undesirable unconscious habits, one could choose to test and, based on the results, edit their conditioning to better support what they deem important. Through conscious discrimination against habitual assumptions, behaviours and perceptions, one continues to expand their free will.[197] One's free will potential is entirely shaped by their mindful ability to properly and thoroughly assess all factors of the situation including their own internal processes and by using their self-control to act or not act in some way. Dennett's compatibilism is enhanced through dedicated mindfulness. For those of us concerned with the debate of whether we have free will or live in a deterministic world, I think the playful answer is yes to both.

We began this discussion recognizing that the consideration of and ability to have free will is important because individuals who believe they have no freedom to act from their intrinsic motivations fail to thrive and can become diminutive versions of themselves in all aspects of their lives. I acknowledge that there are a great many people who are getting on with their lives without problem as they have never considered these concerns. Whether they are on "auto-pilot" or not, I believe the saying "Live and let live" aptly applies here. For anyone concerned about potentially living in a deterministic world, I propose that the pragmatic approach is to challenge the boundaries of what you think you know, stay curious and explore as much as you can of your own deterministic ways and of those in your different environments. When one considers the scientific approach to gaining knowledge, we understand that we must observe and state all that is known of the object being studied; we must be included in the objects we are studying. I have demonstrated that self-control, informed by mindfulness, is our best tool for existing wilfully in some parts of our deterministic world. By applying empirical observation

---

[195] While it is common for mindful meditators to have heightened self-awareness, I am not suggesting that everyone instantly awakens to observing all details of every thought and behaviour they experience in response to every external experience. However, this approach illustrates that increasing self-evaluation is critical to expanding mindfulness and thus, free will.

[196] Austin, *Zen and the Brain*, 125-129.

[197] Patañjali, "Sutra 28," in *The Yoga Sutras of Patanjali. A Study Guide for Book II,* ed. and trans. B. H. Dass (Sri Ram Publishing, 1999), 113-117.

of ourselves, our habits and our patterns as well as testing ourselves against the moment, we can continue to gain and expand a limited free will. If we do not practise mindfulness, we may live our entire lives as a reflection of the templates each of us was born into and conditioned by. So my friends, free your mind, and some of the rest will follow. I wish you well on your journey.

Bibliography

Adams, N., T.D. Little and R.M. Ryan. "Self-Determination Theory." In *Development of Self-Determination Through the Life-Course*, edited by M. Wehmeyer, K. Shogren, T. Little, and S. Lopez. Dordrecht: Springer, 2017: 47-54. https://doi.org/10.1007/978-94-024-1042-6.

Austin, J. H. *Zen and the Brain: Toward an Understanding of Meditation and Consciousness*, 125-129. MIT Press, 1998.

Bayne, T. "Self-consciousness and the Unity of Consciousness." *The Monist* 87, no. 2 (2004): 219-236. http://www.jstor.org/stable/27903867.

Blok, L. E. R., Boon, M., van Reijmersdal, B., Höffler, K. D., Fenckova, M., and Schenck, A. "Genetics, Molecular Control and Clinical Relevance of Habituation Learning." *Neuroscience & Biobehavioral Reviews* 143 (2022): 104883. https://doi.org/10.1016/j.neubiorev.2022.104883.

Deci, E. L. and R. M. Ryan. "The 'What' and 'Why' of Goal Pursuits: Human Needs and the Self-Determination of Behavior." *Psychological Inquiry* 11, no. 4 (2000): 227–268. https://doi.org/10.1207/S15327965PLI1104_01.

Dennett, D. *Elbow Room, The Varieties of Free Will Worth Wanting*. MIT Press, 2015.

Hölzel, B. K., Carmody, J., Vangel, M., Congleton, C., Yerramsetti, S. M., Gard, T., and Lazar, S. W. "Mindfulness Practice Leads to Increases in Regional Brain Gray Matter Density." *Psychiatry Research: Neuroimaging* 191, no. 1 (2011): 36–43. https://doi.org/10.1016/j.pscychresns.2010.08.006.

Latham, N. "Meditation and Self-Control." *Philosophical Studies* 173, no. 7 (2016): 1779–1798.

McKenna, M., and J. Coates. "Compatibilism." In *The Stanford Encyclopedia of Philosophy.* Stanford University, 2021. https://plato.stanford.edu/entries/compatibilism/

Patañjali. "Sutra 39." In *The Yoga Sutras of Patanjali*. *A Study Guide for Book I,* edited and translated by B.H. Dass, 99-100. Sri Ram Publishing.

Patañjali. "Sutra 28." In *The Yoga Sutras of Patanjali*. *A Study Guide for Book II,* edited and translated by B.H. Dass, 113-117. Sri Ram Publishing.

Repetti, R. *Buddhism, Meditation, and Free Will: A Theory of Mental Freedom*. Taylor & Francis Group, 2019.

Reawakening Bayesian Dilemmas: The Role of Assumption in the Sleeping Beauty Problem

Joshua Levy

Introduction

Probability theory is a cornerstone of scientific reasoning and decision-making, and Bayesian probability stands out as one of its most influential branches. Bayesian probability theory rests on the notion that rational agents have degrees of belief, aptly termed "subjective probabilities" or "credences," which can be represented mathematically. This framework asserts that credences should comply with the laws of probability: a rational agent should, for instance, maintain that the probability of any given event lies between 0 and 1, and the sum of the probability of an event and the probability of its negation equals 1.

This paper delves into a captivating paradox known as the Sleeping Beauty problem—a puzzle that has sparked vibrant debates within the field of probability theory. The crux of the issue lies in a schism between two philosophical camps—the Halfers and Thirders—each proposing disparate levels of credence for the same event within an apparently coherent logical framework. This divergence is a testament to the nuanced intricacies of probabilistic reasoning, pointing towards deeper undercurrents within the foundations of Bayesian inference. As we examine the contrasting viewpoints of the Halfers and Thirders, our objective is to disentangle the fundamental presumptions embedded within their differing interpretations.

The complexities revealed in this discourse demonstrate that the Sleeping Beauty problem serves not merely as a philosophical puzzle but as a catalyst for evaluating the boundaries of Bayesian inference. If both sides act in accordance with the conventional framework and come to mutually exclusive results, the scope of the model must come into question. The Sleeping Beauty problem exemplifies an epistemic state that cannot be accurately

captured by any single, precise probabilistic credence function, warranting the exploration of alternative frameworks capable of reconciling the complexities of context-sensitive uncertainty without such burdensome commitments.

The Sleeping Beauty Problem

To understand the Sleeping Beauty problem, consider the following experiment: a rational agent named Sleeping Beauty learns that she will participate in an experiment in which she will be put to sleep and given a drug capable of erasing her memory. Before being put to sleep, Sleeping Beauty is fully informed of the following details. On Monday morning, the research team will wake her, and she will be unaware of whether it is Monday or Tuesday. Subsequently, she will be informed that it is Monday, her memory will be erased, and she will be put back to sleep. Whether or not she will be awakened again on Tuesday will depend on the outcome of a fair coin toss. If the coin lands tails, she will be awakened again on Tuesday before returning to sleep. However, if the coin lands heads, there will be no second awakening. On Wednesday, the experiment will conclude, and Sleeping Beauty will be informed of the outcome of the coin toss.

Upon awakening with no recall of any previous awakening or evidence to suggest the day, she faces three potential realities:

H_1: The coin will land heads, and it is Monday;

T_1: The coin will land tails, and it is Monday;

T_2: The coin landed tails, and it is Tuesday.

Notably, from Sleeping Beauty's perspective, these three states are indistinguishable.

The central question of the Sleeping Beauty problem is this—upon awakening, what subjective probability or credence should Sleeping Beauty assign to a coin toss outcome of heads? Thirders argue that Sleeping Beauty should assign a credence of ⅓ to the outcome of heads because there are three possibilities ($H_1 \lor T_1 \lor T_2$), each with an equal likelihood, and only one involves an outcome of heads. Halfers contend that a credence of ½ is appropriate

because Sleeping Beauty lacks any new evidence to update her Sunday night credence, which is merely speculating the outcome of a fair coin toss.

Following Arnold Zuboff's introduction of the Sleeping Beauty problem in the mid-1980s, two champions arose to defend their respective positions: Adam Elga–a Thirder, and David Lewis–a Halfer[198]. With no intent to signal a preference for his position, this paper uses notations from David Lewis to represent Sleeping Beauty's credence functions at different times:

P_-: her credence on Sunday evening after receiving a complete explanation of the study;

P: her credence upon awakening;

P_+: her credence after being informed that it is Monday.[199]

Understanding the distinction between Sleeping Beauty's credence functions in each of these states is crucial to unraveling the debate and comprehending why this problem sparks such intense controversy.

Mathematical Common Ground

Before analyzing the nature of the split between these two stances, it is essential to recognize the assumptions that Thirders and Halfers share in their approach to this problem. Both sides fundamentally implement a Bayesian model, which rests upon a set of foundational principles which provide the necessary framework for updating beliefs, computing probabilities, and making rational decisions. One such principle is conditionalization, which lies at the heart of Bayesian inference's adaptability. It stipulates that an agent's credence in a proposition at a given time should be based on their prior credence, conditioned on all the relevant information acquired in the intervening period. Mathematically, conditionalization is expressed through Bayes' Theorem, which quantifies how rational agents should update their beliefs in light of new evidence. Bayes' Theorem states that for the events A and B, the probability of A given B can be found using the following equation:

---

[198]Arnold Zuboff, "One Self: The Logic of Experience," *Inquiry: An Interdisciplinary Journal of Philosophy* 33, no. 1 (March 1990): 39-68.
[199] David Lewis, "Sleeping Beauty: Reply to Elga," *Analysis* 61, no. 3 (July 2001): 171-176.

$$P(A \mid B) \; = \; \frac{P(B \mid A) \cdot P(A)}{P(B)} \tag{3.1}$$

Another essential principle of probability theory which guides Bayesian inference is the Law of Total Probability, which provides a method for calculating the probability of an event by considering all possible ways in which that event can occur. This principle enables comprehensive probabilistic analysis by accounting for all relevant factors and pathways contributing to an event's occurrence.

Let B_1, B_2, ...B_n be events that satisfy:

$$P(\cup_{i=1}^{n} B_i) \; = \; 1 \quad and \quad P(B_i \cap B_j) \; = \; 0 \quad for \; i \neq j. \tag{3.2A}$$

Then, for every event A,

$$P(A) \; = \; \sum_{j=1}^{n} P(A \mid_{B_j}) P(B_j) \tag{3.2B}$$

In their respective approaches to the Sleeping Beauty problem, Lewis and Elga both agree that there are only two classes of worlds in which credentials may be ascribed.[200,201] These are the centered world, which consists of a possible world as well as a perspective, or center, within it (i.e., an agent within that world at a particular time in that world), and the uncentered world, which consists of a possible world untethered to an observer. In the context of the Sleeping Beauty problem, centered worlds are those which consider the perspective of Sleeping Beauty at a given moment—namely, when she awakens and is ignorant of the day and the outcome of the coin toss. These centered worlds capture not just the facts about the universe in that moment (e.g., whether the coin landed heads or tails and whether it is Monday or Tuesday) but also the specific viewpoint of Sleeping Beauty. Her perspective is crucial because it embodies her lack of temporal information, her credal function, and any other subjective elements that impact her probabilistic reasoning. Uncentered worlds, on the other hand, consist merely of the possible

---

[200] Lewis, "Sleeping Beauty: Reply to Elga," 171-176.
[201] Adam Elga, "Self-Locating Belief and the Sleeping Beauty Problem," *Analysis* 60, no. 2 (December 2000): 143-147.

worlds without considering any observer's perspective within that world. These are the scenarios which can be discussed in objective terms—where Sleeping Beauty's beliefs (or lack thereof) play no role.

Both sides also accept a highly restricted Principle of Indifference, which states that in the absence of any relevant evidence, agents should distribute their credence equally among all the possible outcomes of a system. Issues arise when deciding which events in this experiment one should apply this principle to specifically, but as a general rule, each side's acceptance of this principle establishes further common ground.

Propositional Common Ground

From their shared mathematical background, both sides further agree upon certain propositions inherent to the Sleeping Beauty problem. Firstly, it is essential to note that T_1 and T_2 are positioned within the same uncentered world, with their only difference being the temporal location of the agent. For this reason, both sides agree:

$$P(T_1) = P(T_2). \tag{4.1}$$

Further, if the coin outcome is heads, Sleeping Beauty will only be awakened on Monday:

$$P(\text{heads}) = P(H_1), \tag{4.2}$$
$$P(\text{tails}) = P(T_1 \lor T_2).$$

Both sides take that the coin is fair, and thus P_- must be $\frac{1}{2}$ for both heads and tails:

$$P_-(\text{heads}) = \frac{1}{2} = P_-(\text{tails}). \tag{4.3}$$

It is also agreed upon that between the times at which she has credences P and P+, Sleeping Beauty gains evidence, specifically that it is Monday ($H_1 \lor T_1$). Despite their different

rationales, both sides agree that this information is relevant to an outcome of heads because it eliminates the possibility of it being Tuesday from the equation. From the proofs given by each side in the following sections, it is agreed that conditioning upon this new information should increase Sleeping Beauty's credence in heads by $\frac{1}{6}$:

$$P+(\text{heads}) = P(\text{heads} \mid H_1 \vee T_1) = P(\text{heads}) + \tfrac{1}{6} , \tag{4.4}$$

$$P+(\text{tails}) = P(\text{tails} \mid H_1 \vee T_1) = P(\text{tails}) - \tfrac{1}{6}.$$

Both sides also agree that it is routine to express conditional credences as a quotient of unconditional credences, and thus accept:

$$P(\text{heads} \mid H_1 \text{ or } T_1) = P(H_1 \mid H_1 \vee T_1) = P(H_1)/[P(H_1)+P(T_1)],$$

$$\tag{4.5}$$

$$P(\text{tails} \mid H_1 \text{ or } T_1) = P(T_1 \vee T_2 \mid H_1 \vee T_1) = P(T_1)/[P(H_1)+P(T_1)].$$

Lastly, it is taken that Sleeping Beauty gains no new uncentred evidence relevant to heads or tails between the times in which she has credence function P₋ and credence function P. Upon experiencing the sure event of awakening, A, the only evidence Sleeping Beauty gains is the centered evidence that she is presently experiencing an awakening, given by $(H_1 \vee T_1 \vee T_2)$.

$$P(A) = P(H_1 \vee T_1 \vee T_2) = 1 \tag{4.6}$$

Despite these many points of consensus, the Halfers and Thirders take divergent paths in their interpretation and application of these principles, leading to differing conclusions regarding the Sleeping Beauty problem. This divergence necessitates a detailed exploration of the intricacies shaping each side's argument.

Halving

According to Lewis, the centered evidence ($H_1 \lor T_1 \lor T_2$) that Sleeping Beauty gains upon her waking experience A is not relevant to heads or tails and should thus have no bearing upon her credences.[202] Therefore, since there is no new relevant information that Sleeping Beauty gains between P- and P, the Principle of Indifference should be applied to the fair coin toss and thus her credence at P must remain ½. The proof is as follows:

Premise	Inference	Rule
(A1)	$P_-(H) = ½$	Equation (4.4) shows that the coin is stipulated to be fair.
(A2)	$P(A) = 1$ $P(H_1 \lor T_1 \lor T_2) = 1$	According to Equation (4.6), Sleeping Beauty knows that she will be awoken at least once and will be aware of her waking state when she is, where A is the event that she is awoken.
(A3)	$P(H \mid A) = ½$	Equation (3.1), Bayes' Theorem allows us to conditionalize upon this sure event.
(A4)	$P(H) = ½$	Equation (4.6) entails that A represents the entirety of the new evidence she gains from P- to P.

It is important to note that for (A4) to be valid in the Halfer argument, Sleeping Beauty must believe that the knowledge she possesses upon awakening, prior to learning the day, is identical on Monday and Tuesday.

---

[202] Lewis, "Sleeping Beauty: Reply to Elga," 171-176.

Thirding

Thirders use a centered approach to argue that Sleeping Beauty's credence in heads upon awakening should instead be ⅓. By breaking down Sleeping Beauty's waking experience A, the Thirder argues that the Principle of Indifference should be applied to the three indistinguishable waking states rather than the coin toss. By this logic, Sleeping Beauty should have equal credence in H_1, T_1, and T_2, yielding a result of ⅓. The proof is as follows:

Premise	Inference	Rule
(B1)	$P(T_1) = P(T_2)$; $P(T_1 \mid T_1 \vee T_2) = P(T_2 \mid T_1 \vee T_2) = $ ½.	Equation (4.1) shows that the probability of T_1 and T_2 are equal, so the probability of Monday and Tuesday given tails must be as well.
(B2)	$P(T_1 \mid H_1 \vee T_1) = $ ½; $P(H_1 \mid H_1 \vee T_1) = $ ½; $P(H_1) = P(T_1)$.	Given that it is Monday, she will have been awoken regardless of the coinflip. Equation (4.5) gives that H_1 and T_1 are thus equally likely.
(B3)	$P(H_1 \vee T_1) = $ ⅔; $P(T_2) = $ ⅓.	Equation (4.6) allows us to break down Sleeping Beauty's sure event of waking, A, into two jointly exhaustive and mutually exclusive, centered events: "It is Monday" ($H_1 \vee T_1$) and "It is Tuesday" (T_2). It follows from (B2) and (B3) that ($H_1 \vee T_1$) is twice as likely as (T_2).
(B4)	$P(\text{heads}) = P(\text{heads} \mid H_1 \vee T_1)*P(H_1 \vee T_1)$;	Equation (3.2), the Law of Total Probability, allows us to calculate the

	$P(\text{heads}) = (\frac{1}{2})(\frac{2}{3}) = \frac{1}{3}.$	likelihood of heads via the distinct events by which it may be realized.

Once again, it is crucial to understand Sleeping Beauty's epistemic state in the context of this argument. Suppose x represents the knowledge Sleeping Beauty possesses upon awakening. In that case, $(H_1 \lor T_1 \lor T_2)$, the Thirder explains that the probability of learning x given tails is twice that of learning x given heads because Sleeping Beauty has two opportunities to learn x given tails compared to only one opportunity given heads.

When examined individually, both the Halfer and the Thirder arguments appear completely rational. Thus, it is no surprise that these positions have spawned rigorous debate and arguments attempting to refute each other.

Counter Arguments

A central argument against halving cites a violation of the Principal Principle, which states that one should align their credences with known objective chances. For Sleeping Beauty's credence in H_1 to be ½, her credence in T_1 and T_2 must sum to ½, as these possibilities are exhaustive. Since T_1 and T_2 are believed to be equally likely, the Halfer argues that a credence of ¼ should be assigned to these states. However, H_1, T_1, and T_2 are subjectively indistinguishable, and thus given A, she is equally likely to be in any of these three situations. Why, then, should she assign a difference in credence among these situations? The Halfer rebuts by suggesting a variety of counterexamples to the Principal Principle. These include that one may have "inadmissible evidence," their evidence may "undermine" the hypothesis, or, most importantly, self-locating uncertainty may make it impossible to satisfy the Principal Principle.[203,204]

This idea of self-locating uncertainty is especially appealing to the Halfer in this situation, as this problem relies on the arguably problematic usage of indexicals—propositions whose meaning depends on the contexts in which they are used. In the context of the Sleeping Beauty

---

[203] David Lewis, "A Subjectivist's Guide to Objective Chance," *Philosophical Papers: With Postscripts*, Vol. 2 (1986): 83–132.

[204] Michael G. Titelbaum, "The Relevance of Self-Locating Beliefs," *The Philosophical Review* 117, no. 4 (October 2008): 555-605.

problem, the statement, "Today is Monday," is not a legitimate classical proposition because its truth varies depending on the context. According to Richard Epstein in "Classical Mathematical Logic: The Semantic Foundations of Logic," in order for a proposition to be valid in this regard, it must mean the same thing and have the same true/false value in all contexts within the scope of the analysis.[205] In the scope of the Sleeping Beauty problem, the meaning of the indexical "today" is not fixed and stable, demonstrating the challenges that self-locating uncertainty raises when applying traditional principles in Bayesian theory.

Turning our attention to the Thirders, it becomes clear that they, too, invite criticism. It is agreed that Sleeping Beauty should have a credence of ½ in heads on Sunday night or that P (heads) = ½. In modeling Sleeping Beauty as a canonical Bayesian, it is required that she update her credence using only relevant evidence. A Halfer would argue (under the Halfer assumption) that the entirety of the evidence gained is the centered evidence A = (H$_1$ $\vee$ T$_1$ $\vee$ T$_2$), and since this evidence is irrelevant to an outcome of heads, conditionalizing upon it should produce no change in credence. Accordingly, the Thirder violates the Bas Van Fraassen's Reflection Principle, which states that if an agent knows that they will have credence c in a particular claim tomorrow, so long as this agent does not learn any new information in the meantime, they should also have credence c in that claim today.[206]

Elga and other Thirders defend their argument by suggesting that the Sleeping Beauty case is a counterexample to the Reflection Principle. Using the Dutch Booking Argument for Reflection and Conditionalization, Christensen shows that Reflection involves higher-order beliefs, while conditionalization does not, giving us grounds for rejecting Reflection even if we accept conditionalization.[207] Green and Hitchcock come to a similar result, concluding that in less-than-ideal circumstances where an agent suspects he or she will adopt some less rational policy, agents need not obey the Reflection Principle.[208] This critique raises fundamental questions about the scope and limitations of traditional Bayesian principles, suggesting that there may be contexts in which they fall short and opening up space for alternative interpretations and models of rational belief updating.

[205] Richard L. Epstein, *The Semantic Foundations of Logic* (New York: Oxford University Press, 1994).
[206] Bas C. van Fraassen, "Belief and the Will," *The Journal of Philosophy* 81, no. 5 (May 1984): 235–256.
[207] David Christensen, "Dutch-Book Arguments Depragmatized: Epistemic Consistency for Partial Believers," *The Journal of Philosophy* 93, no. 9 (September 1996): 450–479.
[208] Mitchell S. Green and Christopher Hitchcock, "Reflections on Reflection," *Synthese* 98.

Technicolor Beauty

Titlebaum takes the Thirder perspective a step further, proposing a thought experiment called Technicolor Beauty which he claims can recover Elga's answer of ⅓ without invoking the much-debated Principle of Indifference or employing the Principal Principle. This is essential, as these 'norms of rationality' are just as the name suggests. They are merely norms–conveniences and compromises that are unable to wield the unimpeachable logical solidity of well-established probabilistic axioms like Bayes' Rule or the Law of Total Probability. Whether or not Technicolor Beauty truly solves the Sleeping Beauty problem is contentious, but it nonetheless provides us with a conceptual framework which will be essential in unraveling the assumptions that give rise to this disagreement.

Consider an experiment with the same parameters as the original Sleeping Beauty problem but with one difference. Sleeping Beauty has a friend who is part of the research team conducting the experiment, and he tells her that he will do her a favor ahead of time. While the other researchers are flipping their coin, he will roll a fair die, and if odds, he will place a red sheet of paper next to her bed that Sleeping Beauty will see when she wakes up on Monday and a blue sheet of paper next to her bed that she will see if she wakes up on Tuesday. If the outcome is even, he will instead place the blue sheet of paper next to her bed on Monday and the red sheet next to her bed on Tuesday. Sleeping Beauty is certain that her friend will carry out this action, and she is certain that when she wakes up, she will see either a blue or red piece of paper, despite not knowing whether it is Monday or Tuesday. She falls asleep that night and sometime later awakens in this state–unsure of whether it is Monday or Tuesday but staring at a colored piece of paper. The question is the same–at this moment, what credence should Sleeping Beauty assign to the coin landing heads?

When awoken on Monday, she does not know the day, but she does have a unique way of representing it–i.e., "the red paper day." This is crucial, as the presence of this colored paper allows for a direct comparison of Beauty's Sunday night and Monday morning degrees of belief. On Sunday night, both sides agree that the chance of heads and the chance of Monday being the red paper day are both ½, as Beauty is sure that the coin flip and die roll are fair and independent. In the event that Monday is in fact the red paper day, if the coin landed tails, Beauty is

guaranteed to be awoken on the red paper day, with only a ½ chance for heads. If Monday is the blue paper day, she is guaranteed to be awoken on the blue paper day given tails, with only a ½ chance for heads. The experiment specifies that given heads she will only be awoken once (H_1) and will thus only see one colored piece of paper, compared to the two different colored papers that she will see given tails ($T_1 \lor T_2$). Therefore, upon awakening, regardless of what color she is looking at, she is twice as likely to be seeing it given tails. This piece of paper serves as newly acquired evidence of tails upon which Beauty may conditionalize, decreasing her credence in heads from ½ to ⅓. Conditioning upon the newly gained information that "today" is the "red paper day" or "today" is the "blue paper day" produces a meaningful change in credence, despite this information being independent of the coin flip.

Titelbaum extends this logic to argue that even if Sleeping Beauty does not learn anything about the coin toss or the world outside her upon waking, she does learn new, self-locating information that justifies her change in credence in the same way observing a colored piece of paper would.[209] This position invites criticism, however, as without this colored paper, no other evidence is specified for Beauty to conditionalize upon in order to produce this change in credence. Brian Weatherson demonstrates this point by altering the Technicolor Beauty case to make the days phenomenologically identical, showing that Technicolor Beauty should then be a Halfer.[210]

Titelbaum does, however, provide something of great value to the original Sleeping Beauty case through his description of a "uniquely denoting context-insensitive expression for 'today.'" With this concept in mind, addressing the fundamental differences resulting in the Halfer and Thirder perspectives is much easier. Technicolor Beauty highlights the Halfer's assumption that no new relevant information is gained upon waking and further demonstrates the Thirder's approach of considering each awakening as a separate piece of evidence. Upon closer inspection, these two fundamental assumptions regarding the consideration of potential uniquely denoting context-insensitive information in the Sleeping Beauty problem result in the schism between Halfers and Thirders.

---

[209] Titelbaum, "The Relevance of Self-Locating Beliefs," 555-605.
[210] Brian Weatherson, "Ross on Sleeping Beauty," *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition* 163, no. 2 (March 2013): 503–12.

Underlying Assumptions

Navigating through this labyrinth of probabilities and epistemic statuses, we arrive at an instructive junction. The rift between the Halfer and Thirder perspectives, as indicated by the Technicolor Beauty case, is not the result of flawed probabilistic reasoning; rather, it emerges from foundational assumptions about the acquisition of new evidence upon waking. These assumptions, implicitly or explicitly, dictate the course of the Sleeping Beauty's credal adjustments. Cisewski, et al. break down the Sleeping Beauty problem using a standard Bayesian account of rational belief with conditioning with no special handling for self-locating beliefs and centered propositions to demonstrate the fundamental basis of both viewpoints.[211] This is crucial because it demonstrates the necessary and sufficient conditions for halving and thirding using strictly probability theory without possible world semantics or indexicals. Conditioning upon what Sleeping Beauty knows upon awakening during the experiment reveals the underlying assumptions that Thirders and Halfers rely on in reaching their respective answers.

In order for sleeping Beauty to be a Halfer upon awakening, she must make use of the necessary and sufficient condition:

> i: If the coin lands tails, Beauty believes that what   (HA)
>
>   she knows on Tuesday at time t is identical to what
>
>   she knows on Monday at time t.

This condition implies that every piece of knowledge Beauty has on Monday, including every sensation, thought, and bodily function, will be known again on Tuesday if the coin lands tails. In this sense, the experiment should look more like the experimenter time traveling than Sleeping Beauty forgetting. This assumption forms the foundation of the Halfer perspective that regardless of the day, the probability of the coin landing heads must remain ½. No matter what, she will awaken in the exact same position with no way to uniquely denote the day and thus no

---

[211]Jessi Cisewski, Joseph B. Kadane, Mark J. Schervish, Teddy Seidenfeld, and Rafael Stern, "Sleeping Beauty's Credences," *Philosophy of Science* 83, no. 3 (July 2016): 324–347.

new evidence. In this case, there is nothing to conditionalize upon in order to change her Sunday night credence function from ½.

Unsurprisingly, this assumption is incompatible with the necessary and sufficient conditions for Sleeping Beauty to be a Thirder:

     i. If the coin lands tails, Beauty must have different   (TA)
      knowledge or beliefs on Tuesday than she has on
      Monday.
     ii. The conditional distribution of what Beauty knows
      on Monday given heads is the average of what she
      knows on Monday given tails and the conditional
      distribution of what she knows on Tuesday given tails.

The difference in knowledge established between these two days gives Sleeping Beauty a way of describing her epistemic state in reference to this difference. By doing so, she is able to make specific claims about this day, i.e. today is the red paper day and Monday is the red paper day. Now, propositions of the form Monday is the x day, where x represents the difference in knowledge Beauty may have, become relevant.

In the Technicolor Beauty scenario, this difference in knowledge is clearly represented by Beauty's colored paper. However, the same logic prevails in the original case, so long as Beauty experiences any difference between these two days which could serve as a unique context-insensitive identifier. The moment that Sleeping Beauty experiences anything that she believes will not or did not occur during her other awakening, including the most trivial details, her degree of belief in heads will decrease from ½ and approach ⅓, based on how certain she is that this experience is unique.

Despite the divergence in these assumptions, it is not the case that one side blatantly violates the principles of probability theory or Bayesian inference. The validity of both stances within their respective assumptions nudges us towards an intriguing crossroad: the establishment of a model that encapsulates both these viewpoints while maintaining congruence with Bayes' Law and the Law of Total Probability. To do so, this model must account for both Halfer's and

Thirder's assumptions by conditioning upon the knowledge she possesses upon awakening to some event, B.

Axiomatic Analysis

When we analyze the effect of these conditional probabilities on the resulting marginal probabilities, it becomes clear that both positions and their respective assumptions are compatible with the laws of probability theory.

Let's say that, upon awakening, Beauty experiences some event, call it B_{HA}, which satisfies the Halfer's assumption such that if the coin lands tails, Beauty believes that she is currently reliving, or will relive this exact same experience be it Tuesday or Monday. It follows that if she is to experience some B_{HA} on Monday, then she must experience the exact same B_{HA} on Tuesday. Thus, the following possibilities partition her sure event of waking, A:

$$\Omega_{HA} = \{\text{heads} \cap B_{HA}, \text{tails} \cap B_{HA}\}. \tag{HA1}$$

Upon experiencing this event B_{HA}, she has no reason to update her credal function, as she has a 100% chance of experiencing the same B_{HA} regardless of the day and regardless of the coinflip. By the law of total probability:

$$P(B_{HA}) = P(\text{heads}) * P(B_{HA} \mid \text{heads}) + P(\text{tails}) * P(B_{HA} \mid \text{tails}), \tag{HA2}$$
$$P(B_{HA}) = \tfrac{1}{2} * 1 + \tfrac{1}{2} * 1 = 1.$$

Since the event B_{HA} is independent of the coin flip, it follows:

$$P(B_{HA} \mid \text{heads}) = \frac{P(B_{HA} \cap heads)}{P(heads)}, \tag{HA3}$$
$$P(B_{HA} \mid \text{heads}) = \frac{P(B_{HA}) \times P(heads)}{P(heads)},$$
$$P(B_{HA} \mid \text{heads}) = \frac{1 \times \frac{1}{2}}{\frac{1}{2}},$$
$$P(B_{HA} \mid \text{heads}) = 1.$$

When we conditionalize upon Sleeping Beauty's newly gained centered evidence, namely her experience of the sure event B_{HA}, using Bayes' Theorem:

$$P(\text{heads} \mid B_{HA}) = \frac{P(B_{HA} \mid heads) \times P(heads)}{P(B_{HA})}, \qquad\qquad (\text{HA4})$$

$$P(\text{heads} \mid B_{HA}) = \frac{1 \times \frac{1}{2}}{1},$$

$$P(\text{heads} \mid B_{HA}) = \tfrac{1}{2}.$$

This demonstrates that the Halfer assumption functions properly in accordance with Bayes' Rule to produce a credence in heads of ½.

In the Thirder case, however, a new partition is necessary. This is because her experience of some event B_{xTA} that satisfies the Thirder's assumption guarantees that, given tails, she has experienced or will experience some different event B_{yTA} upon her other awakening. If Monday is the B_{xTA} day, then Tuesday must be the B_{yTA} day. Likewise, if Tuesday is the B_{xTA} day, then Monday is the B_{yTA} day. Thus, her awakening can be described by the partition:

$$\Omega_{TA} = \{\text{heads} \cap B_{xTA},\ \text{heads} \cap B_{yTA},\ \text{tails} \cap B_{xTA},\ \text{tails} \cap B_{yTA}\}. \qquad (\text{TA1})$$

Further, the events B_{xTA} and B_{yTA} are equally likely to occur on Monday, as part (ii) of the Thirder's assumption ensures that the conditional distribution of what Beauty knows on Monday given heads is the average of what she knows on Monday given tails and the conditional distribution of what she knows on Tuesday given tails. Since these events are equally likely to occur on Monday, and Beauty is awakened on only Monday given heads, these mutually exclusive events are equally likely and collectively exhaustive in the case of heads.

$$P(B_{xTA} \mid \text{heads}) = \tfrac{1}{2}, \qquad\qquad\qquad (\text{TA2})$$

$$P(B_{yTA} \mid \text{heads}) = \tfrac{1}{2}.$$

For tails, however, she is twice as likely to see B_{xTA} or B_{yTA}, as she is guaranteed to be awoken on both days, compared to the single awakening that takes place in the event of heads.

This is true in both the case of Monday being the B_{xTA} and Monday being the B_{yTA} day. In the case that Monday is B_{xTA} day, Beauty will only be woken up on "B_{xTA} day" if heads:

$$P(B_{xTA} \mid \text{heads}) = 1, \qquad \text{(TAx1)}$$
$$P(B_{yTA} \mid \text{heads}) = 0.$$

While Beauty will be woken up on both "B_{xTA} day" and "B_{yTA} day" if tails, and thus is guaranteed to see both B_{xTA} and B_{yTA}

$$P(B_{xTA} \mid \text{tails}) = 1, \qquad \text{(TAx2)}$$
$$P(B_{yTA} \mid \text{tails}) = 1.$$

However, if Monday is B_{yTA} day, Beauty will only be woken up on "B_{yTA} day" for a coin toss outcome of heads:

$$P(B_{yTA} \mid \text{heads}) = 1, \qquad \text{(TAy1)}$$
$$P(B_{xTA} \mid \text{heads}) = 0.$$

While Beauty will be woken up on both "B_{yTA} day" and "B_{xTA} day" if the outcome is tails, once again guaranteeing her to experience both B_{xTA} and B_{yTA}.

$$P(B_{yTA} \mid \text{tails}) = 1, \qquad \text{(TAy2)}$$
$$P(B_{xTA} \mid \text{tails}) = 1.$$

Thus, regardless of which day is the B_{xTA} day and which is the B_{yTA} day, it is guaranteed that Sleeping Beauty will awaken on both of these days given tails:

$$P(B_{xTA} \mid \text{tails}) = 1, \qquad \text{(TA3)}$$
$$P(B_{yTA} \mid \text{tails}) = 1.$$

Now, we can calculate the likelihood of the Beauty awakening to the event B_{xTA} using the Law of Total Probability:

$$P(B_{xTA}) = P(\text{heads}) * P(B_{xTA} \mid \text{heads}) + P(\text{tails}) * P(B_{xTA} \mid \text{tails}), \qquad \text{(TA4)}$$

$$P(B_{xTA}) = \tfrac{1}{2} * \tfrac{1}{2} + \tfrac{1}{2} * 1 = \tfrac{3}{4}.$$

The same applies for B_{yTA}:

$$P(B_{yTA}) = P(\text{heads}) * P(B_{yTA} \mid \text{heads}) + P(\text{tails}) * P(B_{yTA} \mid \text{tails}), \qquad \text{(TA5)}$$

$$P(B_{yTA}) = \tfrac{1}{2} * \tfrac{1}{2} + \tfrac{1}{2} * 1 = \tfrac{3}{4}.$$

Now it is clear. Beauty's awakening to some unique event B_{xTA} or B_{yTA} serves as evidence of tails upon which she must conditionalize and thus update her credal function. Using Bayes' Theorem, the likelihood of heads given B_{xTA} can be calculated:

$$P(\text{heads} \mid B_{xTA}) = \frac{P(B_{xTA} \mid H) \times P(H)}{P(B_{xTA})}, \qquad \text{(TA6)}$$

$$P(\text{heads} \mid B_{xTA}) = \frac{\tfrac{1}{2} \times \tfrac{1}{2}}{\tfrac{3}{4}},$$

$$P(\text{heads} \mid B_{xTA}) = \tfrac{1}{3}.$$

Likewise, for the likelihood of heads given B_{yTA}:

$$P(\text{heads} \mid B_{yTA}) = \frac{P(B_{yTA} \mid H) \times P(H)}{P(B_{yTA})}, \qquad \text{(TA7)}$$

$$P(\text{heads} \mid B_{yTA}) = \frac{\tfrac{1}{2} \times \tfrac{1}{2}}{\tfrac{3}{4}},$$

$$P(\text{heads} \mid B_{yTA}) = \tfrac{1}{3}.$$

Thus, regardless of which day is the B_{xTA} day and which day is the B_{yTA} day, Sleeping Beauty should assign a credence of $\tfrac{1}{3}$ if she knows that there will be a difference amongst the two days that satisfies this assumption. The Law of Total Probability makes this clear:

$$P(heads) = P(heads \mid B_{xTA}) * P(B_{xTA}) + P(heads \mid B_{yTA}) * P(B_{yTA}), \qquad (TA8)$$

$$P(heads) = \tfrac{1}{3} * \tfrac{1}{2} + \tfrac{1}{3} * \tfrac{1}{2},$$

$$P(heads) = \tfrac{1}{3}.$$

This result is rather appealing, as it demonstrates that both parties are able to reach their respective conclusions from their initial assumptions without violating conditionalization. This shines light onto why both sides find their perspectives so intuitive and appealing, and further demonstrates the need for a reevaluation of the granularity of the standard Bayesian approach.

While this result is rather compelling, it is not sufficient to demonstrate that both sides are entirely justified in reaching their conclusions based on their initial assumptions. They must also act in accordance with equation (3.2) the Law of Total Probability, and its conditions stipulated in equation (3.2A):

$$P(\cup_{i=1}^{n} B_i) = 1 \quad and \quad P(B_i \cap B_j) = 0 \quad for \ i \neq j. \qquad (3.2A)$$

If the event B_j causes Sleeping Beauty to change her credences from $\tfrac{1}{2}$ to $\tfrac{1}{3}$, the Law of Total Probability's conditions fail to be met. This is because, under the Thirder's assumption, the probability of an intersection between Sleeping Beauty's knowledge upon awakening on Monday and her knowledge upon awakening on Tuesday is nonzero.

Under the Halfer assumption, however, this intersection is 0, and thus the Law of Total Probability can be applied to show that one should maintain the credence $P_- = \tfrac{1}{2} = P$. At first, this may seem like a compelling enough argument for halving. However, Cisewski et al. provide proof for a theorem that applies if the condition $P(\cup_{i=1}^{n} B_j) = 1$, is met but $P(B_i \cap B_j) \neq 0$.[212]

Let $B_1, B_2, ...B_n$ be events that satisfy:

$$P(\cup_{i=1}^{n} B_i) = 1. \qquad (11.1A)$$

Then, for every event A,

---

[212]Cisewski et al., "Sleeping Beauty's Credences," 324–347.

$$P(A) = \sum_{j=1}^{n} P(A \mid B_j)P(B_j) - \sum_{j \neq k} P(A \mid B_j \cap B_k)P(B_j \cap B_k) \qquad (11.1B)$$

$$+ \ldots + (-1)^{n+1} P(A \mid \cap_{j=1}^{n} B_j)P(\cap_{j=1}^{n} B_j).$$

This theorem yields a different result in the Sleeping Beauty case when $P(B_i \cap B_j) \neq 0$ (the Thirder's assumption). Unsurprisingly, this result is ⅓. This result is crucial because it shows that it is consistent with the laws of probability theory and conditionalization for Sleeping Beauty to have credences of ½ at a time prior to waking and know that she will acquire evidence to condition upon, and thus update her credence to ⅓, so long as this evidence does not satisfy the condition $P(B_i \cap B_j) = 0$ (the Halfer's assumption) of the Law of Total Probability. In the Technicolor Beauty case, the uniquely denoting context-insensitive expression that Sleeping Beauty gains upon observing the colored piece of paper does not satisfy this necessary condition for the Law of Total Probability and can thus justify her shift in credence. However, in the original Sleeping Beauty case, it is not inherently clear whether or not $P(B_i \cap B_j) = 0$. The Halfer thus continues to argue that Sleeping Beauty's knowledge is identical on Monday and Tuesday given tails, such that $P(B_i \cap B_j) = 0$, while the Thirder continues to argue that there is some difference which results from the difference in her temporal position, such that $P(B_i \cap B_j) \neq 0$.

Conclusion

Despite being mutually exclusive, both the Thirder's and Halfer's assumptions function correctly to yield their respective results within their shared axiomatic approach to probability theory, shedding light on why both sides find their arguments so compelling. While the original Sleeping Beauty problem does not specify which side's set of assumptions is correct, one thing is clear–the traditional Bayesian framework falters when encountering temporal, context-sensitive evidence. Sleeping Beauty's loss of certainty in her temporal position and the effects that this has on her ability to apply precise credences to propositions of interest show how changes in an agent's certainty can produce incompatible results under the standard Bayesian model.

This analysis suggests that a more inclusive framework—one that can model a broad range of scenarios and is responsive to our intuitions about prominent and established

examples—may provide an avenue for reconciling the debate between Halfers and Thirders. This framework should be able to address and incorporate the peculiarities of context-sensitive uncertainty, like Sleeping Beauty's, without compromising the foundational laws of probability theory.

One approach that may prove more effective than the standard Bayesian approach in solving the Sleeping Beauty problem is called "Imprecise Bayesianism." Imprecise Bayesianism allows agents to suspend judgment in contexts in which an agent may lack sufficient evidence to settle on a single, precise creedence, and may instead posit a range of credal possibilities. In the case of the Sleeping Beauty problem, this framework would allow Beauty to express her answer as a range of credences, x, where $\frac{1}{3} \geq x \geq \frac{1}{2}$. This value could be proven to be contingent upon her certainty in the aforementioned assumptions or the conditional likelihood of Sleeping Beauty experiencing a difference amongst her Monday and Tuesday awakenings in the event that the coin lands tails.

If the Sleeping Beauty problem contained more information regarding the nature of the drug, the experiment, or the natural chaos of her waking experience, settling on a precise value may have been reasonable, but in the case of the original problem, an imprecise range of possibilities is the most comprehensive answer available. Moreover, an imprecise framework could also help clarify the role and handling of indexical information in Bayesian inference, highlighted as a potential weakness in existing models, such as Radford Neal's Full Non-Indexical Conditioning.[213]

Nevertheless, adopting Imprecise Bayesianism isn't an unalloyed good; it introduces its own set of dilemmas, most notably of which is the problem of dilation.[214] Consequently, the pursuit of an ideal framework to assess the rationality of subjective probabilities remains an open question, necessitating further exploration to reconcile the anomalies arising from context-sensitive uncertainty.

The Sleeping Beauty problem, while ostensibly a fun philosophical puzzle, extends far beyond its simple framework. It beckons a deeper exploration into the principles and granularity

---

[213] Radford Neal, "Puzzles of Anthropic Reasoning Resolved Using Full Non-Indexical Conditioning," (August 2006).

[214] Despite being the foremost argument against Imprecise Bayesianism, it is well-argued that dilation arises in nearly every current theory of rational credence. Benjamin Eva and Reuben Stern, "Comparative Opinion Loss," *Philosophy and Phenomenological Research* (August 2022).

of Bayesian inference and nudges us toward the development of more sophisticated models able tackle nuanced uncertainties. This problem embodies the dynamism inherent in probability theory and reaffirms its pivotal role in shaping our understanding of complex decision-making processes. It is not merely about identifying the "correct" answer but refining the tools and methodologies we use to approach such dilemmas. Thus, the real value lies not in the paradox itself, but in the questions it continues to raise, fostering the continued evolution of Bayesian inference.

Bibliography

Cisewski, Jessi, Joseph B. Kadane, Mark J. Schervish, Teddy Seidenfeld, and Rafael Stern. "Sleeping Beauty's Credences." *Philosophy of Science* 83, no. 3 (July 2016): 324–347. https://doi.org/10.1086/685741.

Christensen, David. "Dutch-Book Arguments Depragmatized: Epistemic Consistency for Partial Believers." *The Journal of Philosophy* 93, no. 9 (September 1996): 450–479. https://doi.org/10.2307/2940893.

Elga, Adam. "Self-Locating Belief and the Sleeping Beauty Problem." *Analysis* 60, no. 2 (December 2000): 143-147. https://doi.org/10.1111/1467-8284.00215

Epstein, Richard L. *The Semantic Foundations of Logic.* New York: Oxford University Press, 1994.

Eva, Benjamin and Reuben Stern. "Comparative Opinion Loss." *Philosophy and Phenomenological Research* (August 2022). https://doi.org/10.1111/phpr.12921.

van Fraassen, Bas C. "Belief and the Will." *The Journal of Philosophy* 81, no. 5 (May 1984): 235–256. https://doi.org/10.2307/2026388.

Green, Mitchell S. and Christopher Hitchcock. "Reflections on Reflection." *Synthese* 98 (February 1994): 297-324. https://doi.org/10.1007/bf01063945.

Lewis, David. "Sleeping Beauty: Reply to Elga." *Analysis* 61, no. 3 (July 2001): 171-176. https://doi.org/10.1111/1467-8284.00291.

Lewis, David. "A Subjectivist's Guide to Objective Chance." *Philosophical Papers: With Postscripts*, Vol. 2 (1986): 83–132. https://doi.org/10.1093/0195036468.003.0004.

Neal, Radford. "Puzzles of Anthropic Reasoning Resolved Using Full Non-Indexical Conditioning." (August 2006). https://doi.org/10.48550/arXiv.math/0608592.

Titelbaum, Michael G. "The Relevance of Self-Locating Beliefs." *The Philosophical Review* 117, no. 4 (October 2008): 555-605. https://doi.org/10.1215/00318108-2008-016.

Weatherson, Brian. "Ross on Sleeping Beauty." *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition* 163, no. 2 (March 2013): 503–12. https://doi.org/10.1007/s11098-011-9827-x.

Zuboff, Arnold. "One Self: The Logic of Experience." *Inquiry: An Interdisciplinary Journal of Philosophy* 33, no. 1 (March 1990): 39-68. https://doi.org/10.1080/00201749008602210.

The UBC Journal of Philosophical Enquiries