FRAGMENTS

McGill Undergraduate Journal of Philosophy

Issue 36 - 2022

TEAM

This project is generously funded by AUS Financial Management Committee and Philosophy Students' Association.

CONTENTS

Non-Naturalist Responses to the Generalized Modest Humean Principle

Julien Gilchrist da Silva

Moral facts supervene on descriptive facts. Put differently, this means that moral properties and descriptive properties always covary. It seems *prima facie* impossible that two descriptively identical states of affairs could differ morally. A moral naturalist can explain this necessary connection very easily. But, for non-naturalists, an explanation for supervenience is not so easy, and many non-naturalists explain it away as a brute fact. However, Stephanie Leary (2017) offers a unique metaphysical account of supervenience by leveraging the explanatory power of essences. Nevertheless, Tristram McPherson and David Plunkett (2021) remain unconvinced. They argue in favour of the Generalized Modest Humean Principle, the idea "that it should count significantly against a theory if it posits an unexplained fact that rules out the free recombination of discontinuous properties across the space of metaphysical possibilities." In this paper, I explore the two main options non-naturalists have in response to the Generalized Modest Humean Principle. They may either deny that moral supervenience is metaphysically necessary or deny that they violate the Generalized Modest Humean Principle. In the end, I show that the non-naturalist should deny that they violate the Generalized Modest Humean Principle because the theoretical cost of denying metaphysically necessary moral supervenience is too high.

Keywords: moral supervenience, non-naturalism, essence, metaethics

I. Introduction

Moral facts supervene on descriptive facts. Put differently, this means that moral properties and descriptive properties always covary, and as a result it seems *prima facie* impossible that two descriptively identical states of affairs could differ morally. A moral naturalist can explain this necessary connection very easily. On this view, moral facts and descriptive facts are both natural properties (they are of the same type). So, it is trivially easy to show that moral properties supervene on descriptive properties because same-type properties always supervene on themselves. But non-naturalists assert that moral properties are *sui generis*[1] and so they must give an alternative explanation for why natural properties and moral properties supervene. Some moral philosophers think that non-naturalism cannot provide a plausible explanation for supervenience without a problematic reliance on brute facts, and this is the Supervenience Objection to non-naturalist moral realism.

One strategy the non-naturalist can employ to meet the objection is to provide a metaphysical explanation for supervenience. This entails describing a metaphysical picture of the relationship between moral properties and descriptive properties. Stephanie Leary (2017) summarizes several ways that non-naturalists can provide a metaphysical picture of the relationship between moral facts and descriptive facts and shows that none of the views provided by non-naturalists like Moore and Shafer-Landau, among others, meet the Supervenience Objection effectively. At the end of the chapter, Leary offers a unique metaphysical account of supervenience by leveraging the explanatory power of essences and autonomous facts. Leary argues that there are hybrid normative properties that serve as metaphysical "double-sided tape," sticking the descriptive facts to the *sui generis* moral facts. The hybrid normative properties do this work because facts about their essences involve purely naturalistic conditions for their instantiation, as well as sufficient conditions for the instantiation of *sui generis* normative properties. Moreover, essence facts are autonomous and therefore do not require grounding.

If this is all right, then non-naturalists can indeed provide a metaphysical explanation for supervenience. However, Tristram McPherson and David Plunkett remain unconvinced. They claim that objecting to the "Hybrid Gambit" helped them understand the limitations of a non-naturalist metaphysical explanation for supervenience. They argue in favour of the **Generalized Modest Humean Principle**, "that it should count significantly against a theory if it posits an unexplained fact that rules out the free recombination of discontinuous properties across the

1. I will consistently use the phrase '*sui generis*' with the meaning 'of their own type.'

space of metaphysical possibilities."[2] In this paper, I argue that the non-naturalist should reject the proposition that they necessarily violate the Generalized Modest Humean Principle.

First, I give a brief definition of essence and autonomy and explain the difference between brute and autonomous facts. Second, I present the Generalized Modest Humean Principle in full and explain the force it has on non-naturalist metaphysical explanations for moral supervenience. I then evaluate two potential options for the non-naturalist faced with the Generalized Modest Humean Principle. I argue that the non-naturalist can either deny that moral supervenience is metaphysically necessary or deny that they violate the Generalized Modest Humean Principle itself. The first option entails that the non-naturalist effectively adopt a Gideon Rosen-style moral contingentist view. As I will argue, this approach is problematic. On the other hand, the second option requires that the non-naturalist explain why they do not violate the Generalized Modest Humean Principle. I argue that this is possible, and hence that the second option is a viable response to the Generalized Modest Humean Principle.

II. Essence and Autonomy

As the concepts of essence and autonomy ground the discussions in this paper, I will now preliminarily explain them. The complete essence of a thing is the set of all propositions that are necessarily true (the "essentialist truths") in virtue of the thing being what it is.[3] In this way, an essentialist truth is what logically follows from the definition of the thing. For instance, it is in the essence of a bachelor to be unmarried. It is in the essence of a moon that it orbits a planet. In this way, the essence facts carve out the space of all metaphysical possibilities. Additionally, on Dasgupta's view, the essence facts are autonomous, meaning that they are not apt for grounding.[4] To elucidate his meaning, Dasgupta makes a distinction between substantive facts and autonomous facts. He says that a "substantive fact" is apt for being grounded, whereas an "autonomous fact" is not. This means that it would make sense to ask the question "what grounds *x*?" if *x* is a substantive fact, but it would not make sense if *x* were autonomous. So, a brute fact is a kind of substantive fact, where the question "what grounds a brute fact?" is met with the response "nothing." Notice here that the question is still perfectly intelligible, and it is a perfectly intelligible answer to say that brute facts do *not* have a ground.

2. Tristam McPherson and David Plunkett, "Supervenience and the Autonomy of Essence: Lessons from Leary's Hybrid Gambit," (unpublished manuscript, 2021), typescript.
3. Shamik Dasgupta, "The Possibility of Physicalism," *Journal of Philosophy* 111, no. 9-10 (2014): 578.
4. Ibid., 575.

To grasp why, consider that physical states of affairs can be causally grounded in some earlier physical state. However, one might think that the arrangement of particles at the beginning of the universe does not have a causal ground. If nothing caused it to be like that, then it can rightly be called causally brute. Notice here that it is still an intelligible question to ask, "what caused the initial physical state of the universe?" even if the answer truly turns out to be "nothing."[5] Conversely, it is not intelligible to ask what causes mathematical facts. It is not just that nothing grounds these sorts of facts; he argues that questions about causation are incoherent when we talk about abstract objects like numbers. This is the sense in which autonomous facts are not apt to be grounded, and what distinguishes autonomous facts from brute facts.

Finally, Dasgupta argues that autonomous facts and definitions of things are closely related. He motivates this by claiming that it seems incoherent to ask for a proof of a definition. The claim is that "autonomous facts are not apt to be grounded" in the same way that "definitions are not apt for being proved."[6] We will leverage this idea later in the paper.

III. The Generalized Modest Humean Principle

The Generalized Modest Humean Principle is a heavily modified and generalized version of Hume's Dictum. In its original formulation, David Hume writes, "there is no object, which implies the existence of any other if we consider these objects in themselves." In Jessica Wilson's formulation, the principle reads, "there are no metaphysically necessary connections between distinct, intrinsically typed [meaning 'of their own kind'], entities."[7] For example, a non-naturalist takes normative properties to be of their own kind (recall: *sui generis*) and therefore distinct from non-normative properties. So, in its original sense, Hume makes a claim about what sorts of connections actually exist. McPherson adapts this idea into a more modest principle called Modest Humean, stating that it counts significantly against a view if it heavily relies on brute necessary connections between metaphysically discontinuous properties. Here, McPherson makes no claims about what kinds of connections exist, only that a theory is less credible if it relies on brute necessary connections. But Leary uses autonomous facts, not brute facts, in her explanation of moral supervenience. Acknowledging this, McPherson and Plunkett adapt Modest Humean into the much stronger Generalized Modest Humean. Recall here that the Generalized Modest

5. Ibid., 576.
6. Ibid., 577.
7. Jessica Wilson, "What is Hume's Dictum and Why Believe it?" *Philosophy and Phenomenological Research* 80, no. 3 (2010): 595.

Humean Principle claims that it counts significantly against a theory if it posits any unexplained fact (read: brute *or* autonomous) that limits free recombination of metaphysically discontinuous properties. Effectively, the Generalized Modest Humean Principle is just the Modest Humean Principle generalized to brute *and* autonomous facts.

McPherson and Plunkett coin the term "metaphysically discontinuous" to describe differently typed properties.[8] For example, the property of being red and the property of being a square are metaphysically discontinuous. To illustrate, notice that it should be possible to freely recombine two metaphysically discontinuous properties, like being red and being a square. More specifically, in metaphysically possible worlds, we should expect that not all squares are red and not all red things are squares without some good reason. Maybe we can imagine a world where it is metaphysically necessary that all squares are red, but in this world, there is no reason why this must be the case. In short, shape and colour are meant to be different types of properties, and thus metaphysically discontinuous. Moreover, the properties of being red and being a square can be "freely recombined."[9] However, if an object is scarlet, for instance, then it must also be red. There cannot be a scarlet object that is not red in any world. In this way, red and scarlet are metaphysically continuous. The properties of being red and being scarlet cannot be freely recombined.

According to McPherson and Plunkett, non-naturalists believe normative and non-normative properties are analogous to being red and being a square. This is because, to the non-naturalist, normative and non-normative properties are differently typed (recall, normative properties are *sui generis*) and thus metaphysically discontinuous. So, without some good reason, we should be able to freely recombine them.

Of course, non-naturalists suppose that there is a good reason we *cannot* freely recombine normative and non-normative properties; they suppose that there are normative properties that necessarily covary with non-normative properties (this is just moral supervenience). But then McPherson and Plunkett argue that the non-naturalist cannot give a metaphysical account of supervenience that does not posit some unexplained—either brute or autonomous—fact. They argue that an unexplained fact is not a good reason to limit the free recombination of discontinuous properties.

The Generalized Modest Humean Principle is so strong because it targets the only avenue a non-naturalist has for explaining supervenience: describing a metaphysical picture. McPher-

8. McPherson and Plunkett, "Autonomy of Essence," 1.
9. Ibid., 11.

son and Plunkett argue that any attempt to give a metaphysical explanation of supervenience must rely on either brute or autonomous connections between normative and non-normative properties. So, the Generalized Modest Humean Principle presents a seemingly insurmountable challenge to the non-naturalist. To be a non-naturalist in the first place, one must suppose the existence of unexplained facts in one's metaphysical account of moral supervenience. But Generalized Modest Humean says that it counts significantly against the theory to do this. In this way, McPherson and Plunkett have challenged the non-naturalist to an unwinnable game, and so the only winning move is to not play. The non-naturalist will have to find a different way to address the objection. There are two ways of doing this: either (A) non-naturalists can reject that moral supervenience is metaphysically necessary, or (B) they can reject that their position violates the Generalized Modest Humean Principle as it is formulated. Below I explore these two options and conclude that only the latter is viable.

IV. Non-Naturalist Responses

(A) *Reject Metaphysically Necessary Moral Supervenience*

First, the non-naturalist could reject that moral supervenience is metaphysically necessary. This approach follows in the spirit of Gideon Rosen. Rosen argues that, rather than being metaphysically necessary, normativity belongs to its own modality: normative necessity. If this is true, then the non-naturalist sidesteps the Generalized Modest Humean Principle because it only targets moral supervenience taken as a metaphysical necessity. However, Rosen's approach effectively amounts to moral contingentism, something that other non-naturalist philosophers like Jamie Dreier take to be highly objectionable. To see this dialectic play out, let us inspect this option more closely.

Recall that moral supervenience is the idea that there cannot be a change in moral properties without a change in descriptive properties. Rosen's first move is to argue that this is only normatively necessary, not metaphysically necessary. Here, normative necessity is another modal class distinct from metaphysical or even conceptual necessity. Moral supervenience will hold in any set of worlds with the same normative laws. A useful analogy here is with physical laws. For Rosen, moral laws are normatively necessary in the same way that laws of gravity are physically necessary.[10] Laws of gravity are true because they are physically necessary. Something is physically necessary if and only if it is a physical law. In the same way, something is normatively

10. Gideon Rosen, "The Modal Status of Moral Principles," in *Oxford Studies in Metaethics, Vol. 16,* ed. Russ Shafer-Landau, 257-279 (Oxford: Oxford University Press, 2021), 264.

necessary if and only if it is a normative law. Supervenience is thus explained as a normative necessity.

This would solve the problem because under Rosen's moral contingentism, there are no metaphysically necessary connections between normative and non-normative properties; there are only normatively necessary connections. For example, *being a murder* and *being wrong* can now be freely recombined across metaphysical space, as there is nothing metaphysically necessary that stops *being a murder* from *being right*. Only normatively necessary laws in a particular world would limit this.

This approach confers several benefits. First, a Rosen-style moral contingentist account avoids having to explain why moral supervenience is metaphysically necessary. Rosen wants to vindicate supervenience, but he believes, like McPherson and Plunkett, that moral supervenience cannot be metaphysically necessary. Instead, on his view, moral laws are normatively necessary, not metaphysically necessary. But committing to this position is precarious and unpopular. By sidestepping Generalized Modest Humean, moral contingentists leave themselves open to new criticisms. Jamie Dreier argues that if the laws of morality are contingent, then we should consider ourselves lucky if we follow them.[11] For example, he argues that in one world, loving and caring for children would be morally right, and in another, it would be unspeakably evil. The problem is that there would be no way to know. Dreier thinks this sort of moral contingentist luck is incredibly hard to believe, and therefore moral contingentism must be false, or at least equally hard to believe.

Rosen rejects Dreier's objection by simply denying that this objectionable form of moral luck exists. Rosen argues that we know that caring for children is morally right because we have ways of arriving at beliefs about our moral obligations.[12] This is the case even if there are possible worlds where those beliefs are false because the moral laws happen to be different. Rosen argues that it is incredibly unlikely that our moral beliefs are false in this way, however. He thinks that such a possibility is just as unlikely as a Cartesian evil demon world. For our moral intuitions to be leading us so far astray, the actual world would have to be like an evil demon world in which our ways at arriving at moral beliefs are constantly and imperceptibly being tricked. To motivate this further, he appeals to the distinction between "nearby" and "remote" possible worlds. A nearby world is a world similar to our own, whereas a remote world is very unlike it.

11. Jamie Dreier, "Is There a Supervenience Problem for Robust Moral Realism?" *Philosophical Studies* 176, no. 6 (2019): 1407-1408.
12. Rosen, "Modal Status of Moral Principles," 275.

Rosen then argues that a world where our moral intuitions are unreliable is just as remote as an evil demon world. In short, Rosen argues there is no luck involved with being moral because "a person is lucky to avoid a mishap only when a mishap was a serious possibility—that is, when there are nearby worlds in which it happens."[13] So it goes, it is not a serious possibility that our moral intuitions are unreliable and so there is no luck involved.

But if Rosen thinks that an evil demon world in which our senses are constantly being tricked is analogous to a world in which our moral intuitions are always being tricked, then we should unpack what it is to have a moral intuition and how we arrive at moral beliefs. Rosen's argument only holds if he is right that the worlds where our moral intuitions are unreliable are equally as remote as evil demon worlds where our senses are unreliable.

Consider firstly that the laws of nature are metaphysically contingent because there is no apparent metaphysical barrier to them being different than they are. Now imagine that we know the mass of an electron and that we can form a belief about it. If the mass of an electron had been different, we would have formed a different belief about the mass of an electron. This is because there is empirical evidence for creating empirical beliefs and I can verify that my belief is wrong. The reason why an evil demon world is remote is that the demon would have to trick you every time you verify, and you would always form a false belief.

But if someone steals my laptop in this universe, and I conclude that my being stolen from was objectively wrong (that is, I intuit something about the moral fact of the matter) based on my intuition, then I am committing myself to saying that the proof of its being wrong was my intuition.[14] However, imagine another descriptively identical world where the moral laws are nonetheless such that stealing laptops is morally right. In this world, there is a descriptively identical person JG in the exact same circumstances. If they conclude that their being stolen from was objectively wrong, then their intuition would be incorrect in accordance with moral laws in that world and there would be no way to verify that their intuition is incorrect. In fact, I contend that there is no way I could possibly picture JG intuitively concluding that having their laptop stolen was good. Rosen thinks that our intuitions allow us to access what is morally right in accordance with moral laws local and specific to normatively different possible worlds, but I cannot imagine my intuitions leading me and JG to contradistinctive moral beliefs in descriptively identical yet normatively different worlds. Pekka Väyrynen articulates a similar point, arguing that moral con-

13. Ibid., 276.
14. It is worth mentioning here that intuitions may not be the only way to arrive at moral beliefs. But this argument still holds in cases where we arrive at moral beliefs by perception or some other way.

tingentism would entail that the soundness of my and JG's moral conclusions seem to turn on our location on metaphysical space (one universe as opposed to another), rather than any descriptive feature of our respective situations.[15]

Simply put, laws of nature are contingent and our beliefs about natural laws can be empirically verified. But, if moral laws are contingent, then our beliefs about moral laws formed by means of intuitions cannot be empirically verified. If our moral intuitions are not verifiable in the same way, then luck is involved, and worlds where our moral intuitions are unreliable are genuinely unlucky. Note that this is different from simply restating that it is metaphysically necessary that two descriptively identical events (having your laptop stolen) must be morally identical (is wrong). I am arguing that it is impossible to imagine yourself forming the moral *intuition* that having your laptop stolen is morally good, even within a world in which having your laptop stolen indeed *contingently* maps onto the moral law that having your laptop stolen is morally good. This is an objection concerning epistemological access to local moral laws.

In short, if moral intuitions point the way to moral laws, and moral contingentism is right, there must be an objectionable form of Dreier-style moral luck involved. Whereas, if intuitions do point the way to moral laws, but the moral laws are metaphysically necessary, then there is no luck involved. This is because if we intuit what is right based on descriptive evidence, but the moral laws are contingent, then our intuitions must be unreliable in any possible worlds where they lead us to false beliefs about the laws of morality, and there is no way to know.

However, Rosen thinks that our ways of forming moral beliefs *are* reliable in all nearby worlds and *are not* in remote worlds. But, as a result, he must give a plausible story as to how this can be the case. I do not think this is possible. First, Rosen could claim that our intuitions would vary in different possible worlds according to the variance of each possible world's moral laws. But then he is committed to saying that there is some metaphysically necessary connection between our intuitive knowledge about moral laws and those laws themselves. This would be very peculiar, given that a moral contingentist probably does not want to commit to some other type of metaphysically necessary connection within their contingentist account of morality. This would be a serious theoretical cost.

This solution also strikes me as highly peculiar because it rejects metaphysically necessary moral supervenience in the first place but then accepts another, different metaphysically necessary supervenience between moral intuitions and moral laws later. It would mean Rosen

15. Pekka Väyrynen, "Against Moral Contingentism," *Thought: A Journal of Philosophy* 10 no. 3 (2021): 5.

supposes a metaphysically necessary supervenience of descriptive facts (i.e., what our moral intuitions in a given possible world end up being) upon that world's normative laws. Rosen would thus have to claim that moral laws partially cause our moral intuitions, which most would say is a less plausible claim than the former.

But Rosen could also respond differently, by saying that no connection must necessarily exist between our intuitions about moral laws and the laws themselves. Instead, he could say that the only nearby worlds are those in which our moral intuitions happen to correctly map onto our correct moral laws. Conversely, a world is remote if our moral intuitions happen to incorrectly map not onto our correct moral laws, but to incorrect local moral laws. This view holds that possible worlds are remote precisely *because* our intuitions fail us; the very fact that it was good to steal my laptop, and that my intuitions led me astray, *makes* it remote.

This position is especially strange, as there is now a sense in which moral laws are approaching being necessary. To see this, suppose that in the actual world, I am going to form a belief about an action (for example, breaking a promise) being morally wrong. The Rosen-style moral contingentist then says that any world where my belief or intuition does not line up with the moral laws is remote. Now, imagine that the moral laws were slightly different, such that the cost of breaking a promise is just slightly higher. This means that my intuitions no longer line up with the moral laws, and consequently this possible world is now remote. Here, it seems that *any* variation on the actual moral laws makes a possible world "not a serious possibility" for Rosen.[16]

Rosen himself entertains this exact example, granting that there is luck in these cases, but there is no Dreier-style moral luck in general. I think that this small crack begins to undermine the entire enterprise. A reply like this paints a strange picture of metaphysical space where the only nearby worlds are those where the moral laws are identical to our own or at least extremely close, and any substantial deviation on the laws (where our intuitions do not line up) make the world remote. Here, you're effectively on the verge of supposing a necessary connection between descriptive facts and moral facts. This picture of moral contingentism is starting to look very much like moral necessitarianism without going all-in, so to speak. But why not go all-in? It cannot be only to get around worries about supervenience, or by extension, the Generalized Modest Humean Principle. To accept Rosen's argument, I would need some good reason to believe that moral laws are not metaphysically necessary, but that any modification to them makes a world extremely unlikely. In addition, Rosen needs to provide a stronger explanation for why

16. Rosen, "Modal Status of Moral Principles," 276n22.

an extremely unlikely world is not in fact impossible, a fact which would make the contingentism rather necessitarian. I have my doubts that such an explanation can be given. These claims are in such tension that moral contingentism incurs a high theoretical cost here, too.

In this section, I argued that Rosen cannot defend moral contingentism from leading to a new form of objectionable moral luck without incurring serious theoretical costs. Recall that the original reason for adopting moral contingentism was to avoid the costs incurred by violating the Generalized Modest Humean Principle. The question remains, then, whether the shortcomings of moral contingentism are worse than the toll taken by the principle. Someone like Rosen may welcome moral contingentism as a happy alternative that avoids this toll. But I have shown here that the costs of accepting moral contingentism are far too high. So, I think it best to find an alternative.

(B) *Deny Violation of the Generalized Modest Humean Principle*

Here, I argue that the best alternative for non-naturalists is to reject that they violate the Generalized Modest Humean Principle. This is because the Generalized Modest Humean Principle targets only a subset of all non-naturalist positions which are already distinctly susceptible to it.

This section plays out in a dialectical way. Recall that McPherson and Plunkett defend the Generalized Modest Humean Principle: that unexplained facts which rule out the free recombination of metaphysically discontinuous properties significantly count against a theory. First, I attack the Generalized Modest Humean Principle head-on by denying that this always must be the case and showing that essences can appropriately leverage the explanatory power of autonomous facts. Then, I entertain a reply by McPherson and Plunkett addressing this objection. Here, McPherson and Plunkett simply broaden the Generalized Modest Humean Principle to include essential connections between "metaphysically discontinuous properties." I conclude this section by responding in turn that their understanding of "metaphysically discontinuous properties" does not characterize all non-naturalist positions. Here, I will show that a non-naturalist who adopts a certain metaphysical model of normative and non-normative properties (example: Leary's hybrid model) does not violate the Generalized Modest Humean Principle at all, nor the newly broadened version thereof.

First, recall here that McPherson and Plunkett employ the weaker "Modest Humean" principle in their paper, which claims that relying on *brute* facts which rule out the free recombination of metaphysically discontinuous properties counts significantly against a theory. If

Generalized Modest Humean is effectively the same principle extended to autonomous facts, the first move a non-naturalist could make is to argue that such this generalization of the principle to autonomous facts is a mistake. Specifically, they could substantiate this point by arguing that, when used appropriately, autonomous facts can do useful explanatory work in a theory as opposed to counting against it. This is because, unlike brute facts, autonomous facts (example: facts in the essences of things) do not require an explanation. In the hybrid account of non-naturalism, essence facts are autonomous because they are not the sorts of facts that can have an explanation, so it does not count against the theory if they are left unexplained. In short, autonomous facts are doing explanatory work in the theory that a brute fact would not.

The idea that autonomous facts could bestow explanatory power (read: credibility) to a theory should be persuasive to McPherson and Plunkett. They are quick to say that the Generalized Modest Humean Principle is an epistemological claim. They are concerned about the "epistemic credibility" of views that embrace unexplained necessary connections between metaphysically discontinuous properties.[17] So, if I can show that autonomous facts can be used effectively to increase the credibility of a theory, then this would be enough to show that the Generalized Modest Humean Principle is too strong. I believe that this is possible through Leary's account.

To start, let us acknowledge that, of course, not *all* use of autonomous facts would be effective. McPherson points out that it would be "transparently unsatisfying" to claim that supervenience is just an autonomous fact and leave it there. After all, the goal of Hume's Dictum (and its "modest" derivatives) is to prevent positing brute necessary connections whenever it seems convenient or easy.[18] But I cannot imagine a single non-naturalist who would disagree. That is why a non-naturalist account of supervenience like Leary's goes to such great lengths to describe a robust, satisfying, and ultimately credible theory. If it were as easy as claiming that supervenience is an autonomous fact or cooking up some autonomous facts whenever it was convenient, then it would be easy to agree with the Generalized Modest Humean Principle. But no serious metaphysician would regard such a strategy as viable.

Recall that Dasgupta claims it is incoherent to ask what grounds an autonomous fact for the same reason that it is incoherent to ask for the proof of a definition. If this is right, then essence facts are the correct candidate for being autonomous. Indeed, Dasgupta even uses the term

17. McPherson and Plunkett, "Autonomy of Essence," 5.
18. Tristram McPherson, "Ethical Non-naturalism and the Metaphysics of Supervenience," in *Oxford Studies in Metaethics*, *Vol. 7,* ed. Russ Shafer-Landau, 205-234 (Oxford: Oxford University Press, 2012), 26.

essence interchangeably with "real definition."[19] In other words, if essences are like definitions, then they are autonomous, meaning that they are not apt to be grounded. So, Leary makes use of autonomous facts in the essences of hybrid properties because they are the right kind of thing to be autonomous in the first place. This means that, overall, she appropriately leverages autonomous facts to explain the connections between the different types of properties.

Consequently, it is less obvious whether it should count significantly against a view if it makes use of autonomous facts to explain necessary connections. In light of the previous explanation, it seems like it depends whether they are being used appropriately or as an ad-hoc explanatory crutch. I have argued that their mere existence within a theory does not count against the view. On the contrary, I have shown that the hybrid model uses them effectively and benefits from their explanatory power to increase the credibility of the theory. To be sure, one may still want to object to the hybrid model. But one could do so while accepting the inclusion of autonomous facts.

In reply, McPherson and Plunkett further tweak the Generalized Modest Humean Principle slightly to account for essence facts. The Modest Humean-Essentialist Principle retains the spirit of the Generalized Modest Humean Principle, except that it disallows the use of unexplained essential connections between discontinuous properties.[20] They argue that they can accept a metaphysical picture of supervenience involving essences as the bottom rung of the metaphysical ladder, but they claim that "essentialist explanations, to be credible, must appeal to plausible claims about essences."[21] In brief, they do not think it is plausible that essences can explain metaphysical connections, and so in practice the possibility of doing so is ruled out.

Let's take stock of where we are so far. First, a non-naturalist could think that autonomous facts can be used appropriately to explain connections between metaphysically discontinuous properties by leveraging them in essential connections. But McPherson and Plunkett shut this down quickly by simply generalizing their principle to essential connections, too. This makes it count significantly against the theory if *essential connections* rule out the free recombination of metaphysically discontinuous properties. Now, even non-naturalists who rely on essential connections violate this new broadened principle.

But I think there is still room to object. Here, the non-naturalist reply is to shift gears somewhat and interrogate what McPherson and Plunkett mean when they say that normative

19. Dasgupta, "Possibility of Physicalism," 578.
20. McPherson and Plunkett, "Autonomy of Essence," 13.
21. McPherson and Plunkett, "Autonomy of Essence," 14.

properties and non-normative properties are metaphysically discontinuous. This is because McPherson and Plunkett think that a non-naturalist position necessarily entails conceiving of normative and non-normative properties as metaphysically discontinuous. But this does not characterize every non-naturalist theory.

Recall that the term "metaphysically discontinuous" is meant to capture the proposition that normative properties and non-normative properties are differently typed. But McPherson and Plunkett's wording and examples seem to suggest that discontinuous properties do not relate metaphysically *in any way*. This is the case for our toy example, "being red" and "being a square." But, if you understand the Generalized Modest Humean Principle this way, the problem is that not all non-naturalist theories claim that normative properties are discontinuous from the natural properties in the same way, i.e., completely metaphysically unrelated.

We can turn to Leary's hybrid model for an example. Even though the non-naturalist can claim that the essences of normative properties are not fully specifiable in terms of natural properties, and even though this means that it is not in the essence (read: definition) of the normative properties that there are natural (read: non-normative) sufficient conditions for their instantiation, this claim is still compatible with thinking that there is some way that the non-normative and the normative are metaphysically related. In other words, if what it means to be metaphysically discontinuous is to be like the example case "being red" and "being a square," as seems to be the case from McPherson & Plunkett, then simply put, the non-naturalist does not think that normative and non-normative properties are metaphysically discontinuous in this way. "Being a murder" and "being wrong" are, therefore, not in fact metaphysically discontinuous on this view.

One way to see this is to consider other sorts of facts determined by non-normative properties and see whether we find them to be metaphysically discontinuous or not. Consider two sets of facts, one relating to physics and the other to economics. It is not so clear that these facts can be freely recombined. So, on the one hand, economic facts seem to be metaphysically continuous with physical facts. Indeed, given another physical state of affairs, the set of economic facts may be very different. In this example, we might think of normative properties like the economic facts in that they are determined by natural, non-normative facts. We may also think, then, that it would be wrong to say that normative properties are metaphysically discontinuous from descriptive properties. The two are not freely recombinable. This view suggests that normative properties and non-normative properties are actually metaphysically continuous.

Of course, there is a way in which normative properties are still different from economic facts. This is because, according to the non-naturalist, normative properties are *sui generis*,

whereas economic facts can be considered extremely derivative natural facts. Put differently, something in the essence of some normative properties make them different, namely that their essences do not specify any non-normative sufficient conditions for their instantiation.[22] It is plausible that, by contrast, the essences of economic facts *do* specify entirely non-normative conditions for their instantiation.

So, this helps explain the difficulty with the term "metaphysically discontinuous" in the Generalized Modest Humean Principle. If we take metaphysically discontinuous to mean completely metaphysically unrelated, then the principle could easily be true, but then a non-naturalist who appeals to essences and essential connections in their account of metaphysics does not need to violate the principle. This is because the non-naturalist can take non-normative properties to metaphysically determine the normative properties through some essential connections. So, contrary to the initial characterization of the non-naturalist's theoretical commitments by McPherson and Plunkett, there is metaphysical continuity between the normative and non-normative properties, at least on the essentialist view.

In sum, the non-naturalist who adopts essences in their metaphysical framework can accept a certain definition of metaphysically discontinuous and continuous properties. Accordingly, they can say that the essences (read: definitions) of some normative properties (a) "cannot be specified entirely in non-normative terms" and (b) "do not specify any non-normative sufficient conditions for their instantiation." For example, (a) goodness requires some description, but it is not reducible *to* that description, and (b) no description by itself is sufficient for instantiating goodness.[23]

At this point, McPherson and Plunkett raise a new concern. They worry that if the non-naturalist takes there to be essential connections between normative and non-normative properties, then they cannot abide by their theoretical commitment that normative properties are *sui generis* and just too different from the non-normative properties which ground them.[24] They claim that Leary's hybrid model ends up being a sort of "non-reductive naturalism," and so not the type of view "targeted by the supervenience challenge in the first place."[25]

---

22. Stephanie Leary, "Non-naturalism and Normative Necessities," in *Oxford Studies in Metaethics, Vol. 12*, ed. Russ Shafer-Landau, 76-105 (Oxford: Oxford University Press, 2017), 96.
23. Leary gives the example for (a): "even a non-naturalist must admit that, it is part of the essence of *being right* that, if x is right, then x is an action" (Leary 96). Being an action is a non-normative property. So, the essence of goodness needs to have *some* descriptive properties. But it cannot be fully specified in descriptive terms.
24. McPherson and Plunkett, "Autonomy of Essence," 15.
25. Ibid., 10, 15.

At this point, one cannot help but wonder whether their supervenience challenge is rigged up such that it targets only those non-naturalist accounts which cannot respond to it. At some point, we start splitting hairs about which theories are non-naturalist and which theories are not. I will not comment here on whether Leary's hybrid view or indeed, any essence-based account of non-naturalism, turns out to be naturalist in one form or another. However, as I have shown, it does not seem inherently contradictory for the non-naturalist to claim that normative properties are *sui generis* but metaphysically determined by descriptive facts. After all, normative facts must contain at least some non-normative properties. For example, "it is part of the essence of being right that if *x* is right, *x* is an action."[26] McPherson and Plunkett seem to be saying that to be a non-naturalist, you must assert that there are no metaphysical connections *at all* between the normative and non-normative. But I do not see why a non-naturalist would need to restrict themself in this way. If they decide to, then they cannot dream of explaining supervenience, as McPherson and Plunkett have shown. In other words, I argue that if McPherson and Plunkett are not willing to allow non-naturalists to claim that there are at least *some* non-normative properties in the essences of normative properties, then their supervenience challenge is overly restrictive. Indeed, the challenge would then only target the subset of non-naturalist theories that fail against it in the first place.

Here, I have shown that some non-naturalists can deny that they violate the Generalized Modest Humean Principle. They could claim that autonomous facts bolster their credibility rather than hinder it or assert that the principle only succeeds against a narrowly construed version of non-naturalism in which normative and non-normative properties are completely metaphysically unrelated. Indeed, a non-naturalist may even grant that the Generalized Modest Humean Principle holds for versions of non-naturalism that construe normative and non-normative properties as completely metaphysically unrelated. But some non-naturalist theories, including Leary's hybrid model, deny this metaphysical story. So, non-naturalists who posit metaphysical continuities between the normative and non-normative can deny that they violate the Generalized Modest Humean Principle (and the Modest Humean-Essentialist Principle, too).

V. Conclusion

Above I have summarized two responses to the Generalized Modest Humean Principle put forward by McPherson and Plunkett. My contention is that McPherson and Plunkett have

26. Leary, "Normative Necessities," 96.

targeted an impoverished notion of moral non-naturalism that lacks the tools to defend itself from the supervenience challenge. There is no doubt that the version they target does incur a significant cost under the Generalized Modest Humean Principle. In other words, if there really are no metaphysical connections at all between the essences of normative and non-normative properties, then supervenience effectively becomes impossible to explain without incurring a theoretical cost. So, this exploration into non-naturalist responses to the Generalized Modest Humean Principle teaches us a great deal about what it takes for a non-naturalist theory to be tenable and credible.

First, I argued that moral contingentism was not tenable. Second, I argued that non-naturalists should consider moral properties and descriptive properties to be unlike "being red" and "being a square." Thinking this way is too limiting for a non-naturalist position. On the contrary, thinking in terms of essential connections between the normative and non-normative properties while maintaining the *sui generis* nature of the non-normative seems possible, and drives the discussion forward for an ever-more-credible non-naturalist metaphysical explanation for supervenience.

Bibliography

Dasgupta, Shamik. "The Possibility of Physicalism." *Journal of Philosophy* 111, no. 9-10 (2014): 557–92.

Dreier, Jamie. "Is There a Supervenience Problem for Robust Moral Realism?" *Philosophical Studies* 176, no. 6 (2019): 1391–1408.

Leary, Stephanie. "Non-naturalism and Normative Necessities." In *Oxford Studies in Metaethics, Vol. 12*, edited by Russ Shafer-Landau, 76-105. Oxford: Oxford University Press, 2017.

McPherson, Tristram. "Ethical Non-naturalism and the Metaphysics of Supervenience." In *Oxford Studies in Metaethics, Vol. 7*, edited by Russ Shafer-Landau, 205-34. Oxford University Press, 2012.

McPherson, Tristram and David Plunkett. "Supervenience and the Autonomy of Essence: Lessons from Leary's Hybrid Gambit." (Unpublished manuscript, 2021), typescript.

Rosen, Gideon. "The Modal Status of Moral Principles." In *Oxford Studies in Metaethics, Vol. 16*, edited by Russ Shafer-Landau, 257-279. Oxford: Oxford University Press, 2021.

Wilson, Jessica. "What is Hume's Dictum, and Why Believe it?" *Philosophy and Phenomenological Research* 80, no. 3 (2010): 595-637.

Väyrynen, Pekka. "Against Moral Contingentism." *Thought: A Journal of Philosophy* 10, no. 3 (2021): 209-217.

Andy Egan on Quasi-Realism and Moral Disagreement

Nathan Michel

Quasi-realism (as a sophisticated version of expressivism) occupies some peculiar territory in metaethics. Unlike naive expressivists, quasi-realists are especially committed to accommodating ordinary moral discourse without biting first-order ethical bulletws. Some philosophers have argued that this commitment gets quasi-realists into trouble when it comes to ordinary moral discourse surrounding moral error and disagreement. In this paper, I focus on one such argument against quasi-realism from Andy Egan and argue that it is unsound. In Section I, I outline Egan's argument against quasi-realism. In Section II, I challenge one of the argument's premises. In Section III, I respond to potential objections to my claims from Section II. I conclude that quasi-realists have nothing to fear from Egan's argument and others like it.

Keywords: metaethics, quasi-realism, non-cognitivism, moral disagreement, moral error

I. Introduction

One challenge that expressivist theories in metaethics face is explaining how talk of moral error fits into our ordinary moral discourse. If in making moral judgements all we are doing is expressing our attitudes, as expressivists of all stripes claim, it's not clear how we can be right or wrong about morality. But this is clearly at odds with our ordinary moral discourse. We certainly talk as if people can be right or wrong about morality, and we often worry that we ourselves might be mistaken in some of our moral beliefs. Expressivists need an account of why this is. Andy Egan's paper "Quasi-Realism and Fundamental Moral Error" discusses a particular objection to quasi-realism related to these worries. In the paper, Egan outlines an influential and promising expressivist account of first-person present moral error that comes from Simon Blackburn. He then goes on to develop an argument against the expressivist view known as quasi-realism. He argues that, given Blackburn's account of first-person present moral error, quasi-realism generates unacceptable results in certain cases of moral disagreement. In this paper, I argue that Egan's argument is unsuccessful. Before we proceed, I will first lay out Blackburn's account of moral error and explain Egan's argument against the quasi-realist view.

One datum that expressivists need to explain when it comes to our ordinary moral discourse surrounding moral error is that we often worry that some of our present moral beliefs might be mistaken; that is, we don't just attribute moral error to other people and to ourselves at different times (say, in the past), but we feel that our present moral beliefs are subject to error.[1] Expressivists need an account of why this is (i.e., an account of *first-person present error*). The general expressivist strategy for addressing this problem comes from Simon Blackburn. Blackburn suggests that when I'm concerned that a moral belief of mine might be mistaken, what I'm really concerned about is whether I would revise my moral belief given certain improvements in my situation. For example, if I happen to value better information and coherence (in the expressivist sense of value), then when I worry that my moral belief that P might be mistaken, what I'm really worried about is whether I would revise my belief that P if I had access to better information and was being perfectly coherent. And Egan grants this much. He accepts that Blackburn's account seems to successfully explain ordinary moral discourse surrounding first-person present

1. Note that quite satisfactory accounts of the former abound, as Egan outlines in his paper. Expressivists often claim that when we attribute moral error to other people about a given moral matter, for example, we are simply *expressing* our own attitudes about that moral matter while recognizing the other person's conflicting attitudes. The same goes for scenarios in which we attribute moral error to *ourselves* in the past. In these scenarios, it is claimed, we again are *expressing our* own attitudes about whatever moral matter is at hand while recognizing our past self's conflicting attitudes.

error.[2]

Having outlined and accepted Blackburn's account of first-person present moral error, Egan argues that quasi-realist expressivism has implausible implications in certain cases of moral disagreement. While I will define quasi-realism in more detail below, for now it suffices to say that quasi-realism is a special kind of expressivist view in metaethics that attempts to earn the right to realist-sounding moral discourse while remaining fundamentally anti-realist and expressivist. Unlike naive expressivists, quasi-realists are all about accommodating ordinary, first-order[3] moral discourse.[4] It's this commitment, Egan thinks, that gets quasi-realists into trouble.

Egan's argument is fairly straightforward. First, he distinguishes between two kinds of beliefs: stable and unstable. A belief is stable, he stipulates, just in case no changes which the believer would recognize or endorse as improvements in their situation would lead them to revise their belief. Next, he asks us to consider what he calls a *fundamental* moral disagreement—that is, a disagreement where stable beliefs are involved on both sides. Suppose you and I are having some moral disagreement where I stably believe P and you stably believe not-P; according to Egan, this is a *fundamental* moral disagreement. At this point, he points out that quasi-realists should be committed to the following principle:

> FUNDAMENTAL FALLIBILITY: it's possible for other people's stable beliefs to be mistaken.[5]

Because the quasi-realist is especially committed to capturing ordinary moral discourse, and because moral disagreement is generally taken to imply moral error, the quasi-realist must admit that in cases of fundamental moral disagreement someone must be mistaken. And because fundamental moral disagreements are defined to be moral disagreements where people's stable beliefs in particular are at stake, quasi-realists have to grant FUNDAMENTAL FALLIBILITY.

Egan next reminds us of what Blackburn's expressivist account of first-person present er-

---

2. Simon Blackburn, "Truth and A Priori Possibility: Egan's Charge Against Quasi-Realism," *Australasian Journal of Philosophy* 87, no. 2 (2009): 210. https://doi.org/10.1080/00048400802362182.
3. First-order moral discourse has to do with questions about what is right and wrong, good and bad, and what we ought and ought not to do. It is a matter of *applied* or *normative ethics*. In contrast, second-order moral discourse is a matter of *metaethics*. It has to do with the nature of morality and questions about what moral sentences really mean and what their metaphysical status is.
4. Richard Joyce, "Moral Anti-Realism," in *Stanford Encyclopedia of Philosophy*, ed. Edward N. Zalta (Winter 2021 Edition), 24. https://plato.stanford.edu/entries/moral-anti-realism/.
5. Andy Egan, "Quasi-Realism and Fundamental Moral Error," Australasian Journal of Philosophy 85, no. 2 (2007): 213. https://doi.org/10.1080/00048400701342988.

ror is. Again, under Blackburn's account, what it means for my moral belief that P to be mistaken is for there to be changes that I would recognize and endorse as improvements in my situation that would lead me to abandon or revise P. As Egan points out, if this is granted, it seems to follow that my stable moral beliefs *cannot* be mistaken, because what it means for a moral belief of mine to be stable is for there to be *no* improvements that I would recognize as such which would lead me to revise my moral belief. So, Egan concludes, quasi-realists must also endorse the following principle:

> FIRST-PERSON IMMUNITY: I have an a priori guarantee against fundamental moral error.[6]

Now, it is these two principles that, Egan argues, seem to lead to implausible results. Taken together, FUNDAMENTAL FALLIBILITY and FIRST-PERSON IMMUNITY imply that in cases of fundamental moral disagreement, while the other person may be subject to fundamental moral error, I have an *a priori* guarantee of immunity against fundamental moral error. But this violates another plausible, intuitive principle:

> NO SMUGNESS: There isn't any sort of moral error to which others are subject, but against which I have an a priori guarantee of immunity.[7]

For Egan, this leaves quasi-realists in a bad place. Whatever happens, it seems like quasi-realists are going to have to give something up and thereby fail to capture something about ordinary moral discourse. If they deny FUNDAMENTAL FALLIBILITY, they will have to give up on the very thing that motivates quasi-realism—namely, earning the right to realist-sounding moral discourse. If they deny FIRST-PERSON IMMUNITY, then they'll have to give up on Blackburn's account of first-person present moral error—generating the need for a new expressivist account of error. If they deny NO SMUGNESS, they will fail to capture a common-sense, intuitive principle that our ordinary moral discourse endorses—it doesn't seem like we should ever have the kind of infallibility that quasi-realism guarantees us in cases of fundamental moral disagreement. All of this, Egan concludes, gives us good reason to reject quasi-realism.

II. The Problem

I will now argue that Egan's argument is unsound; the premise I will target in order to do

6. Ibid., 214.
7. Ibid., 210.

so is FIRST-PERSON IMMUNITY. More specifically, I will argue that FIRST-PERSON IMMUNITY is vague. Once this premise is made clearer in light of the aims and distinctive features of quasi-realism, Egan's argument no longer succeeds.

To provide some background for my objection, consider first the quasi-realist deflationary account of moral truth, facts, properties and beliefs. What sets quasi-realism apart from naive expressivism is that quasi-realists have a way to account for how and why it is that we say things like "it is a fact that abortion is wrong," "I believe that abortion is wrong," "abortion instantiates the property of wrongness," and so on. At least *prima facie*, these sorts of moral sentences seem to be better explained by realist, cognitivist views than by expressivist ones. Expressivist views admit of no moral facts, properties, truths, beliefs, and so on. Why, then, for expressivists, do we seem to talk as if these things exist? Quasi-realists have a special way to account for this part of our ordinary moral discourse: deflationism. Quasi-realists propose that when I assert "P is a moral fact" (or "P is true," "I believe P," "P has the property of wrongness," etc.), all I am doing is asserting "P" *simplicter*. When I say "it is a fact that abortion is wrong," for example, all I am doing is asserting "abortion is wrong." In other words, "abortion is wrong is a fact," for the quasi-realist, *if and only if* "abortion is wrong." These sentences are equivalent in meaning. This deflationism is what sets quasi-realism apart from other expressivist views.[8]

Returning to Egan's argument, quasi-realist deflationism illustrates how, given quasi-realism, moral sentences can be "read" in different ways. Consider the following moral sentence: "It's a fact that the death penalty is wrong." For the quasi-realist, this sentence can be "read" in two different ways. On the one hand, it can be read in an ordinary, ethical, realist sense, such that it will mean exactly what it says taken at face value—namely, that it really is just a fact that the death penalty is wrong. On the other hand, it can also be read in a metaethical, metasemantic sense that is explicitly aware of quasi-realism's deflationist machinery—namely, only that the death penalty is wrong and nothing more (which, remember, is just an expression of attitudes). Recall that this is what quasi-realism is all about—keeping our *ethics* intact while only changing the *metaethics* that is happening behind the scenes.

But how is this relevant to FIRST-PERSON IMMUNITY? Recall that, for Egan, FIRST-PERSON IMMUNITY is supposed to be true because, given Blackburn's account of moral error, it is built into the concept of stable moral beliefs that I can't be mistaken about them, because I know that, per their definition, there is no potential improvement in my situation that

8. Ibid., 207.

would lead me to revise them. I also know that I only have reasons to be concerned about my moral beliefs insofar as I might revise them given improvements in my situation. So, it seems like I have an *a priori* guarantee that my stable moral beliefs aren't mistaken. But here's where the distinction I drew between the different ways of reading moral sentences becomes relevant. What I want to ask Egan is this: just how am I supposed to understand "improvements in my situation"? When thinking about my stable moral beliefs, am I thinking about moral beliefs of mine that I would not abandon given any improvements in my situation where "improvements in my situation" is read in the ordinary or metaethical sense? That is, am I thinking about "improvements" in the ordinary, realist sense, or as improvements that *I personally* would recognize and endorse as such?

On the one hand, if read in the ordinary sense, "improvements in my situation" will just mean genuine, objective improvements in my situation in the realist sense. On this reading, FIRST-PERSON IMMUNITY is trivially true. Even the realist cannot deny that if no genuine, objective improvements in my situation would lead me to revise a moral belief of mine, then that moral belief of mine cannot be mistaken. This is the same sort of conceptual guarantee against moral error that one gets in conditionals like *if my moral belief that P was produced by my moral faculties when they were working perfectly, then my moral belief that P cannot be mistaken*. But now what's crucial about this reading is that it need not lead to a violation of NO SMUGNESS, since other people have this same guarantee of immunity against moral error (and I can have it about *them* as well). If no real, genuine, objective improvements would lead another person to revise their moral belief that P, then it seems obvious that their moral belief that P can't be mistaken. Hence, on this reading, we don't get a violation of NO SMUGNESS and Egan's argument is invalid.

That leaves reading "improvements in my situation" in a strictly metaethical sense. Read this way, "improvements in my situation" will be interpreted from a metaethical, metasemantic point of view, as changes in my situation that *I personally would recognize to be and endorse as improvements*. This follows from quasi-realism's deflationism and Blackburn's account of first-person present error. In the context of this metaethical machinery, "improvements" no longer functions as a sort of indexical that mimics realist terms; instead, it takes on the meaning of improvements that *I would recognize as such*.

Does this reading entail that I have an *a priori* guarantee that my stable moral beliefs aren't mistaken? Here, I think the quasi-realist has a story to tell. The quasi-realist can point out that often improvements that *I would recognize as such* turn out not to be improvements at all.

Sometimes I get improvement wrong—I know that *I* get it wrong. Moreover, I know that other people have dubious beliefs about what improvement constitutes, so I am inclined to treat my own beliefs about improvement with caution and scrutiny. Because I worry about being wrong about what constitutes real improvement, I must also worry about moral beliefs that would survive changes that I would recognize to be improvements. Thus, on this reading, I do not have an *a priori* guarantee of immunity against moral error where my stable beliefs are concerned, which means that FIRST-PERSON IMMUNITY is false and Egan's argument unsound. Since neither reading of FIRST-PERSON IMMUNITY yields a sound argument, Egan's argument should be rejected.

III. Potential Objections

One objection Egan might have to my argument is that reading "improvements in my situation" in the metaethical sense I described creates new problems for the quasi-realist. I argued that we often worry that changes that we would recognize as improving changes might not actually be improving changes at all, and that it is this that prevents us from having an *a priori* guarantee that our stable moral beliefs aren't mistaken. Put another way, we often worry that our *improvement-beliefs* (i.e., beliefs about what constitutes improvement) might be mistaken. But this generates the need for a quasi-realist account of how it is that our improvement-beliefs might be subject to error. Here, it doesn't seem like Blackburn's account will do. Blackburn's account of moral error is all about whether we would revise our moral beliefs if our circumstances were to improve, where "improve" is defined by our improvement-beliefs. Applying this account to evaluative error related to improvement, we would have to say that when we are concerned that our improvement-beliefs might be mistaken, we are really just concerned about whether our improvement-beliefs would survive our improvement-beliefs. But does this make sense? While I've successfully fended off Egan, it may seem I've also introduced a new and equally daunting challenge for quasi-realists, which can only be solved by coming up with an entirely new account of a certain kind of evaluative error.

To this I reply: is there really anything wrong with applying Blackburn's account of moral error to our improvement-beliefs? I think we can imagine cases in which it makes sense to worry about whether improvement-beliefs of ours would survive our improvement-beliefs. For example, suppose I have an improvement-belief that doesn't really cohere with my other improvement-beliefs, one of which happens to be coherence. In this case, it makes sense that I could revise one improvement-belief in light of the others. Another example is this: suppose I

really value better information and that getting such information would lead me to revise some of my improvement-beliefs. In this case, it also makes sense to talk about revising one improvement-belief in light of the others. So, I don't think quasi-realists need to develop a new account of certain evaluative errors anyway—Blackburn's account seems perfectly up to the task.

At this point, Egan might try to press the point and say that only a special kind of improvement-belief is relevant when we're thinking about our stable moral beliefs. When we're thinking about our stable beliefs, he might say, we aren't thinking about "improvement" read in the ordinary, realist sense nor "improvement" read metaethically as improvements that we would *currently* recognize as such, but "improvement" read as improvement that we would recognize as such *after having maximally considered our improvement-beliefs*. It's the stable moral beliefs that we'd have under these maximally considered improvements-beliefs, Egan might say, which afford us an *a priori* guarantee of immunity against moral error.

However, it's hard to see how this suggestion doesn't collapse back into the same problems that afflict the original argument. Am I to read "maximally considered" in the ordinary or metaethical sense? Read in the ordinary sense, "maximally considered" is too strong—I won't understand my maximally considered improvement-beliefs to be any different from genuinely and objectively correct improvement-beliefs. Read in the strictly metaethical sense, "maximally considered" is too weak—I will still be left worried that the improvements-beliefs that I would have after putting in my best efforts might be mistaken. So, it does not seem like this objection undermines my argument.

Finally, one might also be tempted to accuse my objection of being *ad hoc* because of how it deals with moral language. Is it really sensible to read and interpret moral sentences in different ways, as I've discussed? My response to this is simply to say that I don't think this objection really has anything to do with what I've argued here. Instead, it seems to be reiterating a different problem that quasi-realism seems to have—namely, its strange account of how we use moral language. That's an independent problem to take up with the quasi-realists and the deflationism they are committed to. All I am arguing here is that if quasi-realists can hold their ground with respect to those independent issues, then they have resources to deal with Egan's argument.

IV. Conclusion

In this paper, I defended quasi-realists from Andy Egan's argument against quasi-realism. I challenged one of the central premises from Egan's argument and responded to potential

objections to the claims I made throughout. I conclude that—at least with respect to this charge and others like it related to moral disagreement and error—quasi-realists have a promising strategy to work with. A careful analysis of how we use moral language in different ways and across domains (i.e., metaethics versus ethics) provides quasi-realists with a helpful tool when dealing with objections related to moral disagreement and error. What are the implications? If what I've argued is correct, and if Egan's argument happens to serve as a strong representative of arguments from moral disagreement and error against quasi-realism, then it seems quasi-realists should not be worried about the metaethical data to be explained concerning disagreement and error.

Bibliography

Blackburn, Simon. "Truth and A Priori Possibility: Egan's Charge Against Quasi-Realism." *Australasian Journal of Philosophy* 87, no. 2 (2009): 201–213. https://doi.org/10.1080/00048400802362182.

Egan, Andy. "Quasi-Realism and Fundamental Moral Error." *Australasian Journal of Philosophy* 85, no. 2 (2007): 205–219. https://doi.org/10.1080/00048400701342988.

Joyce, Richard. "Moral Anti-Realism." In *Stanford Encyclopedia of Philosophy*, edited by Edward N. Zalta. Fall 2018. https://plato.stanford.edu/entries/moral-anti-realism/.

The Implausibility of Artificial Intelligence

Baptiste van Gaver

Many of the leading theories in the philosophy of mind either implicitly or explicitly admit that the creation of artificial consciousness is possible. This constitutes an important explanation of the existing enthusiasm at the interstice of consciousness studies and artificial intelligence research, and of the current optimism surrounding the emergence of artificial consciousness. In this paper, I seek to articulate and evaluate some of these leading theories in order to show that this optimism rests on some philosophically unfounded assumptions and that the emergence of artificial consciousness is implausible and worthy of skepticism.

Keywords: artificial intelligence, consciousness, mind, cognition, physicalism, reductionism

I. Introduction

Most of the progress in artificial intelligence research has been fueled by an underlying optimism surrounding human beings' ability to create conscious experience in an artificial system. This optimism is itself grounded in certain philosophical assumptions about the nature of consciousness and its relation to matter. In this paper, I will argue that we ought to be skeptical as to the possibility of artificial consciousness, because none of the leading explanations in its favor currently seem promising. First, I will attempt to provide clear definitions of 'consciousness,' 'the mind,' and 'artificial consciousness' in order to better grasp exactly what it is that arguments in favor of the possibility of artificial consciousness are claiming. Next, I will provide both an overview of the main views in the philosophy of consciousness that admit the possibility of artificial consciousness and a critical evaluation of their central claims and assumptions. Finally, I will explain why the best-supported philosophical attitude one could adopt towards artificial consciousness is one of skepticism towards its very possibility.

II. Consciousness and Cognition

The distinction between consciousness and cognition is fundamental to any attempts to determine whether the creation of artificial consciousness is possible. It is crucial to clearly define the terms 'consciousness' and 'mind' to better understand what is meant by 'artificial consciousness.' Phenomenologically speaking, consciousness is one of the most familiar aspects of human existence, and yet, there has been prevalent disagreement and confusion about what the term 'consciousness' designates, in the academic context and beyond. Some have spoken of consciousness in general terms, encompassing a wide range of mental processes and properties, while others have adopted a narrower understanding of consciousness, focusing exclusively on the subjective, first-person character of human experience. I would argue that consciousness is best defined in narrower terms, precisely as the qualitatively subjective awareness one has of oneself and of the world. More specifically, consciousness is one aspect of human mentation, and as such it should not be confused with other aspects of the mind.

This is in line with David Chalmers' assertion that there are two distinct conceptions of the mind, one phenomenal and one psychological. He defines the former as "the concept of mind as conscious experience, and of a mental state as a consciously experienced mental state" and the latter as "the concept of mind as the causal or explanatory basis for behavior," where "a state is mental…if it plays the right sort of causal role in the production of behavior." Chalmers therefore makes a distinction between the subjective character inherent to mental activity on the one hand,

and the causal, functional character of mental states on the other. This distinction is important because it suggests that each conception of the mind calls for a distinct explanatory paradigm. As Chalmers points out, "the phenomenal and the psychological aspects of mind have a long history of being conflated,"[1] and a consequence of this conflation is that many thinkers have studied the phenomenal aspect of the mind through an exclusively psychological framework of description and explanation, which is a category-mistake.

An alternative way of framing the distinction between first-person phenomenal consciousness and the cognitive and causal aspects of the mind can be found in Ned Block's distinction between "phenomenal consciousness" and "access consciousness." Block writes that "phenomenal consciousness is experience" and that "the phenomenally conscious aspect of a state is what it is like to be in that state," whereas "access-consciousness…is availability for use in reasoning and rationally guiding speech and action."[2] Thus, like Chalmers, Block distinguishes the purely subjective and qualitative aspect of mental states (what it feels like) from their representational, causal and intentional aspects, which are constitutive of cognition.

However, there is a difference between Chalmer's "psychological concept of mind" and Block's access consciousness: Chalmers recognizes that a mental state can be both phenomenal and psychological (depending on the explanatory lens through which it is viewed), while Block seems to claim that phenomenal consciousness and access consciousness interact but are in some sense independent of each other. What both Chalmers and Block contribute to the effort to clarify our understanding of 'consciousness' is the idea that the subjective quality of mentation exists independently of its causal, functional and intentional dimensions. In addition to what these authors have said, I argue that the term 'consciousness' should refer exclusively to the subjective quality of experience, and that instead of speaking of "access consciousness" to refer to cognitive and intentional processes, one ought to only speak of 'cognition.'

Having clarified what I mean by 'consciousness' as opposed to 'cognition,' it is now possible to explain why I choose to speak of "artificial consciousness" and not of "artificial intelligence." Intelligence, which can be defined as the abstract ability to acquire information and apply it rationally and intentionally, is constitutive of cognition rather than consciousness. It *feels like* something to use one's intelligence—intelligence cannot merely be equated with conscious-

1. David J. Chalmers, "Two Concepts of Mind," in *The Conscious Mind: In Search of a Fundamental Theory* (New York: Oxford University Press, 1996), 11-12.
2. Ned Block, "On a confusion about a function of consciousness," *Behavioral and Brain Sciences* 18, no. 2 (1995): 227. https://doi.org/10.1017/S0140525X00038188.

ness because phenomenologically speaking, human beings have a subjective awareness of their intelligence which feels distinct from intelligence itself. The two must be distinguished conceptually. Therefore, asking whether artificial intelligence can be conscious is representative of a general historical confusion between intelligence and consciousness.

It should be noted that claiming that there is a distinction between artificial intelligence and artificial consciousness does not necessarily commit us to the idea that consciousness is not in some sense constitutive of intelligence. Rather, the distinction merely follows from the idea that, methodologically speaking, consciousness and cognition cannot be approached the same way. Distinguishing between the two allows for a clearer idea of the topic of artificial consciousness. The latter is best defined as a qualitatively subjective, first-person experience that is designed and created by human beings, and questions about artificial consciousness must be distinguished from those about artificial intelligence.

III. Reductive Physicalist Views: Identity Theory and Functionalism

I will now summarize the main views in the philosophy of mind which either explicitly or implicitly admit the possibility of artificial consciousness. I contend that all these leading views face significant philosophical challenges. More particularly, I argue that all these views involve an explanatory gap grounded in the theoretical conflation of cognition and consciousness. The first set of views to be examined are those that fit within the reductive physicalist approach. Generally, reductive physicalism reduces mental states and processes to physical states and processes. The ontology of reductive materialism is materialist in the sense that it posits matter as the ultimate substratum underlying the world, and therefore the human body and brain as well. I will take up the identity theory and functionalism as two of the most prominent reductive physicalist approaches.

Identity theorists posit a strict relation of identity between mental states and physical states. Notably, David Lewis writes that the identity theory "is the hypothesis that…every experience is identical with some physical state," and, more specifically, "with some neurochemical state."[3] The first premise of Lewis' argument in favor of the identity theory is that experiences play a causal role in motivating brain states and behavior; he writes that "the causal role of an experience is expressible by a finite set of conditions that specify its typical causes and…typical

3. David K. Lewis, "An Argument for the Identity Theory," *The Journal of Philosophy* 63, no. 1 (January 1966): 17. https://www.jstor.org/stable/i335581.

effects under various circumstances."[4] The second premise of his argument for the identity theory is that of the "plausible hypothesis that there is some unified body of scientific theories, of the sort we now accept, which provide a true and exhaustive account of all physical phenomena."[5] Interestingly, Lewis claims that this second premise is not an ontological thesis in itself, but rather that it only denies that we ever need to explain physical phenomena by nonphysical ones. What he is claiming, in essence, is that all mental states are fundamentally causal and that, given the accuracy of our current materialist scientific paradigm, mental states should be described and explained in terms of the same causal relationships which we observe in the universe at large. A *prima facie* criticism of Lewis' argument would be to point out that while he speaks of "experiences," he does not directly and explicitly account for the subjective, first-person character of experiences. However, any committed identity theorist would have to respond that this subjective character is a sort of illusion—that all there ultimately is to the mind is reducible to objective physical facts. In principle, identity theory admits the possibility of artificial consciousness, by reducing consciousness to physical states or properties which could hypothetically be replicated in an artificial system.

Functionalism, which emerged in opposition to the identity theory as an alternative reductive physicalist conception of the mind, asserts that "rather than being identical to any specific neural process, [mental experiences] could supervene on any number of different neural processes."[6] Functionalists such as Daniel Dennett argue for a substrate-independent understanding of mental experiences, in which mental states are defined in terms of the functional role they occupy. Dennett defines a cognitive function as a kind of internal disposition which provokes or triggers other specific mental events. As indicated by Thompson, Dennett argues that "all conscious experiences are embedded in a complex network of dispositions [where] their roles in this network distinguish experiences from each other."[7]

Furthermore, unlike Lewis, Dennett explicitly tackles the issue of the subjective character of experience, but ultimately agrees with the identity theorist insofar as he claims that while it seems like experiences are characterized by intrinsically defined qualia, these qualia are nothing but illusions which do not really exist. Instead, for functionalists such as Dennett, to say that an experience is conscious is to "point to its functional role in the cognitive processes that make up

4. Ibid., 20.
5. Ibid., 23.
6. David L. Thompson, "What Consciousness Is: The Multiple Drafts Model," in *Daniel Dennett* (London: Continuum, 2009), 35.
7. Ibid., 39.

our perceptual, motor, reflective, and linguistic capacities."[8] While functionalism claims that the functional nature of mental experiences defines the latter independently of any specific and definite substrate, it still requires a substrate in order to instantiate the functional organization giving rise to mental events in the first place. Functionalism therefore inscribes itself within the sphere of reductive physicalism insofar as it reduces mental experiences to a physically instantiated complex of cognitive functions that are causally related to one another. This view also implies that the creation of artificial consciousness is possible, insofar as it acknowledges that an artificial system could instantiate mental events and properties, including the qualitative character of experience, by having a functional organization equivalent to that of a human mind. However, both functionalists and identity theorists would deny that artificial systems could experience qualia in the first place; they would rather claim that such systems could in theory be under the impression that they experience qualia, in the same way that humans do.

IV. Jackson's Epiphenomenalism and Nagel's Panpsychism

I argue that reductive physicalist theories of the mind involve a significant explanatory gap grounded in a problematic understanding of consciousness, and in questionable ontological assumptions. Most notably, reductive physicalism as a whole faces what David Chalmers has famously called the "hard problem of consciousness," the idea that physicalist theories are unable to explain the qualitatively subjective character of experience in objective physical terms. Reductive physicalist thinkers like Lewis or Dennett begin with the naturalist assumption that the objective physical world of matter is the ontological substratum of the mind, and then try to fit mental experiences within the materialist framework they have assumed. By committing to an objectively materialist explanatory framework, which describes the world in causal and functional terms, they conclude that the subjective character of experience is an illusion. However, this reasoning is fundamentally flawed. The hard problem of consciousness only emerges because reductive physicalists do not properly distinguish between cognition and consciousness, attempting to explain the latter with an explanatory paradigm that is proper to the former.

Although reductive physicalism seems to be the approach that most readily admits of the possibility of artificial consciousness, there are alternative views which lend themselves to this while sharing the same conceptual and methodological shortcomings. Views such as non-reductive physicalism, property dualism, and panpsychism are also grounded in an improper under-

8. Ibid., 45.

standing of the concept of consciousness. In discussing these views, I will primarily consider Frank Jackson's theory of epiphenomenalism and Thomas Nagel's panpsychism as representative of these two families of perspectives on consciousness.

Jackson's central claim is that "there are certain features of the bodily sensations [and of] certain perceptual experiences, which no amount of purely physical information includes."[9] These features are the individual instances of subjective conscious experience known as qualia. Jackson argues that qualia "are such that their possession or absence makes no difference to the physical world," and that, as a whole, "the mental is totally causally inefficacious."[10] Jackson therefore accepts the idea that the subjective character of experience cannot be explained through a physicalist causal framework, because while qualia might be caused by physical events, they themselves do not cause any physical events. Jackson is therefore a property dualist in the sense that he acknowledges the existence of two distinct kinds of properties, physical properties and the subjective property associated with qualia (note here that not all property dualists are necessarily epiphenomenalists).

The objection which Jackson brings up against reductive physicalists seems to be exclusively against their methodological arguments with regards to the explanation and description of mental content. Indeed, although he is careful not to make a category-mistake by recognizing that the subjective character of experience cannot be explained in objective terms, Jackson still ultimately grounds qualia in physical phenomena, and in that regard he could be considered a non-reductive physicalist. His epiphenomenal understanding of consciousness therefore seems to admit the conceptual possibility of artificial consciousness, in the sense that an artificial physical system could give rise to qualia if the proper physical causes are present.

However, my main objection against epiphenomenalism is a version of the 'explanatory gap' objection: epiphenomenalists such as Jackson give no concrete indication as to how mental properties emerge out of physical causes; the causal link posited seems mostly hypothetical and unsubstantiated. Epiphenomenalism faces this explanatory gap because it is based on the idea that consciousness has physical causes. While epiphenomenalists acknowledge that the subjective character of consciousness must be distinguished from the physical component of the mind, they are unable to explain how physical properties give rise to mental properties.

A different theory formulated against the reductive physicalist refutation of the existence

9. Frank Jackson, "Epiphenomenal Qualia," *The Philosophical Quarterly* 32, no. 127 (1982): 127. https://doi.org/10.2307/2960077.
10. Ibid., 134.

of qualia is panpsychism. This view can be formulated a variety of different ways, but I will focus on Nagel's understanding of panpsychism in order to articulate some of the view's central claims. Nagel defines panpsychism as "the view that the basic physical constituents of the universe have mental properties, whether or not they are part of living organisms."[11] Nagel characterizes it as an explicitly non-reductionist approach that adopts realism regarding the existence of qualia. Thus, it is an approach which explicitly refutes the reductive physicalist's conclusions.

The structure of Nagel's argument for panpsychism is as follows: (1) every living organism is a complex material system; (2) mental properties are not implied by physical properties alone; (3) mental properties are nevertheless part of the organism, i.e., they exist; (4) there are no emergent properties of complex systems; therefore, (C) the constituents of that physical system "must have nonphysical properties from which the appearance of mental properties follows."[12] Nagel's argument for panpsychism is hence built around a refutation of both the reductive materialist belief in the illusory nature of qualia and the epiphenomenal view of emergence. Importantly, Nagel's first premise shows that his panpsychist theory fully embraces ontological materialism. However, it is a conception of materialism which recognizes matter and consciousness as co-primordial—neither existed prior to the other; rather, they are inextricably linked.

A central notion in Nagel's conclusion is that of 'combination': the qualitatively subjective experience of an individual organism is constituted by a combination of the nonphysical (mental) properties of the elementary material constituents from which the organism is itself constituted. It therefore seems that, theoretically, Nagel could concede that the appropriate combination of elemental material constituents would give rise to consciousness, even in artificial systems, rendering his theory open to the possibility of artificial consciousness. Panpsychists often claim that their approach resolves the "hard problem of consciousness."

But in reality, it may produce an equally hard problem, the "subject-combination problem." The problem is that as of now, the combination of subjects is an unexplainable process. Bernardo Kastrup argues that "to conclude that…the consciousness of a living being…is made up of a combination of lower-level inanimate subjects requires an extra logical step for which… there is no justification." Kastrup goes on to argue that "like the 'hard problem' faced by physicalism," the combination problem "is not grounded in empirical reality, but in the internal logi-

11. Thomas Nagel, "Panpsychism," in *Mortal Questions* (Cambridge: Cambridge University Press, 1979), 181.
12. Ibid., 182.

co-conceptual structure of [panpsychism] itself."[13] His persuasive objection is that, just like the 'hard problem,' the combination problem exists, not because of empirical facts about consciousness, but because of logical fallacies involved in the assumptions about the nature of consciousness from which the problem emerges.

In essence, the 'subject-combination' problem is just another form of the explanatory gap faced by all the theories I have discussed in this paper. Each theory, including panpsychism, encounters difficulties when trying to explain how the unitary, subjective dimension of conscious experience emerge from physical facts. I have argued that such methodological problems arise because of a problematic understanding of the ontological status of conscious first-person experience. Any theory which admits of the possibility of artificial consciousness necessarily implies that consciousness can be explained through a physicalist causal framework, insofar as it must be designed and created by conscious beings. Therefore, I would conclude that the best attitude one could adopt towards artificial consciousness is to be highly skeptical of its possibility, considering that any explanation in its favor involves problematic explanatory gaps.

V. Conclusion

The theories I have presented constitute the basis for most contemporary arguments in favor of the possibility of artificial consciousness. I have argued that each of these theories involves an explanatory gap grounded in a fallacious understanding of the ontological status of consciousness. I therefore concluded that the best attitude one could adopt towards the possibility of artificial consciousness is to be highly skeptical, because the major explanations as to why this possibility might be realized are not philosophically convincing. As a final note, I would even suggest that one ought to be skeptical as to the validity of the materialist ontological paradigm as a whole. The first-person character of consciousness is fundamental to human existence, and if materialism is unable to explain it without encountering significant explanatory gaps, it seems necessary to actively search for alternative ontologies. The question of artificial consciousness is therefore of pressing importance because it challenges basic unresolved intuitions about the world and our place within it as conscious beings.

13. Bernardo Kastrup, "The Quest to Solve Problems That Don't Exist: Thought Artifacts in Contemporary Ontology," *Studia Humana* 6, no. 4 (2017), 49. doi: 10.1515/sh-2017-0026.

Bibliography

Chalmers, David J. "Two Concepts of Mind." In *The Conscious Mind: In Search of a Fundamental Theory*, 3-31. New York: Oxford University Press, 1996.

Block, Ned. "On a confusion about a function of consciousness." *Behavioral and Brain Sciences* 18, no. 2 (1995): 227-247. https://doi.org/10.1017/S0140525X00038188.

Block, Ned. MacDonald, Cynthia. "Phenomenal and Access Consciousness: Consciousness and Cognitive Access." In *Proceedings of the Aristotelian Society* 108, no. 1 (October 2008): 289-317. https://doi.org/10.1111/j.1467-9264.2008.00247.x.

Jackson, Frank. "Epiphenomenal Qualia." *The Philosophical Quarterly* 32, no. 127 (1982): 127-136. https://doi.org/10.2307/2960077.

Kastrup, Bernardo. "The Quest to Solve Problems That Don't Exist: Thought Artifacts in Contemporary Ontology." In *Studia Humana* 6, no. 4 (2017): 45-51. doi: 10.1515/sh-2017-0026.

Lewis, David K. "An Argument for the Identity Theory." *The Journal of Philosophy* 63, no. 1 (January 1966): 17–25. https://www.jstor.org/stable/i335581.

Nagel, Thomas. "Panpsychism." In *Mortal Questions*, 181-195. Cambridge: Cambridge University Press, 1979.

Thompson, David L. "What Consciousness Is: The Multiple Drafts Model." In *Daniel Dennett*, 34-57. London: Continuum, 2009.

Marx's Proletariat and the Forcibly Unemployed: Are They One and the Same?

Karina Vandenhoven

Marx's conception of the proletariat has been interpreted and applied to various emancipatory struggles since the publication of *Capital*. Especially in the past century, there has been an attempt to situate many different revolutionary narratives within Marx's conception of the proletariat. One such narrative has been the rise in mass unemployment and subsequent organization of the unemployed in the last century. This rise has created a new debate within Marxism; namely, whether those who are unable to work for formal wages can be considered part of the proletariat. For the purposes of this essay, the "forcibly unemployed" are defined as those who, by virtue of certain physical or mental characteristics, are deemed undesirable as a labourer in the workplace and hence not hired. My argument is that the forcibly unemployed can and should be considered part of the proletariat.

Keywords: Proletariat, forcibly unemployed, labour, labour-power, Marxism, capitalism, emancipation, value, workers' organization

I. Introduction

Marx defines members of the proletariat as those who sell their labour on the labour market for wages to the capitalist.[1] By this definition, anyone who can never, for various reasons, sell their labour, cannot be considered part of the proletariat. Given that Marx defines labourers as the revolutionary agent, left out of the proletariat then are those forbidden from entry into the factory or other formal spheres of work, such as those who are unable to work due to disability or discrimination. For the purposes of this paper, the "forcibly unemployed" are defined as those who are unable to find work due to physical or mental characteristics that render them undesirable as potential laborers. This may include those with physical or mental disabilities, socially marginalized segments of the population, and others. As a result, in referring to the "forcibly employed," this paper does not intend to include workers barred from work due to market pressures (e.g., lack of work or external shocks). In general, Marx's definition of the proletariat seems hyper-focused on the factory workplace itself, while seemingly paying insufficient attention to the ways that the capitalist system influences life and behaviour outside of the factory walls.

So, this paper's purpose is to first and foremost highlight a potential problem in Marxist thought; namely, that the forcibly unemployed may not be considered part of the proletariat (the working class) and therefore the main revolutionary agent which Marx foresaw as overthrowing capitalism or otherwise altering it. As the number of dispossessed peoples increases—due to increased automation on an unprecedented scale and the advent of artificial intelligence—and the historical capitalistic process of 'creative destruction' begins to break down,[2] Marx's traditional explanatory value risks becoming obsolete. The Marxist's main revolutionary agent—the proletarian—will comprise an increasingly smaller percentage of those with legitimate revolutionary struggles within the capitalist system.[3]

---

1. Karl Marx, *Capital: A Critique of Political Economy, Volume I*, ed. Alexander Chepurenko. Vol. 35 of Marx/Engels Collected Works (Progress: Moscow, 1996), 119.
2. This term is borrowed from Joseph Schumpeter, but cf. Robert L. Heilbroner, 160 and 164. Heilbroner describes Marx's preoccupation with the idea that capitalism always brings itself closer to its own inevitable destruction in two ways: first, through constant economic crises that are necessary to its existence; and second, through the creation of an increasingly disaffected class of people—the proletariat.
3. As 'low-skill' jobs become increasingly automated, workers increasingly become unemployed with very little hope of re-entering the workforce. Even "high-skill" jobs are at risk as artificial intelligence substitutes human intelligence. While it is true that, historically speaking, this has more or less been happening every couple of generations as inventions erase old jobs and create new ones, the modern industrial revolution doesn't seem to provide that much space for humans to re-enter the workforce in considerable numbers. It is not simply that there will not be enough jobs for the number of people seeking them, but that the jobs which do exist will be highly technical and limited to people with very specific qualifications.

In Section I, we will introduce the problem, asking why it is a problem and what motivations we might have in seeking to address it. In Section II, we will explore three potential readings of Marx which each present a *textual solution* to the stated problem, additionally highlighting how these solutions could apply to the real world. Section III will show how the third solution, which combines the proletariat and forcibly unemployed into one group, is the best of the three options, bringing Marxist thought into the 21st century. In Section IV we will conclude and summarize this paper's main points.

II. Why Is This a Problem?

In asking why this supposed problem presents itself, we might also be inclined to ask what motivation there is to pose the question in the first place. In other words, why should we care about the forcibly unemployed? Well, it seems that the forcibly unemployed suffer under the capitalist system, often being subjected to prolonged poverty, illness, and death at least to the same extent as the proletariat. It seems hard to imagine a scenario in which the overthrow of capitalism emancipates the workers but simultaneously leaves the forcibly unemployed to their restless misery. In light of this, there is a sense that the forcibly unemployed might also have a role to play in the anti-capitalist revolution.

This will be the focus of this paper; namely, determining whether the forcibly unemployed, as defined here, can and should be included as part of Marx's conception of the proletariat, and thus, by extension, the Marxist revolutionary agent.

Turning to textual explanations of why the exclusion of the forcibly unemployed from the proletariat is a problem, one begins with Marx's assumption that capitalism's primary class contradiction is that between the worker and the capitalist. Marx writes, "It came to pass that [the capitalist] accumulated wealth, and [the worker] had nothing to sell except their own skins."[4] But what exactly does selling one's skin entail? There is value in examining whether those outside of the factory have the ability to sell their skins in the way that Marx describes it.

A second issue that arises in the text is whether the forcibly unemployed even exist as a separate group, or whether they are contained under Marx's description of the unemployed or those with extremely irregular work.[5] Marx's conception of those who have never worked is

4. Marx, *Capital, Vol. I*, 507.
5. Ibid., 443-444. Marx defines the "industrial reserve army" essentially as the unemployed, a surplus labouring population that exists as a necessity in the capitalist system as a disposable human material always ready for exploitation.

very different from more modern conceptions. For example, he fails to make any mention of the social production of identities apart from class. Scholars such as James Boggs, Selma James, and Laura Jaffee have all demonstrated how unemployment can come about through the construction of identities such as 'disabled,' 'black,' 'woman,' and so on, lodging the relevant individuals into the category of "undesirable worker."[6] While much headway has been made materially to remove the categories of womanhood and racialization from undesirability, almost none has been made for the disabled. This becomes a big problem for Marx, as he defines all those who cannot work as "paupers" inherently tied to criminality.[7] There seems to be a need to explain whether those forces that oppress the worker are the same as those that oppress the forcibly unemployed—and whether there is any possible way for these groups to meaningfully organize.

Before addressing three potential solutions for this problem, we should first address a solution that we will not explore here; namely, including *every* oppressed person under the proletarian umbrella. The forcibly unemployed and the proletariat seem to be fighting for the same thing—to be able to live a life of subsistence without the exploitation of their labour.[8] In having chiefly economic goals, the two share a strong similarity. But the goals of other marginalized groups include social or political liberation, which are not the focus of this paper. While these groups would likely work together in certain revolutionary capacities, we will not discuss varieties of emancipation other than economic emancipation here.[9] Our scope is limited to whether, within a Marxist framework, we can amalgamate the workers and the forcibly unemployed into one singular revolutionary agent—with respect to *economic* emancipation.

To this end, we will look at three potential readings we can give of Marx. Each reading will be split into two parts: first, a textual solution to the problem that we can draw from Marx's writings. Second, whether that response can apply to the real lives of workers and the forcibly unemployed, rather than having it fall victim to 'Armchair Revolution.'[10] The three readings of Marx are as follows:

---

6. Cf. James Boggs, *The American Revolution: Pages from a Negro Worker's Notebook* (Monthly Review Press: New York, 2009), 51, https://libcom.org/library/american-revolution-pages-negro-workers-notebook; cf. Selma James, "Sex, Race, and Class," in *Sex, Race, and Class—The Perspective of Winning: A Selection of Writings, 1952-2011*, (New York: PM Press, 2012), 94-5; and cf. Jaffee, "Marxism and Disability Studies," in *Encyclopedia of Educational Philosophy and Theory*, ed. Michael A. Peters, 1-6 (2016), 4. https://doi.org/10.1007/978-981-287-532-7_279-1.
7. Marx, *Capital, Vol. I*, 450.
8. Jaffee, "Marxism and Disability Studies," 6.
9. However, it is also clear that social emancipation and economic emancipation are closely linked.
10. This refers to an emancipatory theory that is so convoluted and difficult to understand that no member of the group that would be conducting the revolution could ever actually apply it in practice.

(1) *The forcibly unemployed cannot be added to Marx's conception of the proletariat, and a new theory capable of amalgamating the two is required.* — This reading will stress Marx's assertion that labour is related specifically to workers' subsistence and the ability for them to use their labour for factory production. In terms of its potential application, this reading emphasizes the lack of consideration that Marx displays towards the forcibly unemployed, making it difficult to apply traditional Marxist thought to their emancipation.

(2) *The forcibly unemployed and the proletariat could be amalgamated, but the current capitalist system has made their organization materially impossible.* — This reading will center on Marx's description of the battle between the employed and the industrial reserve army or surplus population (the unemployed).[11] Applying this to the real world, it will be demonstrated that the forcibly unemployed are often constrained to tools of workers' rights, such as unions. The presumed necessity for the forcibly unemployed to work within the confines of workers' organization may be misguided.

(3) *The forcibly unemployed can be added to Marx's concept of the proletariat.* — In this paper, we will focus on Marx's conception of labour as it relates to James Boggs' discussion of the unemployed. Works such as that of Laura Jaffee will additionally demonstrate the ways in which Marxist thought and disability theory have intermingled to present a shared goal of revolutionary emancipation. As a result, the forced separation of the forcibly unemployed and the workers may prove surmountable.

III. Readings of Marx and Proposed Solutions

(A) *Solution 1: The Forcibly Unemployed Cannot Be Added to Marx's Proletariat, and This Is a Problem*

The first reading involves examining the way in which Marx describes labour as a process inherently tied to the factory and bourgeois economic system. Indeed, Marx defines the factory as the site of class struggle.[12] His economic analysis relies heavily on the description of capitalism as occurring in the labour-market and inside the factory walls.[13] Marx additionally considers the fact that the worker's labour is directly tied to his subsistence, in that he must work

11. It is acknowledged here that the surplus population that Marx describes and the forcibly unemployed as understood here are not the same. However, 'undesirable' labourers can often have the same effect as the surplus labouring population in depressing the wages of the working class.
12. Marx, *Capital, Vol. I,* 10.
13. Ibid., 119.

to survive.[14] This relates to his description of the capitalist system as a binary between workers and capitalists, writing, "[worker-capitalist relations] arise only when the owner of the means of production and subsistence finds the free worker available, on the market, as the seller of his own labour-power."[15] Here we find pushback against a conception of labor or labor-power as anything that goes beyond a worker and his wages, and thus by extension *against* the assertion that the system is *not* binary.

In terms of applying Marx's binary to the real world, there arises with the first reading the problem that Marx has a very particular view of the role that the unemployed play in the capitalist system, categorizing all but one of his four kinds of 'unemployed peoples' as having done some form of formal work at one time, even if the work in question was incredibly irregular. Those who do not, Marx calls "paupers," describing them as "vagabonds, prostitutes, in a word, the 'dangerous' classes."[16] This definition is problematic in that the association of forced unemployment with criminality is not necessary. Marx's categorization, then, does little to account for "outsiders"[17] or those who can never work in the factory and yet are not criminals. This renders it difficult to apply his work to the possibility of revolutionary action by the forcibly unemployed.

Marx's inability to acknowledge that there may be forcibly unemployed people who are not criminal lends itself to the suggestion that his theory is incomplete, and that we thus require a new theory unburdened by his prejudicial observations. Consequently, this first solution is not so much a solution *per se*, but it recognizes that there is a problem in applying Marx's writings to the situation of today's forcibly unemployed. The problem, however, is that simply stating that Marx failed to consider those with mental or physical disabilities, or that he misidentified a group, does not give us or the anti-capitalist revolutionary much determinate content upon which to act in solving the problem of the exclusion of the forcibly unemployed, or even in envisioning what a solution would look like.

(B) *The Forcibly Unemployed Could Theoretically Be Added to Marx's Proletariat, But Not in Practice*

The second solution entails examining how Marx describes his "surplus population"—that is, the generally unemployed—as being in competition with the proletariat. This competition

14. Ibid., 47.
15. Ibid., 119.
16. Ibid., 450.
17. Boggs, *The American Revolution,* 50.

creates conflict of benefit to the capitalist. It is first important to note that in this definition, Marx is only referring to those unable to work due to the supply of jobs, rather than by virtue of specific 'undesirable' traits. However, there is historical evidence of 'undesirable' groups of labourers still causing conflict and competition among the working class. One of the key examples of this in Marx's time was slavery and serfdom, especially in the American South. Slaves, per Marx's account, produced a large amount of surplus labour in the production of cotton.[18] Indeed, "worker shortages" were often used by European moguls to justify the capture and enslavement of Indigenous peoples to avoid paying domestic workers better wages.[19]

Since the advent of the capitalist system, there has loomed the threat of formal work being taken up by informal labourers and the forcibly unemployed. Though labour and accommodation laws have made this less common in the modern age, manifestations of this phenomenon still exist today. For example, in the United States, an estimated 420,000 people with disabilities are paid below minimum wage, averaging around $2.15 an hour. This is due to a loophole in the 1938 Fair Labors Standards Act that allows employers to justify these wages by claiming that they provide workers with vocational training and jobs for those who would otherwise never find one. Most of the time, these "gigs" are not substantial work, but rather programs that leave the forcibly unemployed with no real career path. Disabled people have even become an attractive alternative to formal labourers in the realm of piecemeal work, as they often have no other options.[20] A comparison to the surplus labouring population is thus beneficial, as it describes a similar dynamic of conflict and competition to that which occurs between workers and the forcibly unemployed.

According to Marx, "if a surplus labouring population is a necessary product of accumulation or of the development of wealth on a capitalist basis, this surplus population becomes, conversely, the lever of capitalistic accumulation, nay, a condition of existence of the capitalist mode of production."[21] The necessity of the existence of the unemployed is something that pushes down on the worker and is overall beneficial to the capitalist:

> The overwork of the employed part of the working class swells the ranks of the [unemployed], whilst conversely the greater pressure that the latter by its competition exerts

18. Marx, *Capital, Vol. I*, 164.
19. Jon Schwarz, "The Business Class Has Been Fearmongering About Worker Shortages for Centuries," *The Intercept: Voices*, 7 May 2021. https://theintercept.com/2021/05/07/worker-shortage-slavery-capitalism/.
20. Sarah Kim, "The Truth of Disability Employment that No One Talks About." *Forbes*, 24 October 2019. https://www.forbes.com/sites/sarahkim/2019/10/24/sub-minimum-wages-disability/?sh=49f89c95c22b.
21. Marx, *Capital, Vol. I*, 444.

> on the former, forces these to submit to overwork and subjugation under the dictates of capital. The condemnation of one part of the working class to enforced idleness by the overwork of the other part, and the converse, becomes a means of enriching the individual capitalists.[22]

Although this passage refers to the generally unemployed, it could certainly apply to the 'undesirable' workers as well, given that the presence of both depresses the wages and working conditions of the working class. There is then somewhat of a contradiction in Marx. Either this analysis can apply to the forcibly unemployed—in which case conflict between the two groups is a necessity to the capitalist system—or he is once again failing to consider the tangible effects that the forcibly unemployed can have on the economic system as informal workers despite their 'undesirable' characteristics.

In applying this in practice, one could argue that it is not necessary for the struggles of the forcibly unemployed to be lumped in with those of the workers for the former to attain their emancipatory goals. Rather, we might have, as Boggs would say, a "revolutionary force or army of outsiders and rejects who are totally alienated from society."[23] A plausible instance of this phenomenon might be the fact that the forcibly unemployed are not allowed to join workers' unions. These groups often act as one of the strongest proponents and forces for better conditions in the workplace. However, without any means of meaningful participation in industrial organizing, the forcibly unemployed are often relegated to afterthought, or worse, no consideration at all.

It could perhaps be said that we do not need to view the forcibly unemployed to cooperate with Marx's proletariat—that they could be the revolutionary agents themselves. This separation of the proletariat (Marx's revolutionary agent) and the growing number of forcibly unemployed (the *nouveau* revolutionary agent) into two forces, and the clash that might subsequently arise, could give the anti-capitalist revolution the momentum it needs.[24] So, it might make more sense for the proletariat and forcibly unemployed to be regarded as separate entities with distinct reasons and methods for emancipating themselves from the capitalist system.

(C) *The Forcibly Unemployed Can Be Added to Marx's Proletariat*

Given that one is unable to get a full picture of the economic role of the forcibly unemployed from Marx, it is beneficial here to examine more recent Marxist perspectives. For many

22. Ibid., 446.
23. Boggs, *The American Revolution*, 50.
24. Ibid., 88.

neo-Marxists, the forcibly unemployed are a crucial aspect of the capitalist system, and their separation from the workers is often regarded as a surmountable problem of organization rather than a substantial difference. Selma James writes:

> Some of us wore the blinkers of the white male Left, whether we knew it or not. According to them, if the struggle is not in the factory, it is not the class struggle. The real bind was that this Left assured us they spoke in the name of Marxism. They threatened that if we broke from them, organizationally or politically, we were breaking with Marx and scientific socialism….We found that redefining class went hand-in-hand with rediscovering a Marx the Left would never understand.[25]

Here, James is trying to move Marxist political thought beyond the confines of the factory—of the workplace—and into other kinds of emancipatory struggles, such as those of race. In fact, by regarding the forcibly unemployed and the proletariat as separate entities, and by keeping the revolutionary struggle for emancipation squarely *inside* the capitalist workplace and its hierarchical class structure, we fail to appreciate the emancipatory struggles led by historically marginalized people taking place *outside*. At a certain level, the struggles are chiefly the same, both involving one group of people being exploited by another group in a more powerful position than them. Accordingly, it seems that the difference among struggles lie merely at the organizational level: while traditional workers want economic respect, the outsiders, marginalized from traditional capitalist work, want the social, political, *and* economic respect that results from entering the workplace. Since both the proletariat and the forcibly unemployed identify an identical oppressive force from which they must be emancipated (the capitalist), their revolutionary motives are necessarily connected, making them both part of the same revolutionary agent if only through cooperation and a shared enemy.

This solution is supported by Laura Jaffee in her work *Marxism and Disability Studies*. Jaffee describes how the severely intellectually disabled "will always be beyond the purview of recognized labor power within the capitalist mode of production." The continued association of one's ability or usefulness to society with their capacity for production of labour is one way in which the capitalist system works to oppress the forcibly unemployed outside of the factory. While many social movements have worked to debunk the racist or sexist assumptions with regards to people's capacity to produce labour, disabled people have largely been left out of this structural reframing.[26]

---

25. James, "Sex, Race, and Class," 94.
26. Jaffee, "Marxism and Disability Studies," 4-5.

In terms of applying this solution practically, it is quite compatible with action being taken by neo-Marxists and other radical activists. It does not seem to stray too far from Marx's conception of labour and value, making it quite compelling to those who subscribe to his original ideas. Indeed, if the issue is only that workers and the forcibly unemployed require better organization for their dual emancipation to be realized, this is, as James writes, a surmountable problem. Workers and the forcibly unemployed could organize together to remove the socially constructed barriers that cause manufactured conflict, both improving the conditions of their labour and preventing their exploitation. The creation of emancipatory organizations beyond workers' unions could work to repair a relationship that has historically been fractured by bourgeois interests. For example, the creation of accommodations in hiring practices for those with disabilities has been one of the few times in which workers and the forcibly unemployed have come together to advocate for the improvement of working conditions. Though this sort of joint organization clearly has yet to reach complete economic emancipation, its successes have been quite powerful in demonstrating the ability of workers and the forcibly unemployed to cooperate with one another.

However, one must be careful when considering whether Marx himself would agree with this solution. While Marx's failure to consider the forcibly unemployed in his writings might be seen as an opportunity to coherently expand his original theory, it is equally possible that Marx would not agree with this reading. But even if the latter is the case, we should nevertheless appreciate the changing circumstances of our time and adopt this solution anyway.

IV. Evaluation of the Solutions

It is difficult to say for certain whether Marx would have been content with any of these solutions. However, all but the third solution exhibit major issues with prescription, as well as being overly generous readings of Marx's economic theory. Marx's inability to consider the forcibly unemployed as anything other than criminal is a profound issue and a clear instance of the problematic views of certain marginalized peoples that he held. The second solution only emphasizes that Marx's failure to consider the forcibly unemployed has left them and workers with little ability to meaningfully organize. However, this solution too easily attaches the role of the unemployed to that of the forcibly unemployed, while failing to consider the substantial difference between the two groups' productive capacity.

While none of these solutions are textually more 'right' than the other, the third reading seems to have the most relevance to, and utility for, the modern day. The forcibly unemployed do

suffer under the capitalist system, even outside the factory. As Jaffee mentions in her work, the very conception of ability versus disability is rooted in the capitalist assumption that one's utility to society is founded in their ability to act as a labourer.[27] Continually classifying their labour as unproductive is just one way in which the capitalist system directly oppresses the forcibly unemployed. As such, there is a strong motivation to say that the proletariat can, and indeed must, comprise all those that capitalism—as an economic system—oppresses, including the forcibly unemployed.

27. Ibid., 4.

Bibliography

Boggs, James. *The American Revolution: Pages from a Negro Worker's Notebook*. Monthly Review Press: New York, 2009. https://libcom.org/library/american-revolution-pages-negro-workers-notebook.

Conroe, Sean Thomas. "The Disabled Proletariat and Ableist Bourgeoisie: Tracking the Dialectic through Political Action." Senior thesis, Carroll College, 2021. https://scholars.carroll.edu/bitstream/handle/20.500.12647/10460/ConroeS_Final_2021.pdf?sequence=1&isAllowed=y.

Heilbroner, Robert L. *The Worldly Philosophers: The Lives, Times, and Ideas of the Great Economic Thinkers*. New York: Simon and Schuster, 1995.

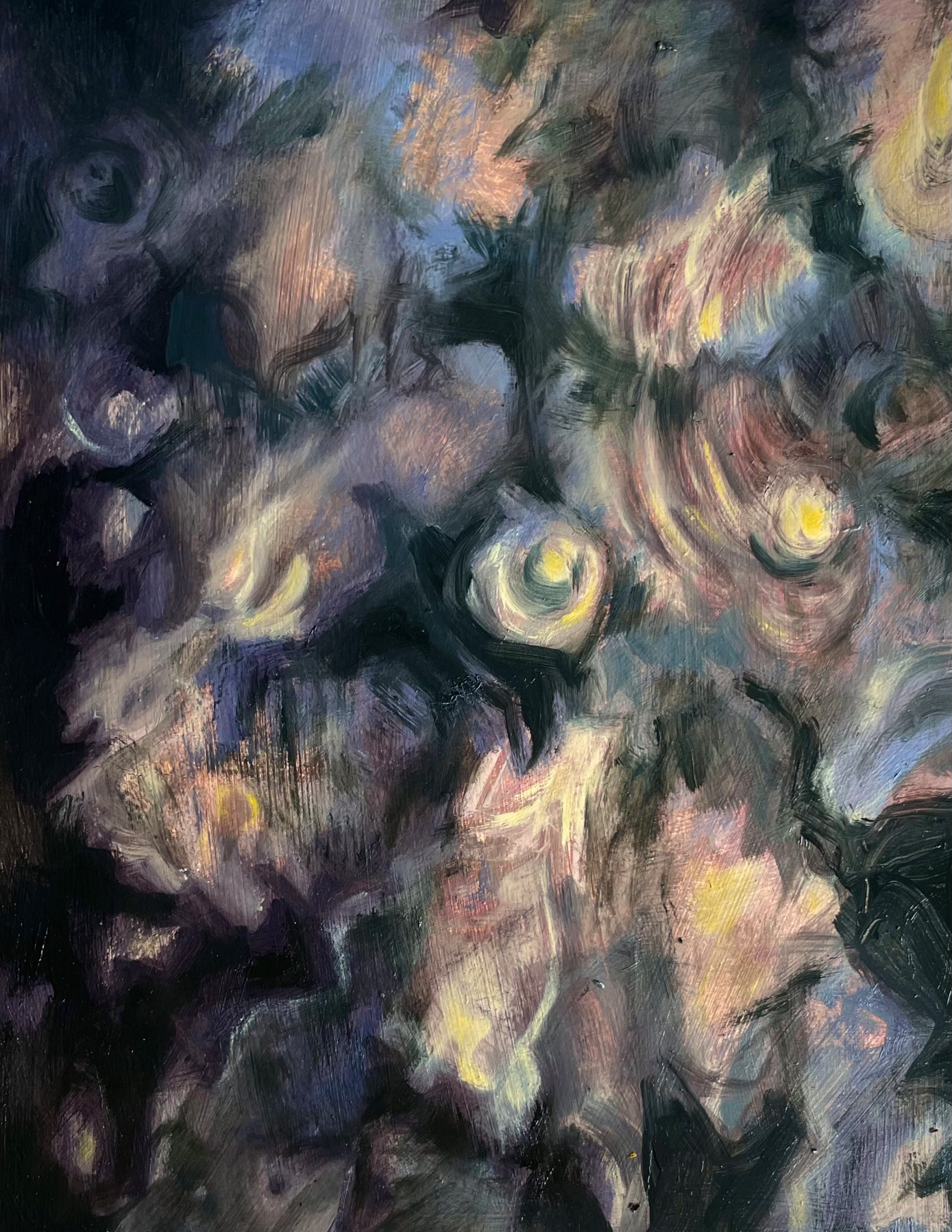
Jaffee, Laura. "Marxism and Disability Studies." In *Encyclopedia of Educational Philosophy and Theory*, edited by Michael A. Peters, 1-6. 2016. https://doi.org/10.1007/978-981-287-532-7_279-1.

Kim, Sarah. "The Truth of Disability Employment That No One Talks About." *Forbes*, 24 October 2019. https://www.forbes.com/sites/sarahkim/2019/10/24/sub-minimum-wages-disability/?sh=29db4703c22b.

James, Selma. "Sex, Race, and Class." In *Sex, Race, and Class—The Perspective of Winning: A Selection of Writings, 1952-2011*, 92-101. New York: PM Press, 2012.

Marx, Karl and Alexander Chepurenko, ed. *Capital: A Critique of Political Economy, Volume I*. Vol. 35 of Marx/Engels Collected Works. Progress: Moscow, 1996.

Schwarz, Jon. "The Business Class Has Been Fearmongering About Worker Shortages for Centuries." *The Intercept: Voices*, 7 May 2021. https://theintercept.com/2021/05/07/worker-shortage-slavery-capitalism/.



Memory Loss Doesn't Have to Entail the Loss of Personal Identity

Devin Galway

Beginning from the question 'what happens to the self when someone loses their memory?' this paper argues that memory loss does not destroy personal identity. Hoping to show that memory is an important part of who we are, I argue that personal identity should be understood as a narrative—a blend of history and fiction in the form of a chronological sequence of experiences, characters, themes, and feelings. First, I review canonical beliefs about personal identity that would deny the selfhood of people suffering from memory loss. Second, I explain how narrative depends on memory by introducing Paul Ricœur's notion of narrative identity, which makes space for the possibility of self-same subjectivity despite constant change. Third, I argue that Ricœur's account of personal identity should be accepted because it meets the challenge that memory loss poses to personal identity, and I discuss the important role of others in the creation and preservation of the self. Fourth, I reflect on some risks entailed by the narrative model, asking whether people suffering from memory loss might be vulnerable to others intentionally manipulating their life story.

Keywords: memory, Ricoeur, personal identity

I. Introduction

The phenomenon of forgetting reveals a profound lack of clarity in the notion of personal identity. For many memory theorists, the loss of recollections entails a loss of self. I believe this account is deeply unsatisfactory because it pathologizes forgetfulness and suggests that people suffering from memory loss no longer have a stable or complete self. I argue that personal identity should be understood as a *narrative* because the narrative form can merge two seemingly opposed claims: memory is an important part of the self, yet memory loss does not destroy the self.

I introduce the problem of forgetting by examining early modern accounts of personal identity. In response, I place Paul Ricœur's notion of narrative identity in conversation with the problem by explaining how narrative depends on memory. On my account, narrativizing identity can preserve the self in cases of temporary and permanent forgetting. From this claim, I explain how other people help restore forgotten elements of the self. Finally, I address the risk of manipulated identity to reflect on the limits of narrative.

II. Locke on Memory as a Prerequisite for Personal Identity

Towards the end of the *Ethics*, Benedict de Spinoza recounts an anecdote about a Spanish poet. After travelling around the world, he suddenly lost most of his memory. When others told him about his past exploits, he refused to believe he had performed them. From this, Spinoza asserts that "sometimes a man undergoes such changes that I should hardly have said he was the same man."[1] In other words, in a marginal note of the *Ethics*, we find the significant claim that memory loss changes or even destroys personal identity.

Today, efforts to ground identity in memory are typically associated with John Locke, who made similar claims seventeen years after Spinoza in the second edition of *An Essay Concerning Human Understanding*. His account will be the starting point for our analysis of forgetting and personal identity. Locke thinks that the self is nothing but constant consciousness. While he never fully defines consciousness, one of its main attributes is that it "always accompanies thinking" and is therefore what allows a person to call herself a self. Bringing in memory, Locke claims that "as far as [one's] consciousness can be extended backwards to any past action or thought, so far reaches the identity of that person."[2] It thus seems like consciousness and memory

1. Benedict de Spinoza, *Ethics*, ed. & trans. Edwin Curley (London: Penguin, 1996), IVp39(s).
2. John Locke, *An Essay Concerning Human Understanding* (Batoche Books, 2000), 2.27.9.

are the same.[3]

At the very least, Locke suggests that memory is a precondition for personal identity. This means that the loss of recollections destroys parts of the self: if someone loses many or all of her memories, she might no longer be the same self.

In a manner of speaking, consciousness is the 'glue' that makes self-constancy possible. Wondering what happens when a person is asleep, Locke claims that "if the same Socrates waking and sleeping do not partake of the same consciousness, Socrates waking and sleeping is not the same person."[4] On this account, then, the physical body is not enough to ground stable personal identity. This passage has important ethical implications: as two distinct subjects, Socrates awake and Socrates asleep are different moral agents, meaning they should be judged and treated separately.

Now, consider people with Alzheimer's disease. When they lose certain memories, they no longer partake of the same consciousness. For Locke this means that as they forget, they lose part of their identity. If we accept this view, then people with Alzheimer's disease are no longer the same moral subject. For this reason, I think that Locke's conception of personal identity is excessively fragile. For the remainder of this work, I will attempt to rethink the connection between memory and personal identity to assert that people with Alzheimer's hold onto their identity and should therefore be seen as stable selves. This is not a reversal of Locke's position. My claim is not that personal identity never changes, but rather that it exists somewhere between stability and change. To elucidate this point, I refer to Ricœur's notion of narrative identity.

III. Ricœur on Narrative Identity and Memory

(A) *Narrative Identity*

I propose that personal identity must be rethought if we do not want to say that it is permanently lost when memory fades. Nevertheless, it seems incorrect to suppose that memory plays no role in constituting the self. Ricœur's work on *narrative* identity can bridge this gap. For Ricœur, "narrative" is a middle path between history and fiction. It is a "fictionalization of history and a historization of fiction" that can be written or spoken. In the domain of personal identity, the interweaving of history and fiction produces "narrative identity," which is a constructed account of the self. The self, for Ricœur, is a narrative of a life—populated with experi-

3. Paul Ricœur, *Memory, History, Forgetting*, trans. Kathleen Blamey and David Pellauer (Chicago: University of Chicago Press, 2004), 105.
4. Locke, *Essay Concerning Human Understanding,* 2.27.19.

ences, characters, themes, feelings, and other elements belonging to the realm of storytelling.[5]

But how does this answer the problem of forgetting? Historically, philosophy has conceived of the self in two opposing ways: (a) an unchanging kernel (e.g., Plato), and (b) a constantly changing illusion (e.g., Hume). Narrative identity is neither (a) nor (b). Let me review Ricœur's reasoning. First, he acknowledges that there is indeed a static self as in (a), which he terms "idem identity." For example, the name 'Martin Heidegger' always picks out the same philosopher. Second, not wanting to dissolve the dynamic self as in (b), Ricœur introduces a second type of identity. He presupposes a shifting self that nonetheless remains "self-same" [*soi-même*] and calls it "ipse identity."[6] For example, Martin Heidegger would have always perceived himself as the same self even if his beliefs changed drastically with time. Unlike its idem counterpart, ipse identity comes purely from within. Third, Ricœur claims that narrative identity fuses idem and ipse identity in a coherent story of the self. In other words, personal identity is narrative identity. To *narrativize* is to blur the line between unchanging and changing identity.

(B) *Relation of Narrative Identity and Memory*

Despite writing about forgetting in other works, Ricœur does not explicitly connect narrative identity with his account of memory. My goal in this section will be to explain how and why narrativization requires memory. I will argue that memory is essential to narrative in a discussion of three acts involved in the creation of the self: construction, preservation, and repetition.

First, narrative identity is *constructed* out of recollections. For Ricœur, the elements of literary and personal narratives are identical. I propose that characters, themes, and other elements of the narrative form come from memory or are memories themselves. In a word, my argument is that memory is the content of narrative identity, and thus cannot be separated from narrative construction.[7]

Contrary to my argument, one could object that some elements of narrative identity are not recollections. My mother is an important character in the narrative that ties me together, yet it seems reductive to say that her affective role in my identity is a mere series of recollec-

5. Ricœur, *Time and Narrative*, trans. Kathleen McLaughlin and David Pellauer, Vol. III (Chicago: University of Chicago Press, 1988), 246.
6. Ibid., 246.
7. My argument could fall apart if we deny that narrative identity even exists. However, I am not arguing that narrative identity is somehow true whereas Locke's account is false. I am merely claiming that narrative is a promising theory of personal identity because of its implications for our relationships with others, especially those affected by forgetfulness.

tions. However, the objection fails because the elements of narrative must come from memory, since they are (at least roughly) temporally situated in the structure of an ordered whole. I am not attempting to turn my mother into memory. Ordering a narrative—an act of construction called "narrative configuration"—looks to the past for recollections. Within my identity, the idea of my mother comes from the past, that is, the idea of her is a memory. This does not mean that my mother is merely a collection of memories, because there is a significant difference between characters in the narrative self and the real people to whom they correspond. Memory is simply the medium of the parts that make up narrative identity. Narrativization synthesizes numerous literary elements into a "discordant concordance" embedded in time. In creating a narrative, I take a series of heterogeneous memories and "integrate [them] with permanence in time."[8] Recollections are diverse and discontinuous, yet they paradoxically form a single linear structure of selfhood: they are me. Memory is essential to narrative identity, then, because narrative is constructed out of recollections.

Second, narrative identity is *preserved* in memory. Given that elements of personal identity are temporally situated and made up of memories, a narrative must persist for us to recall it. We can only recall what is in our memory, so if it is possible to recall parts of one's narrative, then those parts must be contained within the faculty of memory.

It is worth clarifying the boundaries of memory. What if narrative identity can be preserved in external physical receptacles, like an autobiography? This kind of written record traces notable events, stories, and relationships, thus presenting a cohesive portrait of a person. Saint Augustine's *Confessions* preserves the theologian's identity in written form. As he tracks his relationship with God, he exclaims, "In my memory too I meet myself—I recall myself, what I have done, when and where and in what state of mind I was when I did it."[9] By reading the *Confessions*, we can meet Augustine without ever accessing *his* memory. Maybe, then, personal identity is stored in written records, rather than in memory. This scenario casts confusion on the boundaries of memory.

In response, I propose that anything capable of storing a narrative *is* memory, though this is of course not an exhaustive definition. Consequently, external physical receptacles are also memory. The memory of others, as well as "collective memory," which is a shared narrative of

8. Paul Ricœur, *Oneself as Another,* trans. Kathleen Blamey (Chicago: University of Chicago Press, 1992), 140-1.
9. Augustine of Hippo, *Confessions*, ed. Michael P. Foley and trans. F. J. Sheed (Indianapolis: Hackett, 2006), 196.

sorts, would count as well.[10] This assertion may seem strange—equivalent to claiming that Augustine's personal identity can be preserved in a memory that does not belong to him. However, there is a distinction between preservation and actualization. The *Confessions* preserves Augustine's narrative identity, but Augustine himself no longer exists to actualize, or enact, that identity. The point here is that narrative identity is preserved in memory, both individual and shared.

Third, narrative identity is *repeated* in memory—it is not a one-off story. Ricœur uses a hermeneutical-phenomenological method in his work; of course, (re)interpreting experience requires repetition. After a narrative has been configured, its possibilities are repeated and acted upon in a process called "narrative refiguration." In my view, this stage is only possible if the recollections comprising someone's identity can be repeatedly retrieved from memory, since they must be accessed and acted upon. Configuration makes possible a certain narrative understanding of oneself that is static. Refiguration, by contrast, creates a form of self-understanding that is more fluid because it occurs when someone seizes possibilities from their constructed and preserved narrative and translates them into action. This stage of narrative identity "marks the intersection of the world of the text and the world of the hearer or reader; the intersection, therefore, of the world configured by the poem and the world wherein real action occurs."[11]

To seize possibilities is to hermeneutically travel back and forth from "text" to action. "Text," as Ricœur means it, does not have to be linguistic, but means rather any assembly of signs that convey some meaning. Here, the world of the text is the world of our personal narrative. The self is dynamic and open to change because identity is repeated in several different ways, thanks to the hermeneutic of self-construction.[12] Action requires that elements of narrative be repeated, and they can only be repeated in memory. In response, one could argue that refiguration betrays a conceptual contradiction at the heart of narrative identity. How can we say that an event or a character repeats itself in memory if it never reappears in the same way? It seems we can only act on the possibilities of a configured narrative when those possibilities are static. Brought to its limit, this argument could even imply that narrative identity makes it impossible for personal identity to change.

This is an important objection, but ultimately, it is undermined by the distinction between

10. Maurice Halbwachs, *On Collective Memory*, ed. and trans. Lewis Coser (Chicago: University of Chicago Press, 1992), 45.
11. Paul Ricœur, *Time and Narrative*, trans. Kathleen McLaughlin and David Pellauer, Vol. I (Chicago: University of Chicago Press, 1984), 71.
12. Interestingly, this might imply the same event can be recollected in several, or even infinitely different ways. Ricœur does not make this move, but it is a reasonable assertion given what I have argued.

idem and ipse identity. In accepting this distinction, we eliminate any fundamental incompatibility between claiming that on the one hand, I repeat the same memory to myself, and on the other hand, the repeated memory shows itself differently in the repetition. The key is to extend the idem/ipse distinction to recollections. This would mean that recollections are simultaneously static (idem) and changing (ipse). Consequently, when we replay our life story and act upon its possibilities, our memories could show themselves differently while still denoting the same experiences, characters, themes, and other narrative elements. More generally, if we accept that a recollection is the same despite change, then parts of our narrative identity could unfold in new ways while still contributing to the same story. Clearly, memory is integral to narrative identity because we repeat recollections as we think about who we are and what we will do next.

IV. Memory and Forgetting

(A) *The Metaphysics of Memory Traces: Two Forms of Forgetting*

If memory is as essential to narrative identity as I am claiming, then the problem of forgetting becomes pressing. *How can narrative identity help us say that a person who forgets remains the same person?* Before interpreting this question, it will be helpful to clarify what I mean by "forgetting." At the heart of the phenomenon is the notion of "trace," a kind of artifact left behind by memory. There are three kinds, according to Ricœur. First is the written trace, quite literally a memory stored in the form of a textual archive, like Augustine's *Confessions*. Second is the psychical trace, which is a conscious or unconscious impression on the psyche made by a striking experience. While Ricœur cannot prove that it exists, he asserts that the brain alone cannot explain the experience of memory.[13] Third is the cortical trace, which is the physical imprint that is left on the brain when an experience is turned into a memory. This is an object of the neurosciences.

Psychical traces are particularly affective—there is a feeling to remembering and forgetting about which written and cortical traces cannot tell us anything—which is why they are the kind of trace we are mainly interested in. Putting written traces aside, Ricœur asserts that psychical traces can never be erased, whereas cortical traces can be.[14] This allows him to distinguish

13. Ricœur, *Memory, History, Forgetting*, 421. Ricœur's argument might resemble "anomalous monism," a position in philosophy of mind according to which each mental event is identifiable with a physical event but not subject to the strict kinds of natural laws governing physical events (Davidson 118). Consequently, mental events are based in the physical world, but cannot be exhaustively explained in purely physical terms. In our present discussion, this would mean that memories are identifiable with cortical traces, but not reducible to them.
14. Ibid., 417.

between two kinds of forgetting. "Reversible forgetting" is when psychical traces are hidden but not lost. It implies that at least some lost memories can be recovered. "Definitive forgetting," by contrast, is when cortical traces are completely erased. It implies that lost memories are irretrievable. Neither conception rules the other out, since we access psychical traces from the point of view of our own subjective experience, whereas cortical traces are understood objectively.[15]

The distinction between reversible and definitive forgetting helps us anticipate a possible objection to Ricœur's account. Simply put, one could claim that reversible forgetting doesn't exist and conclude that memory loss is permanent. If a person cannot draw on memories because they have been erased, then it is unclear how she could ever narrativize, that is, construct and preserve a stable sense of self. However, I will reply that narrative identity resolves the problem of personal identity in both cases of forgetting.

(B) *Forgetting and Narrative Identity*

Narrativization provides an excellent response to forgetting. In cases of *reversible* forgetting, a trace always remains. Consequently, thinkers like Spinoza and Locke cannot say that memory loss changes or destroys personal identity, since no trace is ever lost. Of course, this does not mean we cannot forget. But when it comes to reversible forgetting, it seems that we can recover memories when we interpret the traces at our disposal.

I propose that the project of interpreting memory is only possible through narrative identity. Narratives are like a stop-motion animation: they are composed of many discrete moments which, taken together, create a stable but dynamic portrait of change over time. Traces can be thought of as details that belong to each frame in the animation. One detail in isolation says nothing of its corresponding frame. However, seeing some or all of the animation will help us interpret which frame the detail belongs to. Narrative identity thus provides context for the act of interpreting traces.

As for *definitive* forgetting, it is less clear how narrative identity can respond to the Lockean challenge. It might even be argued that if reversible forgetting exists, then definitive forgetting is a logical impossibility, though Ricœur would disagree. Unfortunately, Ricœur has little to say about this kind of forgetting. He describes the permanent erasure of traces as a "misfortune which beckons us more to poetry and to wisdom than to science."[16] While this may be true, I think that narrative identity can provide a better response. With definitive forgetting, the trace is permanently erased; a particular recollection is forever lost. Imagine that some of the de-

15. Ibid., 428.
16. Ibid., 427-8.

tails of the stop-motion animation are destroyed. Without those details, the picture will still play. Analogously, I maintain that even if memory traces are permanently lost, it does not follow that one's narrative will be destroyed, since it is composed of several recollections. Note that while I originally formulated the problem of personal identity as one of losing recollections, it is just as much about losing memory traces insofar as traces point to recollections.

An important objection might come to mind: if someone were to permanently lose every single trace or recollection, there would be nothing left to narrativize, and they could no longer be called the same person. My account reaches a limit at this point. If someone were to permanently lose their entire memory, their narrative identity would indeed be irretrievable.

However, having a stable identity is not a necessary condition for being treated as a moral subject. Someone suffering from complete amnesia should retain dignity despite lacking *metaphysical* selfhood. Individual entities ought to be regarded and treated as if they possess inherent moral worth. We could argue that as social beings, we are entitled to cultivate relationships with individuals and communities as *social equals*. Or maybe we could reason that people deserve dignity because we are alive and have an inner mental life. Here I do not intend to put forth an exhaustive defense of dignitarianism, but rather to appeal to dignity to ensure the wellbeing of people whom we cannot understand as the same self even with a narrative conception of identity. The limit case of total memory loss is a useful point to transition towards the role of others in narrative identity.

(C) *Forgetting and Other People*

I have obliquely touched on the importance of others in the process of self-construction, but now it comes to the fore. Identity is constructed, preserved, and repeated by oneself, but also by others. Consequently, 'personal' identity is somewhat of a misnomer. Following P.F. Strawson, Ricœur suggests that if personal mental predicates are to be self-ascribable, then they must also be other-ascribable. The reason for this rule is that we view other people as the kind of entities to whom personal mental predicates can be assigned. For example, if I wish to express that I am affected by the predicate 'depression,' I must also think that others can be affected by the same predicate. For Strawson and Ricœur, a predicate must mean the same thing when applied to oneself or another to satisfy the rule. Crucially, this equivalence of meaning "preserves the asymmetry between self-ascribable and other-ascribable."[17] Recollecting and forgetting count as personal mental predicates that must be ascribable to others. And yet an asymmetry endures:

17. Ibid., 124-5.

> The problem of two memories is not abolished. It is framed. What distinguishes self-ascription is appropriation under the sign of mineness, of what is my own….And it is this capacity to designate oneself as the possessor of one's own memories that leads to attributing to others the same mnemonic phenomena as to oneself.[18]

Ricœur is saying that my memory is certainly and distinctly *mine*, but that I must also recognize that other people possess the same faculty in order for me to understand the *mineness* of my memory. Baked into the very idea of personal memory, then, is an imperative to see others as possessing the faculties that make it possible to narrativize. Even after losing their memories, people will make new ones.

In addition to helping us understand what memory means, I propose that other people also play an active role in the construction and preservation of narrative identity. The reason for this is quite simple: everyone is always already embedded in relations with others. In keeping with the theme of narrative, these relations are historical; they are elements of our own history. Ricœur writes that "the life history of each of us is caught up in the histories of others."[19] By their entanglement, narratives influence each other.

Two points follow. First, narrativization does not occur outside of a context of relations. We borrow characters and share experiences with others. The very ideas that become part of us are given by the group. As Maurice Halbwachs puts it, "when people think they are alone, face to face with themselves, other people appear and with them the groups of which they are members."[20] Hence, narratives are always constructed with the help of others.

A person with Alzheimer's disease may forget a character or an idea that belongs to her narrative. However, the forgotten element of her identity might be retrievable through the help of others. Some people with Alzheimer's disease endure the distressing process of forgetting their son or daughter. But since their conception of their child is partly co-constructed with others, it seems possible that they could be re-introduced to that conception by someone who was involved in their child's life, or even by a written trace.

Second, narrative identity is not preserved by one person alone. If we adopt the narrative model, then part of our narrative identity is preserved in memory. As I have suggested, memory can be shared. Given these propositions, I claim that narrative identity is preserved at least in part by others. This of course connects to the question of forgetting. A person with Alzheimer's

18. Ibid., 128.
19. Ricœur, *Oneself as Another,* 161.
20. Halbwachs, *On Collective Memory*, 49.

disease may have forgotten a lot, but parts of their narrative might have either been preserved in someone else's memory, or in the form of a written trace. As such, lost elements of narrative might be recoverable. In sum, then, other people are essential to narrative identity. We cannot understand what memory is without them, and they seem to play an active role in the construction and preservation of narrative.

Before considering the limitations of Ricœur's account, I will address two potential concerns about my position on the co-preservation of narrative. It seems that narrative identity is only partially preserved by others. Due to the incompleteness of memory, it might only be possible to restore a fraction of personal identity to someone who has forgotten. However, my intention in this paper is to provide a framework that can account for stable identity despite memory loss. I agree with the concern: it seems unreasonable to maintain that personal identity can be comprehensively reproduced and returned.

Also, if nobody remembers an agent, it is hard to see how their personal identity could be retrieved. Again, I have no trouble accepting this possibility. The following conditions would have to be met for someone's identity to be permanently destroyed: (a) someone forgets most of their narrative; (b) nobody remembers them; and (c) there are no written records of them. Even in this case, at the very least, we still ought to treat them with dignity despite the metaphysical state of their selfhood.

V. Overdependence on Others and the Risk of Manipulated Identity

If we accept that identity is unified through narrative, then our sense of self might be at risk if our narrative is intentionally distorted by another. The story of a life is unified but also "an unstable mixture of fabulation and actual experience."[21] It will not always accurately reflect the past. It might contain discontinuous or even opposed elements. There is nothing inherently wrong with such tension. But for some, such as those with Alzheimer's disease, memories can be fragile. Should we be worried that someone could conceivably create false memories of another and destroy true ones? Such manipulation would pervert our life story and hence our identity.

In a 1995 study, a group of researchers successfully implanted false memories in five participants out of a group of twenty-four. People recalled being lost in a mall when they were young, but the story was entirely made up.[22] If false memories can be implanted, then narrative identity can be manipulated. This is a real problem for my account of narrative identity. Ricœur

21. Ricœur, *Oneself as Another*, 141, 162.
22. Elizabeth Loftus and Jacqueline Pickrell, "The Formation of False Memories," *Psychiatric Annals* 25, no. 12 (1995): 721–722. doi:10.3928/0048-5713-19951201-07.

himself is aware of this issue when he exclaims, "It is always possible to weave different, even opposed, plots about our lives….In this sense, narrative identity continues to make and unmake itself….Narrative identity thus becomes the name of a problem at least as much as it is that of a solution."[23] The more someone forgets, the more they depend on others to help them recall and reconstruct their identity. As such, the forgetful are vulnerable.

Perhaps this would not be a problem if ipse identity were to precede idem identity. If this were the case, my personal identity would be grounded primarily in the fact that I perceive myself as the same despite constantly changing. In this state, I would attribute less importance to accounts of myself that are given by others and be more capable of resisting narrative inaccuracies and tensions. However, this argument undervalues the role of others in constructing, preserving, and repeating our narrative. The problem of forgetting shows that some selves depend heavily on others to understand who they are.

Concerns about manipulated personal identity within the frame of narrative are valid. Though the narrative model is more nuanced than Locke's, it still uses memory to construct the self. If false memories can be implanted, our sense of self can be twisted. This aporia cannot be overcome. But we can ignore the metaphysical tension and treat those suffering from memory loss with dignity despite a lack of narrative 'wholeness.' Perhaps their increased vulnerability obligates us to help them maintain their narrative structure and resist manipulation.

VI. Conclusion

I have argued that the narrative model allows us to connect memory and personal identity while resisting the implication that forgetfulness entails a loss of self. This improves upon traditional conceptions of identity by allowing the self to persist through change. At the same time, the narrative self is also more dynamic than Locke implied. It is a mosaic of memories which are constructed, preserved, and repeated. An important takeaway concerns dependency: constructing narrative identity and storing relevant memories is easier and richer with the help of others. I have not considered the possibility that some parts of the self might not conform to the rules of narrative. By reckoning with this challenge, I believe that an even better account of narrative identity can emerge.

23. Ricœur, *Time and Narrative,* Vol. III, 248–9.

Bibliography

Atherton, Margaret. "Locke's Theory of Personal Identity." *Midwest Studies in Philosophy* 8, no. 1 (1983): 273-293, doi:10.1111/j.1475-4975.1983.tb00470.x.

Augustine of Hippo. *Confessions*. Edited by Michael P. Foley and translated by F. J. Sheed. Indianapolis: Hackett, 2006). www.ebookcentral.proquest.com/lib/mcgill/reader.action?docID=711602.

Davidson, Donald. "Mental Events." In *The Essential Davidson*, 105-118. Oxford: Oxford University Press, 2006. www.ebookcentral.proquest.com/lib/mcgill/detail.action?docID=422871.

Halbwachs, Maurice. *On Collective Memory*. Edited and translated by Lewis Coser. Chicago: University of Chicago Press, 1992.

Locke, John. *An Essay Concerning Human Understanding*. Batoche Books, 2000. www.ebookcentral.proquest.com/lib/mcgill/detail.action?docID=3117747.

Loftus, Elizabeth, and Jacqueline Pickrell. "The Formation of False Memories." *Psychiatric Annals* 25, no. 12 (1995): 720-725. doi:10.3928/0048-5713-19951201-07.

Ricœur, Paul. *Memory, History, Forgetting*. Translated by Kathleen Blamey and David Pellauer. Chicago: University of Chicago Press, 2004.

——. *Oneself as Another*. Translated by Kathleen Blamey. Chicago: University of Chicago Press, 1992.

——. *Time and Narrative*. Volume I. Translated by Kathleen McLaughlin and David Pellauer. Chicago: University of Chicago Press, 1984.

——. *Time and Narrative*. Volume III. Translated by Kathleen McLaughlin and David Pellauer. Chicago: University of Chicago Press, 1988.

Spinoza, Benedict de. *Ethics*. Edited and translated by Edwin Curley. London: Penguin, 1996.

The Gift of Forgiveness

Milan Thiessen

This paper addresses the topic of forgiveness by exploring the philosophical contributions made to the issue by Paul Ricœur, Jacques Derrida, and Vladimir Jankélévitch. I mainly deal with the different ways in which forgiveness can be understood as a gift, using this metaphor of a gift to frame a discussion of the concept it represents. The question of whether repentance on the part of the wrongdoer is a necessary precondition for forgiveness is pertinent to any conversation on this subject, and is consequently one that all three theorists address. I defend the view that in practice remorse and repentance by the wrongdoer make what we typically consider forgiveness possible, but that in its ideal and true form, forgiveness does not require that it be requested by the one who is forgiven.

Keywords: forgiveness, grace, disparity, asymmetry, Ricœur, Derrida, Jankélévitch, evil, love, gift, unconditional, conditional, request, remorse, repentance, Christianity

I. Introduction

Forgiveness represents a unique challenge in the realm of moral philosophy, representing something that is both necessary for life to go on, without which each of us would be left eternally suffering the consequences of a single wrongdoing, and an impossible task that asks us to forgive what is unforgivable. Jankélévitch, Derrida, and Ricœur offer insight into the complex issues surrounding forgiveness, each of them building on the others' ideas to develop some philosophical understanding of the forces at play.

This paper will discuss these three thinkers' ideas and bring them into conversation with one another through the metaphor of forgiveness as a gift. This metaphor is only used and explored thematically by Ricœur, but I argue that it can be applied to Derrida and Jankélévitch as well. I will first explain Jankélévitch's concepts of the dually all-powerful forces of love and evil, as well as his separate discussion of truly unforgivable offences in the context of Nazi Germany's crimes against its Jewish population during the Second World War. I will then discuss Derrida's argument that forgiveness begins its journey with the committing of an unforgivable act, and that this unforgivable quality is what sets up the possibility for forgiveness. I will then elucidate Ricœur's thoughts on this issue, put forward in the "Epilogue" to *Memory, History, Forgetting*, wherein he describes the asymmetry between the one who forgives and the one who requests forgiveness, as well as the debated necessity of the request for forgiveness as a condition to what ought to be an unconditional gift. Finally, I will discuss my own stance on forgiveness and offer some thoughts on the matters discussed by Jankélévitch, Derrida, and Ricœur, culminating in a defense of a particular sense of forgiveness as a gift. Ultimately, I argue, forgiveness is most intelligible on the Christian interpretation, which necessarily implicates a Holy Spirit. Other conceptualizations of forgiveness as a gift, discussed explicitly only by Ricœur, are less philosophically valuable because they overlook something central to the phenomenon in question.

II. Jankélévitch's Dualist View of Forgiveness

Vladimir Jankélévitch contends in the conclusion of *On Forgiveness* that the only thing that stands in need of forgiveness is precisely those acts which are unforgivable. 'Venial' misdeeds, such as spilling hot coffee on someone or arriving late to a meeting, do not require forgiveness as they are not serious enough to warrant it; rather, the job of forgiveness is to combat truly evil deeds, like an assault or betrayal of some kind, with a 'mad' response of love.[1] The

1. Vladimir Jankélévitch, *Forgiveness*, trans. Andrew Kelley (Chicago: University of Chicago Press, 2005), 156-7.

inherent senselessness of forgiveness is thus a central feature for Jankélévitch. By no means an excuse for what has transpired, nor a way of denying its seriousness, forgiveness stands instead as the impossible response to truly inexcusable misdeeds.

Importantly, Jankélévitch does not take forgiveness to be unconditional. He argues that the remorse and 'insomnia' of the wrongdoer is a necessary precondition for forgiveness if we want the concept to have any coherence at all, writing, "[w]hen the guilty person is fat, well nourished, prosperous, and takes advantage of the economic miracle, then forgiveness is a sinister joke." Jankélévitch is therefore of the view that forgiveness may only be bestowed upon one who feels remorse for their actions and requests to be forgiven. Otherwise, forgiveness is 'a sinister joke' in that the wrongdoer is seemingly granted impunity for their wrongdoing.

However, Jankélévitch goes on to use language which seems to contradict this point when he states that the two parties in question, the wrongdoer and the wronged, each have their own tasks to fulfill. He writes specifically that it is not the criminal's concern whether he is forgiven, and that it is not the forgiver's concern whether the criminal repents; rather, the criminal must "redeem himself all alone," while the task of the offended party is to forgive.[2] This makes both the acts of wishing for forgiveness and granting it totally private affairs, respectively taking place within the hearts of each party. It would not seem to be an exchange from one party to the other so much as a mutual but independent process of healing.

This tension between formulating forgiveness as something given from one person to another in response to a request (which we might liken to the exchange of a gift), versus as an individualized phenomenon, is not due to any oversight by Jankélévitch. Rather, it is indicative of the complex structure of forgiveness that he characterizes in this text; namely, that love and evil are two equally all-powerful forces, forgiveness being the 'mad' response of love in the face of evil deeds. Only evil deeds stand in a position to be forgiven, yet their evilness is what makes them unforgivable. This tension plays out in each situation where forgiveness appears 'possible,' possible not because a deed is easily excused, but because the gravity of the offense calls on forgiveness as the only force capable of resolving the tension.

Jankélévitch is extreme in his insistence that forgiveness knows no limits, at least in his conclusion to *On Forgiveness*. He maintains that "[a]ll misdeeds are curable to infinity" and that "where misdeed flows, grace overflows." This is based upon the idea that forgiveness is specifically possible in response to truly evil deeds—deeds which can have no explanation or under-

2. Ibid., 157.

standing, but simply force us to recognize that wickedness exists. Since true evil is what uniquely requires forgiveness, there can be no act so awful as to pass beyond the realm of the forgivable. Jankélévitch describes this as a perpetual grappling of love with wickedness, as these two forces are equally all-powerful; something is forgiven by the infinite love that is the source of forgiveness, and then human misdeed re-establishes itself once more.[3]

The battle described here is very general, speaking of no specific evil deed that can be forgiven. This generality makes forgiveness seem like an appealing option in a situation where some wrong has been committed, painting it in a clean light as a tremendously inspiring part of the human experience. This stands in notable contrast to the more focused discussion of a particular historical event that Jankélévitch gives in his paper "Should we pardon them?" Here, he assumes a very different tone as he addresses the wretched evil of the genocide committed against the European Jewish population in the Second World War. This crime is aptly classified as a 'crime against humanity,' he says, because it was the very essence of humanity that was the object of the efforts of extermination.[4] Jankélévitch gravely writes, "the novel inventions of cruelty, the most diabolical abysses of perversity, the unimaginable refinements of hate, all of this leaves us mute and above all confounds the spirit. The bottom of this mystery of gratuitous evil has never been sounded."[5] This passage illustrates the gravity of acts of evil. Such wrongs are not recognizable as forgivable when one is confronted with them in the raw emotional moment; there is a feeling of horror that takes one's words away and leaves one with an acute awareness of the existence of evil. The normal, and perhaps only acceptable human reaction when encountering such evil is to be horrified, angry, and to crave something resembling justice.

This paper by Jankélévitch displays a feeling of horror and mourning for the deep and irreversible cruelty that has been committed, and one is notably not tempted to raise the possibility of forgiveness when reading it. Encountering evil deeds beyond comprehension makes the suggestion of forgiveness seem outrageous—even cruel. This is a quality not felt when speaking of forgiveness in general, as a concept.

Jankélévitch writes, "[t]o forget this gigantic crime against humanity would be a new crime against the human species."[6] Here, even forgetting the crime amounts to forgiving, since casting aside evil too effortlessly in the name of moving on and freeing its perpetrators from the

3. Ibid., 162-4.
4. Vladimir Jankélévitch, "Should we pardon them?" *Critical Inquiry* 22, no. 3 (Spring 1996): 555.
5. Ibid., 558.
6. Ibid., 556.

consequences of their actions means forgiving the crime too easily and too quickly, undermining the foundations of morality altogether. Forgiveness is not identical to forgetting, but the former does seem to pave the way for the latter in that an unforgiven act can hardly be forgotten, whereas forgiveness allows the parties involved to move on and 'go about their business' in a way that certainly lends itself to a kind of forgetting. What Jankélévitch seems to mean here is that forgiving the crimes of the Holocaust would enable a kind of forgetting or fading away of memory that simply cannot be risked, given the gravity of the offenses in question. Forgiving risks forgetting, and forgetting risks the recurrence of the same evil.

The seeming contradiction between Jankélévitch's two writings on forgiveness, on the one hand asserting that nothing is beyond forgiveness and on the other insisting that crimes against humanity are unforgivable, speaks to the exact equality of the forces of love and of evil described in *On Forgiveness*. "Forgiveness is strong like wickedness; but it is not stronger than it," Jankélévitch writes.[7] This battle plays out brutally within the individual when confronted with true evil while still believing that forgiveness is a profoundly important source of value in human life, and that there is an 'imperative of love,' as Jankélévitch does.[8] Nevertheless, he is consistent across the two pieces in stating that the request for forgiveness, the distress of the wrongdoer, is a necessary precondition for forgiveness to be considered.

III. Derrida's Irreducible Polarities

Jacques Derrida echoes exactly Jankélévitch's sentiment that the unforgivable is the only thing that stands in a position to be forgiven. The notion of forgivable sins presupposes that forgiveness can ever be expected or deserved, and consequently renders the very idea of forgiveness null. Derrida argues that forgiveness is not normal, writing, "it *should* remain exceptional and extraordinary, in the face of the impossible—as if it interrupted the ordinary course of historical temporality."[9] This view accords with the idea that forgiveness is in an important sense an impossible task: only the unforgivable is capable of being forgiven, yet forgiveness is seemingly ruled out by the unforgivability of the crime. So, forgiveness is always an infinitely surprising occurrence—it cannot be expected or deserved, only *given* in a way that interrupts the usual course of things. Derrida cites both Jankélévitch and Hannah Arendt to emphasize how integral this para-

7. Jankélévitch, *Forgiveness*, 165.
8. Ibid., 162.
9. Jacques Derrida, "On Forgiveness," In *On Cosmopolitanism and Forgiveness*, trans. Mark Dooley and Michael Hughes (London: Routledge, 2001), 32.

doxical quality of forgiveness is to the 'heritage' surrounding it.[10] This point of overlap between three prominent philosophers allows us to see that forgiveness is at the same time an exceptional and surprising act with every occurrence, since it runs counter to the natural human instinct to seek vengeance when one is wronged, as well as a cathartic experience for both the wrongdoer and the victim. This evokes the view of forgiveness as a gift that frees the agent from the consequences of his own actions *and* frees the one who gives it from suffering eternally as the victim of a misdeed. We will address these senses of the 'gift' below; Derrida only uses the word 'gift' when referring to the Christian sense of forgiveness as something graciously granted by the Holy Spirit.

Derrida introduces the central conflict between two pervasive formulations present in the inherited tradition of forgiveness, chiefly stemming from Christianity. On the one hand, forgiveness necessarily has an *unconditional* character, called 'aneconomic' by Derrida to emphasize the absence of mutual exchange in the granting of forgiveness. In this characterization, the guilty party is forgiven *as guilty*—they are not distanced from their actions through promises of reformed behaviour or expression of remorse, but rather are forgiven precisely as the agent of the misdeed. On the other hand, there is the *conditional* sort of forgiveness granted on the condition that the guilty party recognize their fault and request forgiveness, promising to reform their behaviour in the future. In the latter case, Derrida argues that the guilty person is not being forgiven as such; rather, a different subject from the one who committed the offence is being forgiven. This gives rise to the dilemma that forgiveness seemingly must be unconditional despite the common intuition that the wrongdoer must repent before forgiveness can be granted. Derrida disagrees with Jankélévitch that the repentance of the sinner is necessary for forgiveness to be granted, arguing instead that forgiveness is either conditional or unconditional and logically pointing out that it cannot be both while remaining one and the same concept.[11] He particularly takes issue with the traditional idea that forgiveness must have an absolute meaning, that it must *make sense*, preferring instead to recognize the inherent separation between forgiveness as a phenomenon with conditions attached and as something unconditional. Ultimately, Derrida will argue that the irreducibility of these two poles of forgiveness is precisely where the closest thing to an understanding of it can be gleaned. Since forgiveness is intrinsically paradoxical, this ten-

10. Arendt writes in *The Human Condition* that forgiving "is the only reaction that does not merely re-act but acts anew and unexpectedly, unconditioned by the act which provoked it and therefore freeing from its consequences both the one who forgives and the one who is forgiven" (Hannah Arendt, *The Human Condition*, 2nd ed. (Chicago: University of Chicago Press, 1998), 241).
11. Derrida, "On Forgiveness," 34-5.

sion between conditionality and unconditionality can never be truly resolved or understood.

Jankélévitch concludes in "Should we pardon them?" that the irreparability and inexpiability of the crimes of the Holocaust means that they are also unforgivable, but Derrida insists that this is not a logical progression—the juridical notions of reparation and justice are not analogous to forgivability. Still referencing Jankélévitch, Derrida writes, "'Forgiveness dies in the death camps,' he says. Yes. Unless it only becomes possible from the moment that it appears impossible."[12] He is reinforcing the claim, made earlier, that forgiveness is precisely a response to unforgivable acts; it is a useless and meaningless concept if applied to acts deemed 'forgivable.' This is the point made by Jankélévitch himself in *On Forgiveness* and seemingly walked back in "Should we pardon them?" It is precisely the impossibility of forgiveness which makes it a miracle, something never expected or normal. So, forgiveness is a human power that does not 'make sense' according to our normal understanding of the world in Derrida's view. It can never make sense, yet it is a perpetual possibility.

Derrida further develops the two opposing poles of unconditional and conditional forgiveness. He reiterates that sometimes forgiveness is a gracious gift with no element of exchange and no conditions external to it, while at other times it is given on the condition that the guilty party repents and demonstrates transformation. He maintains that these two poles cannot be reduced to one another; they are "absolutely heterogeneous." The first sense has a quality of 'madness' in that it forgives with no promise of changed behaviour or remorse of the sinner, seeming somewhat to make light of the evil committed by releasing the wrongdoer from responsibility for the wrong they have done. The second sense is less purely understood as forgiveness, and Derrida asserts that it is *not*, in fact, true forgiveness; yet, conditions such as repentance are practically necessary for the ritual of forgiveness to result in change. This is particularly the case for institutionally mediated performances of forgiveness, aimed at reconciliation, reparations, or some other end resembling, but not identical to, forgiveness after an historical wrong.[13] Derrida does not attempt to resolve this conflict in his section on forgiveness, arguing that it is simultaneously the case that a kind of understanding must exist between the two parties (wrongdoer and wronged) for forgiveness to be possible, and that the non-identification of the two parties remains intact.[14] These relations are all developed further by Ricœur, to whose insights on forgiveness we will now turn.

12. Ibid., 36-7.
13. Ibid., 44-5.
14. Ibid., 48-9.

IV. Ricœur's Asymmetry

In the "Epilogue" to *Memory, History, Forgetting*, Ricœur posits the idea of a vertical disparity between the 'depth of fault' and the 'height of forgiveness.' These two ends of a vertical axis have as their basis the two speech acts which come face-to-face in the moment of forgiveness. Firstly, the agent binds himself to his action and recognizes himself as accountable. Ricœur names this speech act 'the avowal.' Secondly, the proclamation of love and joy indicating the possibility of forgiveness is called 'the hymn' by Ricœur.[15] The religious connotation of the word 'hymn' is intentional, given that the language of forgiveness which dominates today has its origins in the Abrahamic religions, especially Christianity. The hymn that gives voice to the existence of forgiveness formulates it in the proclamation "there is forgiveness," where this voice "has no need to say who forgives and to whom forgiveness is directed" because it is a power that is higher than the level on which these details present themselves. Forgiveness belongs to the same family as love, Ricœur says: it is a spiritual gift granted by the Holy Spirit spoken of in the Bible. It does not belong to any given person *per se*, but 'flows through' the individual, as though from 'above,' in the situation of forgiveness. Ricœur writes, "It is from above, in the way that the admission of fault proceeds from the unfathomable depths of selfhood."[16] These two ends of a vertical relation, where the one who forgives does so from an elevated state akin to a divine capacity for love, and the one who requests forgiveness looks up from the abyss of remorse and recognized fault, are flattened when one considers the horizontal relation of an exchange wherein the gift of forgiveness is given by one and received by the other. The disproportionality between the positions of the two parties involved in forgiveness is what concerns Ricœur primarily, but he develops the ideas contained within the horizontal relation of an exchange in order to explore this disproportionality.

Ricœur formulates the equation of forgiveness as the impossibility of forgiveness replying to the unpardonable nature of moral evil.[17] He notes that the avowal of the wrongdoer bridges the gap between the agent and his action, and is thus a crucial step in forgiveness because it enables the agent to be forgiven *as the author* of the wrongful action, through his own admission.[18] Here Ricœur is taking the side of Jankélévitch in this debate against Derrida, holding that the admission of guilt on the part of a wrongdoer would seem to be essential for forgiveness. Nonethe-

15. Paul Ricœur, "Epilogue: Difficult Forgiveness," in *History, Memory, Forgetting*, trans. Kathleen Blamey and David Pellauer (Chicago: University of Chicago Press, 2004), 457-8.
16. Ibid., 467.
17. Ibid., 458.
18. Ibid., 461.

less, the assumption that a request from the wrongdoer is necessary for forgiveness to be granted is opposed to what Ricœur calls the primary characteristic of forgiveness: its unconditionality. "Yet we believe, on the level of practice," Ricœur writes, "that there does exist something like a correlation between forgiveness requested and forgiveness granted." This turns the relation into an exchange between the two parties, placing them on equal footing, in a one-on-one, intimate relation. Ricœur argues that the vertical relation between height and depth and the tension between unconditionality and conditionality observed by Derrida goes unnoticed if forgiveness is formulated as a bilateral exchange between the request for and the offer of forgiveness—namely, in the formulation of forgiveness as a gift given from one party to the other.[19]

Ricœur goes on to state that the chief issue to be resolved in conceiving of forgiveness as a gift is that of reciprocity. Some have argued that giving a gift necessarily comes with the expectation of something in return from the one who receives the gift. Citing the sociologist Mauss, Ricœur elucidates a triadic relation between the three obligations of giving, receiving, and giving back. But what could the forgiven person give back to the one who forgives them? This alignment of forgiveness with the circle of the gift in this way makes it no different in its logic from retribution, and forgiveness is clearly distinct from this, as Ricœur argues. Forgiveness is not an attempt to equalize the partners; it is better formulated as a gift given in a spirit of complete generosity, with no expectation of anything in return. This sense of forgiveness as giving comes in the commandment to love one's enemies that stands at the foundation of Christianity, and to love them as they are—as our enemies, even ones who have not asked us for forgiveness. This represents the true impossibility of forgiveness—it contradicts the equalizing force of retribution and even the reciprocity of the Golden Rule urging us to treat others as we wish to be treated. It is thus the truest sense of a 'gift': something given freely, with no expectation of return.[20] The commandment to love one's enemies is fulfilled by way of impossible forgiveness, coming from a metaphysically higher source of love and manifesting in a particular situation, as discussed previously. The instinct toward reciprocity is so deeply ingrained in us that it seems impossible to give from such an authentic state of generous love, yet the Biblical understanding of forgiveness (which all three thinkers discussed in this paper take to be primary) holds that it is possible.

Ricœur then raises an objection to his conceptualization of forgiveness as laid out thus far. He states that giving what has not been asked for disrespects the dignity of the receiver by placing the giver in a position of 'condescending superiority.' In this scenario, the recipient is

19. Ibid., 478.
20. Ibid., 480-81.

crushed under the weight of what he cannot repay and is burdened with a duty to be grateful. Ricœur responds that there is some validity to the idea that the dignity of the recipient is not respected in an 'horizontally imbalanced' exchange of a gift from one person to another. To resolve this, he revisits the idea of consideration for the other and argues that the requirement for both parties to be mutually considerate makes the request for forgiveness an important part of the equation. Out of consideration for the other's dignity, we do not give the gift of forgiveness when it has not been requested. When it has been requested, the guilty party willingly receives the gift instead of having it forced upon them. As a result, "the reciprocity of giving and receiving puts an end to the horizontal asymmetry of the gift with no expectation of return, under the aegis of the singular figure constituted by consideration."[21] Consideration is the force which allows for the horizontal asymmetry to be resolved because it represents a mutual acknowledgement of the wrong that has been committed. Further, this acknowledgement maintains the vertical disparity between the victim and the wrongdoer because it recognizes the wrongdoer and the victim as such as a precondition for the exchange of forgiveness. However, this solution offered by Ricœur does not resolve the problem of whether forgiveness can really be considered unconditional when the admission and request of the wrongdoer is deemed necessary, as in this account.

So, the gift must be a response to a request for forgiveness, according to Ricœur. On the horizontal plane, then, the two parties are in an important sense equal. Nevertheless, Ricœur holds onto the vertical dimension that belongs to the moment of forgiveness. He declares that "forgiveness spans an interval between the high and the low, between the great height of the spirit of forgiveness and the abyss of guilt. This asymmetry is constitutive of the forgiveness equation. It accompanies us like an enigma that can never be fully plumbed."[22] The asymmetry between the two parties must be maintained if we wish to understand forgiveness correctly. I will not elucidate the rest of Ricœur's arguments regarding the complexity of forgiveness here, but rather now engage in a brief discussion of my own thoughts on the matter.

V. Forgiveness Is a Paradoxical and Superordinate Gift

The ideas of Jankélévitch, Derrida, and Ricœur provide a rich background against which to navigate the topic of forgiveness. All three figures understand forgiveness in a paradoxical manner: Jankélévitch formulates it as boundless love grappling with absolute evil in an irresolvable battle; Derrida speaks of two mutually irreconcilable polarities of conditional and uncondi-

21. Ibid., 481-2.
22. Ibid., 483.

tional forgiveness; Ricœur calls forgiveness an enigma characterized essentially by an asymmetry between the victim and the perpetrator of a misdeed. Though these formulations are distinct, they each point to a crucial, definitive feature of forgiveness: its paradoxicality. Evil deeds render forgiveness an impossibility because to forgive something evil offends our deepest moral intuitions, overriding our base instinct to seek revenge and bring about some putative form of justice, and thereby seeming to undermine the very foundations of morality. Yet, as Ricœur says, 'there is forgiveness,' and as Derrida says, forgiveness is an extraordinary possibility in the face of the impossible.[23] An impossible feat is, in fact, possible. Forgiveness exists in the face of acts which deprive one of the possibility of forgiving, and because of this is considered a miracle or a gift from God in the Christian tradition. I argue that the clearly paradoxical, incomprehensible nature of forgiveness must be instrumental to the way we think about and categorize it. The dual characteristics of being utterly nonsensical in the context of human affairs, yet clearly a real phenomenon, cause forgiveness to appear a mysterious, gracious gift—a strange yet wonderful capacity that human beings possess.

I would like to explore the different senses in which we might regard forgiveness as a gift in order to clarify how the Christian formulation is the one which makes most sense. First, and perhaps most obviously, forgiveness may be thought of as a gift given from the victim to the perpetrator of a misdeed. To argue that forgiveness is a gift in this sense is certainly valid; to be forgiven for a genuine wrongdoing is a supremely humbling, beautiful, even transformative experience that may well be received in much the same way that one receives a valuable gift. A person who is forgiven after a period of profound remorse for their wrongdoing does not take this gift lightly unless they are deeply corrupted by evil, in which case their remorse cannot have been genuine. This gift received by a wrongdoer frees them from the stain of their past actions and allows them to act anew, to return to their basic disposition of goodness as Ricœur goes on to discuss in relation to Arendt in his "Epilogue." It is not an expectation on the part of the wrongdoer, but rather something graciously granted by the victim. Just as gifts are generally considered optional offerings (if one has been compelled to give a gift, it seems much less genuine), forgiveness is a choice made by a victim to allow a wrongdoer to move past their remorse and to 'act anew.'

Moreover, forgiveness is a human power aptly considered a gift for the one who grants it as much as for the one who receives it, in that it has a healing effect on its bestower. This pass-

23. Derrida, "On Forgiveness," 48.

ing-through goes along with Ricœur's characterization of forgiveness as a heightened thing coming from above the level of any given subject. The proclamation "there is forgiveness" speaks to its being a vast resource that individuals can tap into when the situation impels them to, not as something that has its origin in the individual *per se*. This sense of forgiveness, given by Ricœur, implicates the Holy Spirit of the Christian Bible as its source. Of course, this is a metaphysical and not a literal resource, and the religious terminology may not appeal to the secular reader. However, this issue of forgiveness' centrality to the Christian faith cannot be easily sidestepped in attempting to understand it.

One who forgives is freed from the consequences of the wrong committed against them (in a psychological rather than material sense), allowing them to escape their victimhood. Jankélévitch refers distinctly to the benefit of forgiving for the forgiver in *On Forgiveness* when he distinguishes the criminal's task of redemption and the victim's task of forgiving, such that forgiveness is not the criminal's concern.[24] This part of his analysis fits well with the analogy of a self-given gift. That forgiveness is here considered a task for the victim to undertake in solitude suggests that it is an important psychological or emotional process for an individual to undergo after experiencing some harm, completely independent of whether the wrongdoer repents. This further explains how forgiveness can constitute a self-given gift, or a gift received from the higher source of love, in cases where the wrongdoer has not repented. Forgiveness is a gift to oneself in the sense that it signifies the individual's power to overcome the pain inflicted by another and to manifest the higher power of love in a particular situation.

However, in my view, the most powerful sense in which forgiveness may be considered a gift is the simple fact that we are capable of it at all. To be capable of forgiving, being forgiven, and comprehending forgiveness despite its inherent incomprehensibility strikes one as a miraculous phenomenon of mysterious, superordinate origins. The ability to forgive something completely evil seemingly cannot be classified as a purely human capacity; perhaps more than any other social, psychological, or cultural phenomenon, forgiveness brings one face-to-face with the idea of a power that is higher or better than humanity in its basic state, or at least a feature of humanity that is purely and powerfully good. The source of abounding love that finds expression in forgiveness is named the Holy Spirit in the Bible, which is the context that Ricœur invokes in his "Epilogue." Regardless of whether one understands the power to forgive through the particular framework of meaning offered by Christianity, it is an undeniably powerful element of hu-

24. Jankélévitch, *Forgiveness*, 157.

man experience that raises questions outside the realm of ordinary life, despite being something whose possibility only ever confronts us in the most day-to-day of scenarios. But again, the close tie between forgiveness and Christianity is not easily shed—forgiveness understood as a gift given to humanity by an unspecified, perhaps transcendent power (God) facilitates our enactment of love toward our fellow human, which is a core component of the Christian faith.

More specifically in relation to the issues discussed by the three thinkers, I think that unconditional forgiveness, stemming from unconditional love, is the pure and ideal form of forgiveness as Derrida maintains, though the condition of repentance establishes the interpersonal intimacy of the act and greatly facilitates the ability of any human to grant the gift of forgiveness (in the interpersonal sense of 'gift'). It is not that this condition is explicitly necessary, but rather that humans are limited in their capacity to unconditionally love their neighbour, such that observing remorse from the person who has wronged them reduces the emotional burden of forgiving. To forgive one who feels no remorse is a saintly act whose magnitude is difficult to comprehend. Conditional forgiveness thus lends itself to a useful framework by which to maintain healthy relationships and to facilitate changes in behaviour within the capabilities of imperfect human beings. Nonetheless, the ideal of unconditional forgiveness ought to be remembered. Of the three thinkers I discussed above, this aligns me most closely with Derrida, not least because he does not attempt to offer a firm resolution to the problematic of the two polarities of forgiveness. This does justice to the difficulty of the matter of forgiveness, I think, as any attempt to 'solve' the problematic of forgiveness, elucidated by all three thinkers discussed in this paper, fundamentally misunderstands what forgiveness is—namely, a paradoxical gift from 'above' the level of the human subject.

Finally, I would like to return to Derrida's analysis of forgiveness in *On Cosmopolitanism and Forgiveness* to tie my own view together more thoroughly. Derrida asserts that concepts related but not identical to forgiveness, such as reconciliation, are only comprehensible through reference to unconditional forgiveness, which is a Christian concept. But in our lives forgiveness becomes burdened with conditions of all kinds, such as legal punishments and acknowledgements of wrongdoing by perpetrators. Derrida also includes within this set of 'conditions' the act of understanding why the perpetrator of a misdeed acted as they did, and empathizing with them as a result. Even trying to understand the psychological forces at play in the commitment

of a misdeed is a way of diminishing the pure unconditionality of forgiveness, Derrida argues.[25] Through the addition of conditions of any kind, 'forgiveness' ceases to be pure forgiveness, becoming instead a process of reconciliation. Reconciliation is a hugely important process for people to undergo after a moral atrocity, yet its concrete processes are not to be equated with the true and pure phenomenon of forgiveness. They draw their meaning and intelligibility from this idea but are not equal to it. Thus, Derrida says, "It is between these two poles, *irreconcilable but indissociable*, that decisions and responsibilities are to be taken," but "it must never be forgotten… that all [reconciliation] refers to a certain idea of pure and unconditional forgiveness."[26] He notes, much as Jankélévitch and Ricœur do, that the paradox presented here begs to be investigated. For the Western reader, this evokes the Christian language of sin as humanity's fallen state, whereby it fails to fulfill the ideal of the divine even while this ideal remains in place.

Since forgiveness as we have delineated it necessarily assumes a transcendent being of some kind, the discussion given thus far leads us to conclude that whether unconditional forgiveness is possible may be a matter of personal faith. The statement 'there is forgiveness' does not mean that every individual will be capable of forgiving any misdeed committed against them—it does not even mean that any person has ever actually been capable of it. It only (but significantly) signifies that it is *possible*; that forgiveness is *available* to human beings. Individuals may or may not accept this proposition. In line with Derrida, I think that the pure concept of forgiveness has no conditions attached to it, because any condition immediately brings one into a process of interpersonal reconciliation aimed at some understanding or explanation which compromises forgiveness' central quality of coming from a boundless source of love. It goes without saying that forgiveness in the absence of any understanding, empathy, or excuse for the behaviour is very difficult for us to fathom. Accepting truly unconditional forgiveness as a real and valuable possibility seems to compromise the foundations upon which our understanding of the world is built. How can something be incontrovertibly wrong and yet forgiven without a shred of repentance or understanding? It offends our most deeply embedded instincts, such as those toward reciprocity and revenge, to assert that unconditional forgiveness is both possible and an absolute good. Thus, forgiveness stands as an eternal exception. In practice we may only be capable of conditional forgiveness, but it remains the case that this phenomenon is only comprehensible to us by reference

25. Derrida, "On Forgiveness," 49. Derrida writes, "As soon as the victim 'understands' the criminal, as soon as she exchanges, speaks, agrees with him, the scene of reconciliation has commenced, and with it this ordinary forgiveness which is anything but forgiveness."
26. Ibid., 45.

to its incomprehensible variant, which is the one that the Bible regards as a gift from the Holy Spirit.

This sense of the gift is from 'above,' as Ricœur says, not a horizontally imbalanced gift from victim to perpetrator irrespective of repentance, nor a self-given gift with healing properties. In practice, forgiveness is these things, too, but they do not characterize it in its essence. They are downstream formulations of the original form of forgiveness wherein it is the gift of its own possibility. Of course, this 'gift' metaphor implies a transcendent giver of some kind, but I will end this discussion by leaving this question open.

Bibliography

Arendt, Hannah. *The Human Condition*. 2nd ed. Chicago: University of Chicago Press, 1998.

Derrida, Jacques. "On Forgiveness." In *On Cosmopolitanism and Forgiveness*, translated by Mark Dooley and Michael Hughes, 25-60. London: Routledge, 2001.

Jankélévitch, Vladimir. *Forgiveness*, translated by Andrew Kelley. Chicago: University of Chicago Press, 2005.

Jankélévitch, Vladimir. "Should we pardon them?" *Critical Inquiry* 22, no. 3 (Spring 1996): 552-572.

Ricœur, Paul. "Epilogue: Difficult Forgiveness." In *Memory, History, Forgetting*, translated by Kathleen Blamey and David Pellauer, 457-506. Chicago: The University of Chicago Press, 2004.

Defying Questions, Defining Answers: A Sri Lankan-Tamil Canadian Diasporic Pathway to Purpuse

Vishwaamithran Ramakrishnan

This paper seeks to take an inventory of the Sri Lankan-Tamil Canadian diasporic situation to find a collective way forward; in other words, to mark out the forces within the diaspora via a Fanonian framework to facilitate a communal diasporic discussion of the question "what is to be done?" Making use of a plethora of Fanon's works and interviews conducted by the author, this paper seeks to deconstruct the impacts of the national liberation struggle, the migration, the defeat, and the effects of living in Canada's white settler-colonial society on diasporic Sri Lankan-Tamils. Emulating Fanon's "West Indians and Africans" piece, this paper proceeds in a chronological fashion, moving from stating its methods to understanding how the spirit of the national liberation struggle was carried by those who arrived in Canada, to the significant rupture that the 2009 defeat of the armed liberation effort signified.

Keywords: Fanon, Tamil Studies, diaspora, "what is to be done?," rupture

I. Introduction[1]

I cannot speak Tamil well, but my parents, grandparents, great-grandparent, aunts, uncles, and cousins all educated me on our culture, our homeland, and our people. My parents came over in the 1990s and I was born at the end of the decade, and yet I walk a strange line. Here I am: growing up in a white settler-colonial society, tolerated but never accepted, looking at the ongoing violence in the land of my ancestors, marked by the defeat of the national liberation struggle for Tamil Eelam in 2009, recognizing the privilege and power that I have been afforded by not being there. I wonder, where is my place? Do I have a responsibility to those on the island? Am I, as a Sri Lankan-Tamil in Canada, the same as a Sri Lankan-Tamil on the island? How can I or any of the diaspora begin to meaningfully talk about the national liberation struggle given our geographic, social, and temporal distance from it? Do I have to pick up the torch of the national liberation struggle, and, if so, how? If the interviews in this paper are any indication, I am not alone in my bewilderment by our diasporic situation. As a member of the Sri Lankan-Tamil diaspora in Canada, pulled apart by various currents of our transient existence based on our inability to return ("home" or elsewhere) and living in a white settler-colonial space, we yearn for political, theoretical, and social direction. As this paper will illustrate, such desires emerge from a unique set of concrete circumstances and subsequent neuroses among the Sri Lankan-Tamil diaspora in Canada. I am neither presumptuous nor arrogant enough to provide the answers to the yearning, but I want to illustrate just what this yearning is. In doing so, I hope to showcase what conditions are required for our diaspora to facilitate a discussion of the question "what is to be done?" I will begin by outlining the methodology of my research and its purpose.

I undertook my investigations in two stages. Stage one involved establishing a theoretical framework for understanding the multigenerational diasporic national liberation struggle. Stage two involved taking this framework and setting it out towards the experiences of Sri Lankan-Tamil Canadians. I drew my grounding from the works of Frantz Fanon, particularly *The Wretched of the Earth*, his other political writings and essays, and some excerpts from *Black Skin, White Masks*. My reasoning behind having Fanon as my guide takes the form of a twofold justification. Firstly, his work is of relevance to my people. There is a precedent of conducting

1. Many thanks to my interviewees for their time and insight. Without your help, none of this would have been possible. Special thanks to Dr. Alia Al-Saji for their patience, guidance, and invaluable advice, without which I would have probably quit writing philosophy years ago. Thanks to Robert Bennett Flynn for his additional edits and comments. Finally, the utmost thanks to my, parents, sister, grandparents, cousins, aunts, and uncles who ground me and support me, and to whom this work is dedicated.

histories and analyses of the Sri Lankan-Tamil national liberation struggle with Fanon's work.[2] It should then come as no surprise that there is also a Tamil translation of *The Wretched of the Earth* and that the work is quoted by Tamil cultural information services such as tamilnation.org.[3]

The second justification is best understood as a response to the potential criticism of the lack of Tamil documented sources within this paper. Though there are works and theories on the national liberation struggle, few have addressed the subject of diaspora independently. The efforts of Chelvanayakam (the father of Tamil nationalism in Sri Lanka) and the histories of Anton Balasingham do not account for the divides and commonalities that exist between the diasporas and the Tamil people of Sri Lanka.[4] The potential dangers of such a theoretical gap are especially concerning. Those undertaking histories and theoretical works on the national liberation struggle are potentially writing without recognizing their place. This could lead one to falsely assume that the liberation demands and strategies of those on the island and abroad are one and the same. This potentially gets exacerbated with each successive generation of the diaspora. One might lump those of the diaspora who are not born in Sri Lanka (such as myself), who are potentially more socially, culturally, and perhaps even linguistically different from those on the island than their parents were, into the same category. This might also lead one to falsely purport that the diaspora is more educated (due to assumptions of western education's superiority) and thus better able to push for national liberation on the island.

What this research endeavour seeks to understand prefigures any question that the diaspora could ask of the national liberation struggle. Thus, it pre-emptively addresses the conditions which give rise to these potential intellectual pitfalls. It is a matter of establishing where we are as a diaspora concerning those questions. This is inclusive of setting the ethics (the rules of dialogue) and forum to engage these inquiries and thus subsequently find direction as a diaspora. Fanon is meant to fill in this theoretical absence by being a guide for reading Sri Lankan-Tamil Canadian diasporic experiences and to inform a course forward. As a Martinican serving in the FLN, often outside of Algeria, Fanon and his comrades understood what their participation in

---

2. For example: Qadri Ismail, "'Boys Will Be Boys'," *Economic and Political Weekly* 27, no. 31/32 (August 1992): 1677-1679, https://www.jstor.org/stable/4398721; Ravi Vaitheespara, "The Limits of Derivative Nationalism: Marxism, Postcolonial Theory, and the Question of Tamil Nationalism," *Rethinking Marxism* 24, no. 1 (January 2012): 87-105, DOI:10.1080/08935696.2012.635040.

3. Frantz Fanon, ஒடுக்கப்பட்டவர்கள்: விடுதலையின் வடிவங்கள், trans. வி. நடராஜ் (Coimbatore: Vidiyal Pathippagam, 2002).

4. Cf. S. J. V. Chelvanayakam, *A Memorandum from the Tamils of Ceylon to Delegates Attending the 20th Commonwealth Conference in Sri Lanka*, transcr. N. Ethirveerasingam, tamilnation.org, https://tamilnation.org/diaspora/eelam/shan/chapter%2045.htm; cf. Anton Balasingham, *War and Peace* (London: Fairmax, 2004).

national liberation meant as a diaspora. Fanon was astutely aware of his unique and isolated diasporic place as evidenced throughout his works, and perhaps most self-referentially in his "West Indians and Africans" article.[5] Bearing in mind that a Black-Caribbean experience in relation to Africa is vastly different from a Sri Lankan-Tamil Canadian experience, Fanon's self-conscious work still makes him the ideal facilitator for this research.

To pre-emptively sum up Fanon's position, the diaspora plays an important role in the national liberation struggle, but its position is isolated and unique. Carrying a particular cultural and social baggage, which often shifts in white or settler-colonial nations, the diaspora necessarily distances itself from the lived realities on the ground where the national liberation struggle is taking place. Particularly, in the face of a defeat, the culture of the diaspora retreats and essentializes. This includes older cultural traditions as well as new myths and practices from the prior revolutionary period. The results of such moves are a strange cultural melange of some reactionary and revolutionary elements. This diasporic cultural change is further exacerbated by the lack of a post-defeat binding theoretical/institutional structure in the diaspora who, being forced into a state of permanent transience, are rendered subject to the hierarchies and stratifications (whether by class, caste, gender, or sexuality) of their own community. This understanding of Fanon and diaspora is not just an abstract theory; rather, it was experienced and reinforced by those I interviewed. This lived proof not only echoes Fanon, but also implies that the answer to the question "what is to be done?" lies within Fanon's work. Having covered my approach and the reasoning behind my research, I will now illustrate the way I read Fanon and those I interviewed.

Many of Fanon's works had a distinctly propagandistic tone. One might write off the contents of these political works (such as the articles in *El Moudjahid*) as purely functional endeavours, designed to rally support for the revolution rather than serve as espousals of or contributions to his body of theoretical work. Such a move would be erroneous. One must place these works within Fanon's own framework as discussed in *The Wretched of the Earth*. For example, many of the *El Moudjahid* articles serve the consciousness-raising task of the national liberation organization.[6] Such an organization is not a theoretical voice without ground; rather, it is tempered and continually evolving with and by the lived struggle and challenges of those who constitute it (notably the peasantry, among other oppressed classes). In other words, theory (in-

5. Frantz Fanon, "West Indians and Africans," in *Towards the African Revolution*, trans. Haakon Chevalier (New York: Grove Press, 1967), 17-27.
6. A great example of this is his article against De Gaulle's strategies of preserving French control in Algeria. Frantz Fanon, "Gaullist Illusions," in *The Political Writings from Alienation and Freedom*, ed. Jean Khalfa and Robert J.C. Young and trans. Steve Corcoran (London: Bloomsbury Academic, 2021), 75-80.

cluding Fanon's) is built from the people, making all theory praxis.[7] Therefore, one cannot look at Fanon's political writings as theoretically independent of reality when his broader national liberation framework is self-reflexive and shaped by his lived realities. The upshot of this is that one ought to read these political works and situate them within this framework of national liberation, particularly in the comprehensive *The Wretched of the Earth*.

The interviews for this effort were conducted under tight time constraints. I was unable to do more than four; however, each interview lasted for at least an hour (usually more) and went into extensive detail. I selected the participants on two broad grounds. Firstly, they should have mostly grown up outside of Sri Lanka, and, secondly, they must have made some kind of return to the island (whether for visiting or work purposes). The questions I posed to the participants were almost identical and were structured around my central thesis and findings on Fanon's notion of diaspora.[8] In the interest of transparency, I told each participant what my research was about, phrasing it simply as: "Fanon's understanding of the multigenerational diasporic national liberation struggle." I was not particular about the ordering of the questions or the rigidity of the inquiry; rather, I wanted the participants to enunciate and derive their own conclusions in whatever manner they saw fit. After all, discussions of the national liberation struggle and diaspora are deeply personal and cannot be fully captured in blunt answers to rigid questions. As this paper will show, this manner of response is also related to the practical concrete situation of the diaspora.

The first person I interviewed was an educationalist who spent most of their education abroad and is currently working on the island. For the interview and paper, they have requested to be anonymous and will go by 'Mr. X.' It should be noted that they have some experience with Fanon; however, I mainly drew upon his lived experiences on the island as a means of considering how to facilitate asking the question "what is to be done?" Sumu Sathi was my second interviewee. Having left the island at fourteen, she currently works as an anti-racism activist and filmmaker in Canada. She is currently in the process of releasing her latest film, *Thamaraigal*, which is about the ongoing human rights issues on the island. My third and fourth interviewees, who opted to be interviewed together, were Ann and her brother (who has opted to remain anonymous and will go by Mr. G). Both have returned to the island in different capacities. Mr. G and Ann both returned for a family trip during the 2004 ceasefire, but Ann made an additional trip in

7. Frantz Fanon, *The Wretched of the Earth*, trans. Constance Farrington (London: Macgibbon and Kee, 1965), 112-114.
8. See the Appendix for a sample of the questions asked.

2013 to conduct interviews as a freelance journalist.

I will now turn to the groundwork necessary to ask the question "what is to be done?" in the case of the Sri Lankan-Tamil Canadian diaspora. The first step will be to take an inventory of the diasporic state before the defeat. The second step will be to illustrate what the 2009 defeat of the Liberation Tigers of Tamil Eelam (LTTE), the main fighting force and organization for the national liberation struggle, has done to the diaspora. The third step will be to set out the necessary conditions of possibility to meaningfully facilitate asking the question "what is to be done?" Since the interviews back up Fanon's positions, both angles will work in conjunction to support this analysis.

II. What Was Brought, Given, and Made

When many Sri Lankan-Tamils arrived in Canada, they carried the spirit of the national liberation struggle with them. For many, the departure and exile were temporary setbacks. For some, it was a chance to raise the next generation outside of the material dangers of the conflict. For others, it was a chance to earn some money to send back to the families on the island, some of which inevitably went to the war effort. In any case, leaving was not a defeat. Many, including some of those in Ann's family, expected to eventually go back. This sentiment was reinforced as the national liberation struggle's mechanisms followed those who left. The diaspora established numerous businesses and organizations to fund the struggle back on the island. Money earned in dire financial circumstances at minimum wage would ideally be turned into ammunition and bandages. For example, when he was first able to work as a young adult, Mr. G gave half his earnings to the struggle.[9] The struggle also followed the diaspora via media endeavours. Tamil news organizations linked with the LTTE (i.e., National Television of Tamil Eelam) brought daily updates on the struggle in the form of films and newscasts. Many, like Ann, remember the reports of the struggle being aired continuously on their home television sets during the war.[10] We can recognize these manifestations of the national liberation struggle as facets of a crucial liberation mechanism that Fanon discusses—political organization. The Sri Lankan-Tamil did not just bring the effects and spirit of the national liberation to new locations, but also inherited the organizational structure of disseminating knowledge and materially contributing towards the struggle on the island.

9. Mr. G (accountant and brother of Ann), interviewed by Vishwaamithran Ramakrishnan, November 26th, 2021.
10. Ann (freelance journalist based out of Canada), interviewed by Vishwaamithran Ramakrishnan, November 26th, 2021.

Putting aside for the moment the question of whether this organization was truly effective, the fact remains that a focused and unified theoretical/political direction towards liberation remained in the community after arriving in Canada. As Fanon explains with regards to the latter stages of the national liberation struggle, the liberation organization plays a crucial role in maintaining and reflecting the political cohesion of the struggle across the varied spaces of the nation.[11] As a diaspora, our challenges are a bit different from what Fanon was discussing. The need for political cohesion persists but differs because of the distance in circumstances and space between the diaspora and the actual place of struggle. The organization's structure and media offered a way for the diaspora to communicate and enter the liberation struggle despite being across the oceans. Fanon echoes this sentiment in his discussion of the radio's role in the Algerian Revolution. The radio's adoption by Algerian households was due to the nationwide inclusive message that the Voice of Free Algeria's broadcasts offered, linking those from Batna or Nemours to a single national struggle.[12] Tuning-in meant partaking in the struggle by garnering knowledge of what was going on and hence having the ability to discuss the state of national liberation with one's friends and family. Listening to the radio became a way to build a national community, changing oneself from oppressed to free in embodying the spirit of the new nation as its epic unfolded over the airwaves.[13]

Similarly, watching daily reports of the struggle on TV during family get-togethers allowed the Sri Lankan-Tamil community to be a part of the national liberation struggle, to be unified, and to join in the transformation of our identities from Sri Lankan-Tamils to Eelam-Tamils. But there was nevertheless a distance. All activities, whether the fundraising of money, media publication, or outreach, were geared towards the island, and there was thus an inherent directionality of action between the diaspora and the people on the island. In other words, though we were absorbed into the totalizing and unifying effects of the national liberation struggle and organization, we were still on the outside. This directionality therefore created a different set of experiences and conditions for the diaspora versus the people on the island. Furthermore, these effects did not influence all parts of the diaspora in an equal fashion, especially along generational lines.

Sumu Sathi relayed the important role that the LTTE played in her life. Having been born

11. Fanon, *The Wretched of the Earth,* 105-106, 114-115.
12. Frantz Fanon, "This is the Voice of Algeria," in *A Dying Colonialism,* trans. Haakon Chevalier (New York: Grove Press, 1965), 82.
13. Ibid., 84-85.

in the capital, Colombo, her mother reinforced the importance of the national liberation struggle. Because the LTTE was fighting in the north and east, the Sri Lankan government, fearful of bringing the conflict to the state's heart, would not dare touch the Tamils living in the capital.[14] Though he takes great pride in being Tamil, all that Mr. G heard about Sri Lanka was war stories. It painted a picture of a place where he was glad not to be. For Ann, who has some memories of the island, the struggle marked them all. Before 2009 her relation to the land of her ancestors was one of pride and willingness to even go back and fight. For both Mr. G and Ann, this national pride is related to their parents and family who instilled it in them. I too grew up with a sense of pride in my culture, but the interviewees and I, who grew up in Canada, also had to contend with a white society as children. Ann, who vividly remembers her journey to Canada, recalls becoming a member of the diaspora, in the sense of realizing one's displacement from the land of their ancestors, the moment she left her house. Canada was presented as a safe haven, but as soon as she went to elementary school she was called a "paki" and ostracized. "Tamilness," national culture, the broadcasts, and the organizations could not defend her from the gaze of white eyes. She was affixed to her skin into what Fanon would describe as an 'epidermal schema'—an essentialization of the Other to their skin colour by white eyes (structural or interpersonal). This epidermal essentialization is underlined by a historico-racial schema, or in other words the myths, halftruths, and histories constructed by the colonizer). In Ann's case, her being fixed to her brownness by whiteness was underlined by Apu and suicide bombers. Her body schema, her embodied comportment in the midst of the world, was shattered. She existed in a state of transience where her identity was twisted and mangled, forced to recognize that the land and space where this white creation of brownness exists did not want her for herself.[15] Unlike Fanon, who is contending with his historico-racial schema (and in very different circumstances as a black Martinican) for recognition as a human being among other human beings in *Black Skin, White Masks*, Ann was sure of her Tamil identity in relation to the prior mentioned diasporic mechanisms and history, though it was not recognized by the dominant white society. Yet, like Fanon, she underwent a neurosis of herself that very much had to do with her race.[16] Her "brownness" and exoticized manner of speech forced her to look within, as if there were something wrong with her. She still had to navigate a white settler-colonial society to survive while not losing her-

14. Sumu Sathi (anti-racism advocate), interviewed by Vishwaamithran Ramakrishnan, November 22nd, 2021.
15. Frantz Fanon, "The Fact of Blackness," in *Black Skin, White Masks*, trans. Charles Lam Markmann (New York: Grove Press, 1967), 111-112.
16. Ibid.

self. Her response was to become the quiet kid in class and to keep her head down; yet, despite this, the whiteness still burned her as she became aware of herself and her body as a member of the diaspora.[17] It meant walking a line between the epidermal schema of a brown girl in school and her Tamil identity within her heart and home, all the while defending herself by keeping watch of who was watching her.

Mr. G's diasporic realization occurred with a similar public-school experience. However, his response was that of an outburst and the marking of battle lines.[18] In grade 9, he entered high school and encountered several Tamil people who came from surrounding feeder schools. He referred to his time with his Tamil friends as a time of war—it was the Tamil students against "them."[19] "Them" referred to any other cultural groups, though his school was mostly white (with some exceptions, such as the black students who often backed Tamil students up). They would get outnumbered and beaten, then they would retreat and counterattack, operating under the rule that an offence to one is an offence to all. The Tamil students were united and self-aware of their place as a diaspora in this conflict. Recalling *The Wretched of the Earth*, this communal and violent effect of the diasporic neurosis is akin to the communal muscular tension (the physical tension of the colonized in the face of colonial violence pre-empting retribution) of the colonized during the colonial period where clear divides exist, such as the world of the colonized versus the world of the colonizer. We escaped from a place (or at the very least were informed of a place, as in my case) where the whip of oppression embodied in the police stations, riots, and pogroms of Sri Lanka exploded into conflict, to only then be thrown into another colonial situation which, though differently, continues the cycle of muscular tension and violent release.[20]

This is further compounded with the inherent sense of displacement that follows diasporification and our state of geographical transience. The Sri Lankan-Tamil Canadian presence is only provisionally tolerated in Canada. Sumu Sathi's upbringing in the predominantly Tamil areas of Scarborough also speaks to the enactment of this diasporic mechanic. She recalls the bloody and violent feuds between Tamil gangs such as AK Kannan and VVT, whom she interpreted as fighting a war to carve up their respective turfs. This intergang conflict demonstrates an explicit link between the diasporic neurosis' muscular tension and the desire for permanent space.

17. Ibid., 114.
18. Ibid., 114-115.
19. Mr. G, interview.
20. Fanon, *The Wretched of the Earth*, 40-42.

It might be suggested that this is a poor equivalency to make on a diasporic scale, given that some members of the diaspora (such as myself) have never truly known the violence of the island. Though this observation is correct, it does not take proper account of the transmission of tension across generations which forces the inclusion of a literal muscular tension into the diasporic neurosis. To return to a point made above, the violence of the war was something that everyone knew about. Though not all families spoke of their traumas and events with their children—I am grateful that mine did, many of us kids knew our people were being attacked in Sri Lanka and it was framed as the fault of the government or of the Sinhalese. This became especially clear as the protests and media attention blew up in the final days of 2009. Fanon is clear that children who are isolated from the main theatres of violence are not immune to the totalizing effects of war. Fanon's examination of a case in which two Algerian children murdered their European playmate demonstrates violence's transmutability. Neither of the children could come up with a reason to kill their friend other than that he was a European that they could kill. One of the children (the thirteen-year-old) even stated that not a word was spoken against him by the European child's policeman father. In other words, he had no direct link or motive to commit the act apart from the violence and oppression of the colonial situation.[21] I do not mean to insinuate that this extreme example is a perfect analogy for the continued muscular tension in the diasporic neurosis: I mean merely to point out that the cycle of muscular tension and violent release permeates even those who appear isolated. Even Mr. G, who did not remember much of the island, could not pinpoint whether the cycle of violence at school came from the diaspora's collective experience of the war, but what he did know "was that our culture was being attacked."[22]

Beyond muscular tension, communal solidarity, and the discovery of one's own transience, the diasporic neurosis forces a communal introspection and preliminary cultural essentialism, making norms or customs stagnant. Having grown up in Colombo, Sumu Sathi spoke Tamil as well as Sinhalese when she came to Scarborough. She recalls being ostracized by some of the diaspora for her competency in Sinhalese. As for me, I was made fun of by some of my family because I could not speak or write Tamil fluently. Ann and Mr. G did not have early morning weekend cartoons to look forward to, but their mom taught them to speak and write Tamil before she had to leave for work. Beyond language, many Tamil children, like my cousins, were forced by their parents to go to Tamil cultural activities to come to grips with their culture.

It is also worth mentioning that reactionary attitudes such as misogyny (a common exam-

21. Ibid., 217-219.
22. Mr. G, interview.

ple of which is the chastising of women but not men for having sex out of wedlock) are included in this cultural trove. Though perhaps not directly expressed as 'Tamil culture,' they are passed off by the community as merely "the way that we do things." Mr. X confirms this. He notes that when the diaspora left the island, they also took with them pre-Sri Lankan independence era values, including sexism and casteism.[23] Piecing these cultural testimonies together forms a common picture and a diasporic demand to preserve and police one's cultural identity. As Fanon explains, the colonized hold onto cultural artefacts and traditions as a kind of defence mechanism in the face of the dominant cultural onslaught of the colonizer. To hold onto one's contested culture in a colonial situation, even if it is broken into pieces and essentialisms such as requiring Tamil to be spoken in one's home, is to express nationality and resistance.[24] Especially for the Sri Lankan-Tamil Canadian diaspora, coming from a context in which one's culture was deliberately under siege (under the Sinhala Only Act for example), it is understandable why this cultural entrenchment accompanied the journey to Canada. In addition, the inherent conditions of being a diaspora merely tolerated by a white settler-colonial society—one that forces them to speak a foreign tongue, subjects them to higher rates of police stops, and classifies them into a 'brown lump' alongside other racialized groups—the colonized diaspora is pushed to further fortify. In other words, it is a cultural-political necessity to make this essentialist move. Unlike the colony, in which the native could continually live amongst their people, the diaspora member, particularly the young diaspora member educated by white settler-colonial institutions, has a murkier path. They both must be aware of themselves and their "Tamilness" among their own people while trying to survive and be accepted by white society. Thus, self-consciousness seems to be a part of the diasporic neurosis. The diaspora is forced to straddle their transient existence between a white disfigurement of their bodies and the cultural essentialisms born out of communal necessity.

When our people came to Canada, we brought the rhythm of the new nation with us. We set up businesses and media outlets to support the fight on the island as we stayed unified in the political organization provided by the national liberation struggle. Yet, we faced new challenges as a tolerated diaspora in a white-settler colonial society. We were forced into a self-conscious diasporic neurosis in which we found communal solidarity, our transient existence between our

23. Mr. X (educationalist based out of Sri Lanka), interviewed by Vishwaamithran Ramakrishnan, November 18th, 2021.
24. Frantz Fanon, "Racism and Culture," in *Towards the African Revolution*, trans. Haakon Chevalier (New York: Grove Press, 1967), 42-43; Fanon, *Wretched of the Earth*, 190-191.

Tamil identities and our disfigurement by white society, and a continued cycle of violent muscular tension. Though our situation was dire, we held out in hope for tomorrow. However, all this would change in 2009.

III. The 2009 Rupture: The Beginning of the Defeated Diasporic Neurosis and a Split

May 2009 saw the end of the armed struggle for national liberation. The LTTE over the last months of that year lost key areas of the island and eventually their administrative capital at Kilinochchi, culminating in a final stand at Mullaitivu. This turn meant many things for the Sri Lankan-Tamil community in Canada, notably the addition of the 'defeated' status to the existing diasporic neurosis and an irrevocable split between Sri Lankan-Tamils on the island and those abroad. Focusing on the 'defeated' status, one of its most obvious effects is the lived, *permanent* transience of our people. To return to Sri Lanka under the oppressive regime post-2009 was not a feasible option. In the words of Sumu Sathi, "There is no place to call home. We are forever unsettled."[25] Instead, creating spaces, getting relatives out, and surviving in white settler-colonial Canada became the primary objective. The newfound permanency of our transience made it harder to return the white man's calls to "go back to where you came from!" and to direct ourselves to the lands of our ancestors for grounding and identity. We further shrank into our skin as a result. The spatial elements of this 'defeated' status also carried a socio-cultural element. It signified that we would have to be forever culturally entrenched as long as we remained in the space and under the dissecting gaze of Canada's white settler-colonial society. However, the question of how to reckon with the struggle's legacy, to what end and what for, is a non-spatial, purely socio-cultural facet of our status as defeated.

With those final shells and bullets, the structure, direction, and organization that the LTTE provided were also shattered. As Sumu Sathi notes, "we were left with no representation, identity, or direction."[26] Ann echoed these views, describing how the LTTE formed the political focal point of our culture, especially when it came to dismantling casteism and promoting equal opportunities for women. With their defeat, she says, many of these positions were abandoned and all manner of opinions came forward as to what happened during the war. As a result, many started to express their own individualized experiences of the struggle, including anti-LTTE positions, leading to contentious debates. Even if one were to contest just how much of an impact

25. Sumu Sathi, interview.
26. Ibid.

the LTTE had in structuring the struggle, the defeat of 2009 broke the diasporic momentum that had hitherto supported the national liberation struggle. The active armed effort, the main force of political action that engulfed and provided direction to the diaspora, had ended. National liberation on the island was no longer a certain aim for the diaspora. In this state of confusion, hostile discussions continually crop up with regards to who is more "Tamil" than another based on one's position towards the struggle. There is a structure to this confusion—who gets to speak and proclaim "direction," "Tamilness," or "truth of the struggle" is determined along numerous lines of social conflict, such as gender, caste, and class. Each speaker, based on their own personal vested interests, sets how one might invoke the struggle; for example, one's position on the struggle has been essential for many Tamil election campaigns in Canada.

As Sumu Sathi explains, the post-defeat diaspora can be divided into three broad categories with these social hierarchies in mind: the traumatized, the radicals, and the sellouts. The traumatized are those who were so damaged by the defeat that they withdrew themselves from the struggle and the past, searching and striving to move forward by whatever means. The radicals are those who refuse to believe that the armed struggle has ended, holding from their diasporic distance (whether cognizant of it or not) that armed struggle is the only way forward at this time. The sellouts are those who generally aim to profit from the post-defeat situation, in many cases with no regard for the impact of their actions on their own people. These responses to the end of the national liberation struggle, influenced by the hierarchies discussed above, have collectively had the effect of sidelining or deprioritizing the truth of what actually occurred. They have rendered it impossible for the diaspora to collectively address the legacy of the struggle in order to find a path forward. This obfuscation of collective truth is embodied in the 'defeated' status in the form of a continual yearning. Since the national liberation struggle has characterized much of the Sri Lankan-Tamil diasporic self-understanding, this yearning can be expressed as a desire to know oneself and one's people. Sumu Sathi's journey to Sri Lanka, undertaken to conduct interviews and to find her people and their truth amid these convoluted diasporic currents, speaks to this facet of the 'defeated' status.

One might wonder how the rhythm of the nation was impacted by these changing circumstances, especially the acquisition of the 'defeated' status. To investigate this question, one must remember that the rhythm of the nation is the product of the truth (or natural rhythm) of the people being rendered into collective action, the people making themselves anew in the process.[27] In

27. Fanon, *The Wretched of the Earth*, 28, 105-106.

the diasporic situation, one is physically isolated from the realities on the ground. National truth was only understood so long as the national liberation organization and the broader movement informed and dictated it. But once the defeat broke the transmission of the rhythm of the nation from island to diaspora, this truth became a socio-cultural artifact that backgrounds the diaspora. Fanon's analysis of the rural peasantry is especially enlightening in this regard. Despite each defeat, this class holds on to the dear myths and memories of the last anti-colonial action.[28] However, unlike the peasantry, which daily carries the lived truth of colonialism's violence and keeps those cultural memories alive as a resource for struggle, the diaspora has neither the lived truth, nor therefore the *place* to act upon the memories of the prior struggle. Their memories have lost their link to the natural rhythm of the people on the island. Stuck in a space of situational inaction and forced disjunction with the truth, the rhythm of the nation (as felt and spoken by the diaspora) becomes a stagnant one of tired slogans, the occasional rally, and a few privileged analyses that have no bearing towards the lived realities on the ground.

Nevertheless, the artifact, despite its stagnancy, constitutes the diaspora's origins and beginnings and is therefore crucial; without it, no notion of the diaspora could exist. The artifact becomes one of the many cultural relics that the Sri Lankan-Tamil Canadian diaspora must invoke to continually defend themselves from the cultural onslaught of white settler-colonial society. When one introduces themselves as Sri Lankan-Tamil to a roomful of white people, the war automatically follows them, not as a dynamic unfolding journey of becoming a new being, but as a stagnant anchor against the white waves. This (not inherently bad) facet of the 'defeated' status is highly transmittable across successive generations as stories told between families, and yearly remembrance activities serve as social fixtures for the community.[29] Nevertheless, this revolutionary facet of the culture falls victim to collusion with reactionary facets, in that they both serve the purpose of cultural defence. For example, it is not uncommon for a supporter of the LTTE to hold on to patriarchal values in the name of Tamil culture, despite the LTTE's commitment to gender equality in the struggle.

Fanon also engages with the complex position of revolutionary struggle in a people's history. As he explains with regards to the initial Algerian uprising of November 1st, 1954, the significance of the event derives from its place in Algerian history and its aim towards the defeat of French colonialism. Its relation to the destiny of liberation comes from the masses catalyz-

28. Ibid., 90.
29. For example, Mullivaikkal Remembrance Day events serve to remember those who died in the last battle of the national liberation struggle.

ing themselves towards a self-transformative task, giving birth to a new kind of being.[30] When there is no continued catalyzation or activity, any futurity that an initial or prior liberation action entails gets put aside, but it is not thereby removed from the broader history of the nation. Similarly, the artifact of the national liberation struggle for diasporic Sri Lankan-Tamils in Canada only has any relevance to the future of our people insofar as we are catalyzed into acting upon it concretely. Since we are no longer able to directly or meaningfully act upon the artifact due to the disruption of truth and our physical distance from the island, the earlier armed struggle becomes a piece of broader identity-defining history which we seemingly cannot activate on our own. Ironically, it is for diasporic Sri Lankan-Tamils in Canada at once a binding factor and a point of extreme interpretive contention. Yet, regardless of whatever angle one takes, the artifact is predicated on the island's situation. Until such a break or re-animation can be made, it will remain merely stagnant. Thus, the addition of the 'defeated' status to our existing neurosis leaves the diaspora yearning for direction and dynamism.

This turn is underlined by the fundamental split caused by the 2009 defeat. In the words of Ann and Mr. G, "We (the diaspora) felt and witnessed the loss but they're (Sri Lankan-Tamils on the island) living the loss more materially."[31] The disturbance to the national rhythm ruptured the link between those on the island and those abroad. Since 2009, when the older members of the diaspora speak of "back home," their words seem especially hollow. At that point, Sri Lankan-Tamils in Canada and Sri Lankan-Tamils on the island went down two related but separate roads due to the material circumstances of the struggle. Since the artifact of struggle is still stagnant, with neither the circumstances nor the egalitarian framework present to meaningfully reanimate it, the diaspora is left thinking that their path is the same as those on the island. In other words, the diaspora misidentifies its own place in the struggle.

This often produces a dynamic in which members of the diaspora come to view the diaspora as containing leaders who will be able to dictate and lead the struggle from abroad. As Sumu Sathi explains, this is especially the case for the radical parts of the diaspora, who feel that continued armed struggle is the immediate next step. She further explains how much of the remembrance activities and flag-waving politics of the diaspora seem to forget that such actions are not possible in Sri Lanka, where even mentioning actions during the prior struggle can constitute grounds for a disappearance. As Fanon explains, the West Indians (more specifically the Antil-

30. Frantz Fanon, "Algerian revolutionary consciousness," in *The Political Writings from Alienation and Freedom*, ed. Jean Khalfa and Robert J.C. Young and trans. Steve Corcoran (London: Bloomsbury Academic, 2021), 45-46.
31. Ann, interview; Mr. G interview.

lais) who threw themselves into Africa to find themselves were marching towards a dead end. Whether they are accepted or not, the West Indian who leaves for Africa in the hopes of being incorporated into a constructed vision of their missing truth eventually falls prey to a mirage of themself.[32] Though the Sri Lankan-Tamil context is not as extreme as the legacy of slavery and its corresponding search for truth, an analogy can be made. Just as it makes no sense for the West Indian to grasp at Africa on the basis of a stagnant and dead version of the truth in the hopes of being accepted by the African, it makes no sense for the Sri Lankan-Tamil diaspora to approach and guide those on the island based on a stagnant cultural artifact, as those on the island are actively reckoning with the truth to which the artifact fails to measure up. The consequences of this misidentification are that it becomes much harder to address unique diasporic struggles (since we are focused on the mirage) and to imagine what meaningful roles the diaspora could play given our distance from those on the island.

2009 ruptured our people. It added a 'defeated' status to our existing diasporic neurosis, solidified our transience, entrenched the struggle as a cultural artifact, and sparked a yearning for the truth amidst a chaotic discourse populated with vested interests. Yet, having accounted for these factors, it is still hard to reckon with the sheer totality of what 2009 meant to our people. As a member of the diaspora, dealing with the defeated diasporic neurosis, I readily admit my inherent limitations and weakness. However, my outline of the situation itself already contains some steps towards finding a way forward. In the case of the Algerian Revolution, Fanon notes the individual's role in the national liberation struggle. The colonized individual, without the mass national consciousness of a communal effort, is unable to break the experiential bonds of colonialism. Nevertheless, authentic national liberation only moves ahead insofar as the individual sets out on the path of self-liberation (educating themselves, taking revolutionary action, etc.).[33] Therefore, anyone trying to reckon with the situation of our diaspora, in some small and limited way, brings us that much further towards liberation. With the state of the diaspora on the table, this paper now turns to its active component, unpacking how to facilitate asking the question "what is to be done?"

IV. Where to Now? Facilitating the Asking of "What Is to Be Done?"

In unravelling this neurosis, a reckoning with the truth of the national liberation struggle,

32. Fanon, "West Indians and Africans," 25-27.
33. Frantz Fanon, "Decolonization and Independence," in *Towards the African Revolution*, trans. Haakon Chevalier (New York: Grove Press, 1967), 102-103.

while crucial, seems to be only a part of the puzzle. Given the many nuances of the defeated diasporic neurosis discussed above, a reckoning with our diasporic place and situation seems important. Therefore, a return to Sri Lanka does not seem to be the ultimate avenue for facilitating asking the question "what is to be done?" Furthermore, it is not a complete means to answer our question. As Fanon explains, the journeys of black West Indians (who often served in higher-level posts in French African colonial society) back to Africa are sure to be met with African ire because the African sees the West Indian as a sellout who tried to play white.[34] Similarly, Mr. X describes how many of the Sri Lankan-Tamils on the island resent those who left for not seeing the struggle through. Sumu Sathi's experiences with interviews on the island corroborate this, showing how only a certain class of people could afford to leave the island. This added a class-based layer of resentment to the island's relationship with the diaspora, rendering such a straightforward diasporic return all the more untenable.

However, that is not to say that return or re-engagement with the island is off the table. As Mr. X explains, the generational divide between those who left the island during the struggle and those who remained is not as large as that between West Indians and Africans. Much of the diaspora still has connections to the land where they grew up, but it is also a land that is more alien than home for many of those people's children, including myself. This complexity was embodied by Mr. X, who always intended to return to the island and spent most of his life there, but still felt like an outsider due to his experience abroad. Sumu Sathi, on the other hand, described her journey back to Sri Lanka as part of a healing process in which she reckoned with the guilt of not knowing her people and supported those who had fought for her in the liberation struggle. Yet, despite this journey, she still describes the lingering transience and other facets of the neurosis as not entirely resolved.

Fanon echoes this unresolved tension and points towards a specific manner of engaging with the national struggle on the island. As he makes clear from the West Indian and African case study, two categories of people exist who are not truly bound to their skin *a priori*.[35] Putting to one side whether he takes "West Indian" and "African" to be *a priori* national categories, it is clear that Fanon thinks that these are two different peoples which should not be subsumed into the single, white-defined epidermal racial category of "negro."[36] Nevertheless, he is clear that despite Africans' reluctance to accept West Indians into a common fold and towards a common

34. Fanon, "West Indians and Africans," 25.
35. Ibid., 17-18.
36. Ibid.

struggle, trans-national post-colonial solidarity is an absolute necessity. In the concluding portion of "A Letter to the Youth of Africa," he calls upon the youth of the West Indies and Africa to unite and defeat colonialism.[37] Earlier on in that letter, he points out how liberation must be earned on the national soil and that every colonized person must achieve that aim.[38] Granted, one might state that this demand is specific to 1958, the time in which Fanon was writing. However, such a demand for national liberation, as a task grounded on national soil, is echoed throughout 1961's *Wretched of the Earth*, written as optimism for global decolonization struggles was vastly changing directions. Key to these changing winds was the failure of certain revolutions and their independence parties to include and uplift most of the nation (contained in the rural areas) due to their urban centricity. This speaks to the importance of being present on the ground with the whole of the nation insofar as one is engaging in a national liberation struggle. The consequences of failing this inclusive demand include, but are not limited to, the newly independent ex-colonial state quickly turning fascist through the middle class of the capital's domination of the economy at the expense of the dissatisfied rural majority. Thus, isolation, whether in the colonial/ex-colonial metropole or abroad, cannot be the place from which successful national liberation is achieved.[39] Having addressed this focus on national soil, I will now turn to interpreting Fanon's concluding remarks in his letter and his early essay on "West Indians and Africans."

For different national groups, the tasks of decolonization and liberation manifest in unique ways according to the particular set of colonial circumstances at play; the West Indian colonial situation differs from the Congolese, and so the tasks necessary to bring about national liberation in each case will differ accordingly. Each must reckon with their own national situation in solidarity with each other, since all decolonial struggles are linked through the broader goal of defeating colonialism to liberate humanity.[40] To situate these examples of decolonization in relation to the Sri Lankan-Tamil diaspora, we must address in a nuanced manner our own colonial situation of living in a white settler-colonial state while assisting in the struggle on the island. We must address our own unique truth (artifacts of struggle and all) as a diaspora while we also assist and follow the lead of those on the island, fostering our shared link and liberatory destiny

37. Frantz Fanon, "Letter to the Youth of Africa," in *Towards the African Revolution*, trans. Haakon Chevalier (New York: Grove Press, 1967), 118-119.
38. Ibid., 114.
39. Fanon, *The Wretched of the Earth*, 115, 119-120.
40. Fanon, "Letter to the Youth of Africa," 113-114.

as Sri Lankan-Tamils.[41]

Such a positionality partially breaks the island-predicated national liberation cultural artifact from the island. In recognizing the distance and nuance of the diasporic situation, our artifact of the prior national liberation comes to be recognized *as* our artifact, and not a perfect representation of the national liberation struggle as a whole (including its future). In other words, it is a condition of possibility for the diaspora to recognize the artifact as truly its own creation and part of its own unique truth. However, this brings us back to the very beginning of facilitating asking the question "what is to be done?" As possibly complicit immigrants in a white-settler colonial state that is not our national soil, yoked to the artifact of the national-liberation struggle, how do we attain the point of discussing what this liberation looks like, given our distance from those on the island and the nuance that our engagement with the national liberation struggle demands? In any case, one thing is clear: liberation from the defeated diasporic neurosis and our colonial situation does not solely entail a return or pure re-engagement with those on the island. Instead, a frontline exists—notably, in Canada—that must be addressed.

Describing the frontlines of our diasporic liberation is a difficult task in a settler-colonial nation. By virtue of living on the land and working in Canadian society, the diaspora necessarily collaborates within a racially hierarchical colonial system that routinely dispossesses and harms Indigenous peoples within its borders. This settler reality, which turns the diasporic immigrant into a colonial collaborator is something yet to be unpacked by the diaspora. More strikingly, a discussion of colonialism's impacts on the diaspora and the history of the national liberation struggle on the island has yet to become a common discussion in the diasporic community. As Ann explains, our diaspora still talks as though the "enemy" are the Sinhalese or the present government when the root of the problem, which gave rise to the violent call for liberation, is colonialism. We commonly fail to recognize that it was the same British Empire that ruled over the island, carved its borders, and created its modern commodity production system (i.e., tea, among other products) that dispossessed the Indigenous peoples' lands on which we live today.

We further fail to realize how our own people have collaborated with colonialism for their own benefit (i.e., absconding with the funds of the national liberation struggle or engaging in political opportunism). To return to Fanon, we as a diaspora must raise our political consciousness to address the present state of struggle. As stated above, the pre-independence colonial

41. The question of who constitutes a Sri Lankan-Tamil on the island is complex, but the point on the diaspora having a distanced position and directionality towards such persons (whoever they are) and their struggle stands nevertheless.

values still linger on in our society, and the strict Manicheanism of the unequivocally "righteous" militant Tamil and the unequivocally "evil" Sinhalese persists in our community; we have not truly elevated ourselves in this regard.[42] Concrete political tasks of analysis and reflection are thus crucial because they aid us in situating ourselves within the larger colonial system, such that we can defeat it as a diaspora in conjunction with Indigenous peoples and other colonized allies. The relevance of these tasks towards alleviating our diasporic neurosis is self-evident. In aiding our Indigenous and colonized allies, we can challenge the material situation which gives rise to many facets of the neurosis, such as the historico-racial schema. Sumu Sathi's experiences from working as an ally to BIPOC (Black, Indigenous, persons of colour) people speak to this. She describes this allyship as a form of healing amid the diasporic confusion and lack of safe spaces with regards to unpacking the truth and realities of the national liberation struggle. Despite the benefits of such an avenue, she remarks that the fundamental lack of truth and the baggage of the national liberation struggle still weigh heavily upon her. Her experiences illustrate a key point. These calls for solidarity, analysis, and a changed relationship with the Tamils on the island must also address the artifact of the prior national liberation struggle. It entails that we must unpack our origins as a diaspora, self-activate the artifact of the struggle, and make sense of its truth such that we can ground ourselves and thereby know where to march next. Such grand prerequisites for answering the question "what is to be done?" therefore point us towards the difficult task of reorganizing ourselves as a diaspora.

V. Reorganization and Conclusion

The community has lacked unity for some time, yet it must find direction. I do not intend here to begin a discussion on the methods of practical organization, but for such a discussion to occur at all, some preliminary key principles must be considered. As expressed earlier, much of the discussion around the artifact of the national liberation struggle occurs along social hierarchical lines, resulting in voices being silenced and private agendas being pushed to the fore. As Fanon explains with regards to the newly independent bureaucratic one-party state, whether in Africa or elsewhere, the truth of the nation and the struggle is a collective endeavour.[43] We must all search for it and engage with it, while also being careful not to dogmatically force ideas upon those who are behind in their political education. Leadership and hierarchies are not forced here but rather emerge organically and become legitimate through the exercise and engagement of this

42. Fanon, *The Wretched of the Earth*, 110-111.
43. Ibid., 160.

truth.[44] Therefore, what is required to facilitate asking the question "what is to be done?" is not a dogmatic and dictatorial party, association, program, or congress, but rather a collective, non-institutionalized, and democratic endeavour. Through such an organizational structure, we can collaborate, congregate, and analyze the truth of the national liberation struggle and our place therein. In short, it will allow us to reanimate the artifact for ourselves on equal terrain.

My interview subjects further reinforced this position. As Mr. X explains with regards to how the diaspora can best help those on the island, diasporic democratization along economic and social lines is essential. Opening and organizing ourselves along our diverse experiences with the national liberation struggle can facilitate action in a manner that reflects the diverse realities, identities, and states of affairs on the island. As Ann and Mr. G both explain, we need to listen to each other, sharing our stories, identities, and positions in an equitable democratic framework—or else we leave the door open for our diverse experiences to be utilized by political opportunists. This latter point also alludes to a general democratic necessity of the diaspora to be hyper-vigilant and critical of themselves and those who would seek to make a personal opportunity of such an organizing effort.

To sum up what has been said from the investigations, and to suggest on their basis a possible path out of the defeated diasporic neurosis, a vast reorganization and repositioning of the diaspora must be undertaken. It seems that an attempt to relink with the struggle on the island on pre-2009 terms is not practically nor logically desirable. In the wake of the defeat, the distance between the Tamils on the island and the diaspora abroad must be acknowledged. This enables a refashioning of the prior national liberation struggle artifact as one of our own conscious creation and utilization, which will allow for a new relationship of joint struggle among those on the island and the diaspora. A necessary condition of this is the social and political task of unpacking our own diasporic position in a white-settler colonial society aiming at concrete action in allyship with Indigenous and other colonized groups. These tasks—and new angles of relating—require a cohesive collective effort and are predicated upon democratizing the diaspora (including breaking down social hierarchies) to be able to reckon with and ground ourselves in the cultural artifact of the prior national liberation struggle. In doing so, perhaps we can then ask and answer the question "what is to be done?" for our diaspora.

As Fanon reminds the next colonized generation who will fight colonialism, each genera-

44. Ibid., 160-161; Frantz Fanon, "A democratic revolution," in *The Political Writings from Alienation and Freedom*, ed. Jean Khalfa and Robert J.C. Young and trans. Steve Corcoran (London: Bloomsbury Academic, 2021), 38-39.

tion must discover its role and decide whether to betray or fulfil its mission. Given their tools and resources, we cannot fault the prior generation for being "too passive" or for missing the mark.[45] Nevertheless, the diaspora's circumstances and challenges are vastly different from when our parents first arrived. The issues of the defeated diasporic neurosis and the yearning that it entails grow ever more pressing. Questions of place, direction, identity, responsibility, and future in a white settler-colonial society will not disappear anytime soon. Thus, the need to facilitate asking and answering as a diaspora the question "what is to be done?" seems to be the task of the hour. The question remains: are we up to it? Are the traumas of the prior national liberation struggle too great? Are the divisions within our people too deep to discuss? At the very least, this paper's intentions and efforts speak to a possible advantage in our diasporic corner—as a diaspora we can, despite limitations, try and reckon with the diasporic neurosis and our situation. We are not totally alienated. At the end of Fanon's life, when the fault lines of ongoing national liberation struggles in Africa were beginning to widen, we can find guidance in his words on the assassination of Prime Minister Patrice Lumumba: "Amid the crocodile tears that were shed by the fifth columnists and imperialists, Africa had a choice in its first crisis: to advance or to fall back."[46] Either the African nations must intervene in the Congo or let this iteration of pan-Africanist and anti-imperialist efforts die, with the fate of an independent Africa resting resolutely in the Congo.[47] We all know what happened next with the Belgian backed rise of Mobutu and the creation of 'Zaire' in place of the Democratic Republic of Congo. Similarly, we as a diaspora have been in crisis since 2009, and we too are faced with a choice. Either we try to address our defeated diasporic neurosis and our situation to facilitate discussion around direction, or we let our people float adrift, yearning, and dissected by white settler-colonial society. Our being and future are waiting.

45. Fanon, *The Wretched of the Earth*, 166-167.
46. Frantz Fanon, "Lumumba's Death: Could We Do Otherwise?," in *Towards the African Revolution*, trans. Haakon Chevalier (New York: Grove Press, 1967), 192, 194, 196-197.
47. Ibid., 196-197.

Appendix

The following are a sample of the main interview questions that were asked towards the interviewees for this paper. The questions were not asked in a specific order but rather followed the flow and direction of conversation. The same applies to the naturally arising follow-up questions.

1) Tell us a little bit about yourself!
    a) Where and when were you born?
    b) When did you leave the island and why?
2) With your teen years in white Canada or elsewhere, what was your connection to the island?
    a) Did you feel like a diaspora?
        i) At what point did that first dawn on you?
        ii) How did you make sense of it?
        iii) How did it feel?
    b) Did you ever feel like you were missing something as you were trying to fit in white Canada?
    c) Did you expect to go back?
    d) Growing up, how did you react to adults speaking of the island as "back home"?
        i) Did that resonate with you?
3) Why did you go back?
    a) How did you feel?
    b) Growing up, did you ever expect to go back?
    c) Was it essential?
4) Going back, what worked and what didn't? What were the challenges?
5) How did you learn about the island?
    a) How did your family talk about the land and the national liberation struggle?
    b) How did the media and protests shape that image?
    c) How do you remember the national liberation struggle?

Bibliography

Ann. Freelance journalist based out of Canada. Interviewed by Vishwaamithran Ramakrishnan on November 26th, 2021.

Balasingham, Anton. *War and Peace*. London: Fairmax, 2004.

Chelvanayakam, S. J. V. *A Memorandum from the Tamils of Ceylon to Delegates Attending the 20th Commonwealth Conference in Sri Lanka*, transcribed by N. Ethirveerasingam. Tamilnation.org. https://tamilnation.org/diaspora/eelam/shan/chapter%2045.htm.

Fanon, Frantz. "This is the Voice of Algeria." In *A Dying Colonialism*, translated by Haakon Chevalier, 69-98. New York: Grove Press, 1965.

——. *The Wretched of the Earth*. Translated by Constance Farrington. London: MacGibbon & Kee Limited, 1965.

——. *Towards the African Revolution*. Translated by Haakon Chevalier. New York: Grove Press, 1967.

——. "The Fact of Blackness." In *Black Skin, White Masks*, translated by Charles Lam Markmann, 109-140. New York: Grove Press, 1967.

——. ஒடுக்கப்பட்டவர்கள்: விடுதலையின் வடிவங்கள். Translated by வி. நடராஜ். Coimbatore: Vidiyal Pathippagam, 2002.

——. "Gaullist Illusions." In *The Political Writings from Alienation and Freedom,* edited by Jean Khalfa and Robert J.C. Young and translated by Steve Corcoran, 75-80. London: Bloomsbury Academic, 2021.

Ismail, Qadri. "'Boys Will Be Boys'." *Economic and Political Weekly* 27, no. 31/32 (August 1992): 1677-1679. https://www.jstor.org/stable/4398721.

Mr. G. Accountant and brother of Ann. Interviewed by Vishwaamithran Ramakrishnan on November 26th, 2021.

Mr. X. Educationalist based out of Sri Lanka. Interviewed by Vishwaamithran Ramakrishnan on November 18th, 2021.

Sathi, Sumu. Anti-racism advocate. Interviewed by Vishwaamithran Ramakrishnan on November 22nd, 2021.

Vaitheespara, Ravi. "The Limits of Derivative Nationalism: Marxism, Postcolonial Theory, and the Question of Tamil Nationalism." *Rethinking Marxism* 24, no. 1 (January 2012): 87-105. DOI:10.1080/08935696.2012.635040.