



Iowa Research Online

The University of Iowa's Institutional Repository

Honors Theses at the University of Iowa

Spring 2021

Duties to a Future Self

Miranda Vermeer

Follow this and additional works at: https://ir.uiowa.edu/honors_theses

 Part of the [Applied Ethics Commons](#)

This honors thesis is available at Iowa Research Online: https://ir.uiowa.edu/honors_theses/372

DUTIES TO A FUTURE SELF

by

Miranda Vermeer

A thesis submitted in partial fulfillment of the requirements
for graduation with Honors in the Philosophy

Diane Jeske & Richard Fumerton
Thesis Mentor

Spring 2021

All requirements for graduation with Honors in the
Philosophy have been completed.

Carrie Figdor
Philosophy Honors Advisor

Duties to a Future Self

By Miranda Vermeer

I

Introduction

In this paper, I will evaluate the notion of having duties to one's future self. In doing so, I will discuss both ethical theories and metaethical theories. More specifically, I intend to utilize a prima facie duty-based ethical account and a psychological reductionist view of personal identity to argue that one *can* have duties to future selves. Relationships with our friends ground a number of objective agent-relative duties to them. I will argue that our relationships to our future selves are similar to our relationships to friends in significant ways and, as such, also ground objective agent-relative duties.

II

Duties

In order to evaluate the notion of having duties to a future self, it is important to first get a handle on what it means to have a duty. In the following section, I will detail different ethical theories and examine the role of the concept of a duty to each.

Consequentialism (Mill)

The concept of a duty runs counter to the consequentialist approach to ethics. A consequentialist theory will posit an account of what is intrinsically valuable, as well as what constitutes the right way of producing the greatest net sum of intrinsic value. Although consequentialist theories share the idea that what is right and wrong can be understood in terms of

the consequences of actions, there are different versions of consequentialism. There are different ways to formulate a consequentialist theory in that both what is held to be intrinsically valuable and the principle of right action can change from one version to another. For example, a hedonist holds that only pleasure is intrinsically valuable, so the act that maximizes pleasure in relation to any pain is the right action to take. Additionally, while one hedonist might endorse actual consequentialism, another could endorse the idea that the right action is that which the actor believes will produce the greatest net sum of intrinsic value.¹ Another still could endorse the expected utility view, which goes even further in that probabilities of different outcomes are used to choose the action which can be expected to maximize utility. There are more formulations, and different versions can end up with different moral judgments of actions.

To further evaluate consequentialism, consider a rugby match. Sports, like rugby, can bring people pleasure. They can also bring pain. In a match, scoring brings pleasure to a team's own players and their fans. Thus, by the hedonist view, scoring is instrumentally valuable—valuable as a means to achieving intrinsic value. To be the right action, such scoring must produce the greatest net pleasure. One side's pleasure from scoring must be considered in relation to the pain felt by the other side. Even if one team's scoring produces a negative net sum of value, like if the pain felt by the opposing side outweighs the pleasure felt by their own, it is still the right action if it produces a greater sum than any other actor could take. However, if the greatest net sum of pleasure could be produced by an actor doing that which would make the other side's pleasure outweigh their own, such as knocking the ball forward so that it changes possession, consequentialism would require the player to take that action.

¹ Actual consequentialism holds that an action's rightness, or wrongness, must be evaluated by the real consequences it has, regardless of what the actor thought would happen.

That is to say, a consequentialist theory is one of agent neutrality. It does not matter into whose life the value is brought; whether an actor's action brings value to themselves or to a stranger, the right action is still the one maximizing the net sum.² As will be evident elsewhere, many points in this paper could have, and do have, entire papers and books written about them. For example, some argue for forms of consequentialism that do not require agent neutrality. Examining such arguments is beyond the scope of this paper. Utilitarianism, as I will use the term, is a form of consequentialism which maintains that intrinsic value is only attached to the welfare of conscious beings, regardless of an actor's relationship to them.³ As a consequentialist view, that means it is not only *not* an egoistic theory, it also does not consider motive in evaluating the rightness or wrongness of an action.⁴

Consider a case of organ transplantation.⁵ Say there is a hospital in which five people are dying. Each has a different organ in failure than the others, meaning they each need a different kind of transplant to save them. Lying in another bed is a perfectly healthy person who is just there for a physical. This individual is of the proper blood type and their organs are all of adequate health to allow transplantation to each of the other five patients. Successful transplants would save the five patients but kill the donor.⁶ Given the utilitarian's commitment to maximization, they are

² Only non-relative consequentialism will be discussed in this paper.

³ This is a pluralist view, which can recognize knowledge, friendship, and other aspects of life as intrinsically valuable

⁴ For a more thorough detailing of utilitarianism, see Mill, John Stewart. *Utilitarianism*. 1863. Trans. Jonathan Bennet (Early Modern Texts, 2005), specifically Chapter 2: what utilitarianism is.

⁵ Philippa Foot posits a similar case in "The Problem of abortion and the doctrine of the double effect" *Oxford Review* 5 (1967) pp.5–15, while Gilbert Harman, *The Nature of Morality: An Introduction to Ethics* (New York: Oxford University Press, 1977) puts forth a nearly identical one.

⁶ The word 'donor' might be read as suggesting a level of consent/willingness to provide the organs for transplant. Assuming any unwillingness would not have any negative impact on the consequences of saving the organ receivers, it is inconsequential

obligated to say, “cut up the one to save the five,” because the welfare of the five outweighs that of the one—if there were no other considerations.

According to the utilitarian, the motives of the surgeon making the call to perform the operations do not matter for evaluating whether the action is right or wrong. However, motives can indicate whether the surgeon is a good or bad utilitarian actor and relatedly whether her actions are praiseworthy or blameworthy.⁷ A good actor is one intends to do the right thing—like the surgeon performing the transplants out of a motive to try to maximize the welfare of the people involved. This could be considered praiseworthy even if the surgeries failed and all six died because the intent to maximize intrinsic value was there.

The consequences of the actions determine morality for the utilitarian approach. This is not consistent with how society has formed to operate, such as in actual cases of organ transplantation. People on transplant lists wait either for a donation from someone who can continue living without what they have donated or from the recently deceased. Such sentiment reveals the intuitive idea that there are reasons outside of the value produced that factor into whether or not an action should be performed. While the suffering and death of the other five patients is evidently bad, there is a duty not to kill the other innocent patient.

Duties do not align with the consequentialist approach. If someone has made a promise to help a friend, but more net pleasure could be attained by a conscious being if that promiser stays on their couch eating chips instead, the promiser should do so according to some utilitarians. Even in the case of a principle of intrinsic value holding something like keeping promises as intrinsically valuable, the person whose promise is being kept does not matter. If choosing between an action that would fulfill one of their own promises and one that would mean two other people each

⁷ This terminology can be ambiguous, as the *actions* of praising and blaming would entail their own consequences and those would additionally need to be considered.

fulfilling a promise, the actor should choose the latter. This is out of line with the idea of agent-relativity, a notion that will be discussed in more detail later. Seeing as the concept of duty is incompatible with consequentialism, the deontological approach proves favorable.

Deontology

Like consequentialism, there are different deontological views, each of which has its strengths and weaknesses. First, seeing as the second formulation of Kant's categorical imperative is what initially drew me to write a thesis, I wish to examine Kant's moral theory.

Kant

As a deontologist, Immanuel Kant's moral view focuses on agent motivation and adherence to duties. Rather than focus on the results or repercussions of an act to determine its worth, Kantians examine whether such conduct aligns with the actor's duty. According to Kant, the conduct of an agent can only be morally right if they act in line with their duty.⁸ Possibly most identifiable with Kant's view are his formulations of the categorical imperative.

Unlike a hypothetical imperative, a categorical imperative determines what our duties are and applies to an individual irrespective of whether an action will attain outcomes they want.⁹ Kant provides three formulations of the categorical imperative, positing that each formulation is saying the same thing but in different ways. I will focus on the second one.¹⁰ The second formulation mandates that everyone "act in such a way as to treat humanity, whether in your own person or in

⁸ Immanuel Kant, *Groundwork of the Metaphysics of Morals*. 1785. Trans. Jonathan Bennett (Early Modern Texts, 2005)

⁹ A hypothetical imperative takes the form of: if you want X, then you ought to do Y. The imperative here depends on the disposition of the actor. No matter the consequences or desires of the actor, the categorical instructs what one ought to do.

¹⁰ The first formulation, sometimes called the Universal Law Formulation, mandates that one act as though the maxim of their action were to become, through their will, a universal law of nature. The third formulation of Kant's Categorical Imperative states that everyone act as if through their maxim they were a legislating member in the universal kingdom of ends ; Kant, *Groundwork* (1785)

that of anyone else, always as an end and never merely as a means.”¹¹ Although talking about only one of these formulations might suggest the acceptance of all three formulations saying the same thing, considering the equivalence of three formulations is beyond the scope of this paper. It is for the sake of time that I will only examine the second formulation.

Kant uses four cases to evaluate the categorical imperative, including the case of suicide. According to Kant, one cannot commit suicide because to do so would violate a perfect duty to oneself. People may consider suicide for a number of reasons, including immense physical pain or mental distress; an effort to end such suffering may be claimed as an act of self-love. Kant rejects this reasoning and instead holds that such an action could not be in accord with treating humanity as an end in itself because killing oneself to escape suffering would be using that self merely as means.¹²

The second case is of lying, or false promises. A situation perhaps too many people are familiar with is when a relative or friend constantly asks to borrow money, promises to pay back the funds, but never does. When this person makes such a promise knowing they will not have the ability to actually return the funds, they are making a false promise. In doing so, they are violating a perfect duty to others, according to Kant. A false promise violates the second formulation because when making a false promise one uses the person they make the promise to merely as means; this is because the person being promised is not aware of the true ends.

The third case, developing talents, seems especially applicable to one’s future self. According to Kant, a maxim like, ‘I’m going to sit and watch television for hours every night instead of practicing to develop my natural talent of solving complex equations,’ cannot be held. Ignoring one’s natural talents, according to Kant, hurts oneself and the rest of humanity because

¹¹ Ibid, 29

¹² Ibid.

available potential goes unmet. Rather than treating humanity as a mere means, not developing one's talents constitutes a maxim that does not harmonize with humanity.

The fourth and final case Kant presents is of helping others. Kant again argues that willing an uncharitable system, like one in which nobody gives any money to those less fortunate, would disadvantage the individual were they themselves to fall on hard times. As in the case of developing talents, Kant argues for the duty of helping those in need by appealing to the second formulation's criteria of harmonizing with humanity. This notion of harmonizing with humanity relates to the second formulation in that, although actions violating imperfect duties do not treat humanity as mere means, they fail to promote the ability to attain ends.

The first two cases are examples of perfect duties, while the latter two are of imperfect duties. As alluded to, what conflicts with perfect duties for Kant involves a maxim that treats humanity as a mere means. Imperfect duties do not harmonize with humanity as an end. Perfect duties require the individual to always be in accordance with them and are situationally strict; imperfect duties also require adherence but there is more leeway in how they are fulfilled.

The organ transplant case mentioned previously can be considered under Kant's theory. Killing the individual to save the five would be using the one as a mere means, and therefore be a violation of a perfect duty. Thus, according to the Kantian perspective, such an act would be wrong. This result improves on the consequentialist's, but Kant's approach is not without flaws.

An issue for any deontological theory would seem to be how to deal with conflicts between duties. According to Kant, one cannot permissibly act in conflict with one's duty, regardless of what the consequences of doing so might be. Whether someone has come to your door with the intent of murdering someone who only you know the location of, or you would need to lie about your ability to pay back a sum for food to feed your children, Kant considers lying and not keeping

promises as always wrong. According to him, there are not situational variables which could affect the morality of such actions. Like consequentialism, this sort of absolutist deontological theory proves problematic in its extremism, which suggests needing something in between the two. As such, a theory that allows for more situational evaluations while maintaining relativity is needed.

Ross

W.D. Ross posits such a view. As a deontologist, Ross's theory evaluates the rightness of an action by how it aligns with the duties of the agent performing it. Unlike Kant's deontological theory, Ross's provides for conflicting duties by positing *prima facie* duties—non-maleficence, gratitude, fidelity, justice, self-improvement, reparation, and beneficence.¹³ They are not held on the same initial level, as a duty of non-maleficence would generally be stronger than that of beneficence, but they can conflict and situationally outweigh each other.

For example, one may have a *prima facie* duty to help their friend move because they promised to do so. Imagine the person is on their way to help their friend move when they encounter a horrible car crash. They still have that *prima facie* duty to keep their promise, but now it seems they also have a *prima facie* duty to help the victims of the car accident.¹⁴ According to Ross, the person has a *prima facie* duty to keep their promise as well as a *prima facie* duty to help, and, because they cannot do both, they should do whichever they have a stronger moral reason to do. Their all things considered duty—what they ought to do considering the circumstances and the relevant *prima facie* duties—in this case would be helping the injured people. Having an all things considered duty does not mean the other *prima facie* duties are eliminated. Breaking the promise is still *prima facie* wrong and can entail uncomfortable feelings as well as a duty to make up for having broken it, but that does not mean the act was actually wrong.

¹³ W. D. Ross, *What Makes Right Acts Right* (Oxford: Oxford University Press, 1930) 18

¹⁴ *Ibid*, 28

Looking again to the transplant case, Ross's theory could hold the same as Kant's while the weight of the prima facie duties of non-maleficence outweighs those of the other prima facie duties, including beneficence. The surgeon's all-things-considered duty would be to not remove the organs from the healthy patient as long as the amount of good produced by such an action would not outweigh the stricter duty of non-maleficence. However, the good produced as a consequence of actions, though not determinative in Ross's theory, is still a factor for consideration where it is not in Kant's. An adjustment to the case, such as that removing all of the healthy patient's organs would save the lives of millions rather than saving only five, would entail Ross and Kant's theories diverging in their answers. Kant's absolutist views would still not allow for the healthy patient to be sacrificed, while the prima facie duty of beneficence would become the most heavily weighted in Ross's view and thus entail an all-things-considered duty to sacrifice the patient.

One can know of having a prima facie duty to keep a promise simply by reflecting on the notion of keeping a promise rather than needing to infer, and this extends to all prima facie duties. Keeping a promise is justified by its connection to a corresponding prima facie duty and prima facie duties are self-evident. He compares claims of them to those made about geometry. Just as we do not seek to justify the trust in the reasoning for the claim that a straight line can be drawn from any one point to another or that a circle can be constructed with any point as its center and any radius, Ross argues we should not have to justify our trust in this moral reasoning.¹⁵

Kant's case of the lying promise can provide further insight into the benefits of Ross's view. Say Dolores is a single mother and has been unable to hold down a job. She and her three kids are starving, but she knows her landlord has money to spare. Dolores knows that she would

¹⁵ These are axioms of Euclidean geometry

not be able to pay back any borrowed money, but she needs to feed her kids and so asks for a loan promising to return the money. Again, according to Kant, this is never permissible. According to Ross, the prima facie duty to keep her promise still exists. However, so do other duties, like those to her children. Dolores's relationship with her kids, and her responsibilities to them, matter. According to Ross, relationships like those between spouses, friends, child and parent, creditor and debtor, promisee and promiser, and fellow countrymen found prima facie duties.¹⁶ Thus, Ross writes of duties being agent-relative, referencing the "highly personal character of duty."¹⁷

Agent-Relative Duties

A duty can be agent-neutral or agent-relative. A duty is agent-neutral if it exists for anyone, regardless of anything specific about them; a duty is agent-relative if it does not have to exist for everyone and instead exists for a person because of their specific role or relation. Consider a situation in which a child is drowning in the ocean, not far from the shore. In terms of a general duty of beneficence, anyone on the beach who is capable of safely swimming out to rescue the child would be said to have an agent-neutral duty to do so. For identifying agent-relative duties, one must look at the specifics of the people there. One might think of the lifeguard on duty and the guardian of the child as each being in such a role in the situation to give them an agent-relative duty to help. After all, the lifeguard's job is to help save any swimmer and the guardian is in a position of responsibility for the child's welfare.

In addition to distinguishing agent-neutral from agent-relative, agent-relative can be further broken down into the subjective agent-relative and objective agent-relative.¹⁸ An agent-relative

¹⁶ Ibid, p19

¹⁷ Ibid, p22

¹⁸ Diane Jeske makes this distinction between subjective and objective agent-relative reasons in "Friendship and Reasons of Intimacy," *Philosophy and Phenomenological Research* 63, no. 2, (2001) pp. 329-346 to argue for objective agent-relative reasons to care for intimates.

duty can be described as objective when the actor's personal sentiment toward the action and persons involved are not relevant—they do not ground the duty. That which is subjective agent-relative is not so coherently described as a duty. Instead, talking in terms of subjective agent-relative *reasons* makes more sense. Having a subjective agent-relative reason is grounded by valuing the object. Calling back to the example of the child drowning, the guardian's relationship as such to the child grounds an objective agent-relative duty to save them—regardless of how they feel about the child, they are responsible for its wellbeing. The guardian can also have subjective agent-relative reasons to help the child, such as if they care about the child and want them to do well. In Ross's view, prima facie duties, such as that arising from making a promise, are grounds for objective agent-relative duties.

Some theories concerning agent-relative duties require more involved talk about the self/personal identity and so I will come back to them after briefly discussing some metaphysical aspects of the notion of self. From here on I will endorse Ross's ethical theory regarding duties because it allows for the consideration of consequences, unlike Kant's view, without them being the only consideration, unlike a consequentialist view. As such, it also allows for agent-relative—both subjective and objective—considerations.

II

Metaphysics – Self

Defining the self will be important in establishing how a person at one time is related to themselves at a different time. Many efforts have been made at giving an account of what constitutes the self, but not all of them would be consistent with the notion of having a duty to a *future* self.

Descartes

In *Meditations*, René Descartes sets out to reconsider everything he thinks he knows. He wants to establish absolutely secure foundational beliefs on which to build so as to not include any false ones. As these considerations progress, he famously posits, “cogito, ergo sum,” meaning “I think, therefore I am.”¹⁹ From this, it is derived that the essential property of a mental substance (soul or self) is the generic cognitive power of thought, and not any physical instantiation or characteristic. A substance dualist, Descartes holds that the mind is one kind of substance, and the body is another. The body, rather than being defined by the power of thought, is defined by extension. In discussing physical substance, Descartes famously uses the wax example in which he describes the properties of a piece of wax before and after being melted. Descartes talks about changing wax to suggest the only essential property is extension. The thinking mind can persist without such a body, but the two are also joined in some way, by God, so as to allow for sensory experiences, like the pain of stubbing one’s toe.

Important to the discussion of the self is getting a better idea of what this mental substance is. It is attributed only the power to think. Any particular psychological attributes of an individual are not counted as essential by this view. When remembering John Belushi, words like funny and manic come to mind in describing his personality. However, on Descartes’s approach, these are not essential features of the mental substance; it is the fundamental ability to think that matters. When John Belushi died, the mental substance would have persisted, according to Descartes, but would no longer have been attached to time or space via a body.

Making sense of how these two substances could interact and how individual substances could be individuated are two issues for the Cartesian view. If someone were to be reduced only

¹⁹ *René Descartes, *Meditations on First Philosophy* (1641) p *

to their ability to think, irretrievably losing all of their other characteristics and psychological traits, including their memories, it seems that most of us would want to say that they were no longer the same person as they were before this happened. Looking now to Locke, his view is one that supports such intuitions.

Locke

John Locke does not believe people can be identified with the mental substance posited by Descartes; the mental substance cannot determine persistence through time. Instead, Locke's view is a reductionist theory especially focused on memory. He discusses identity most directly in the chapter "Of Identity and Diversity".²⁰ Locke puts forth the case of the prince and the cobbler, and this example will be helpful for illuminating Locke's own theory as well as for comparing Locke's theory to Descartes'.

In the prince and cobbler example, the soul of a prince, accompanied by his consciousness—an ability, inseparable from thinking, that allows an individual to be aware of and consider their reasons and thoughts over time and space—overtakes the body of a cobbler whose own soul has gone. The underlying question here is, who is this person with the 'soul' of the prince but the body of the cobbler? The person has the same memories as the prince, and only the prince. Locke claims that this person is the prince. In doing so, he makes the argument that the individual would be the same *person* as the prince, but be the same *man* as the cobbler, outwardly. Thus, while the identity of a man requires reference to a body, the identity of a person does not and instead involve a continuation of consciousness. This idea of a consciousness that persists through time represents an important concept to Locke's theory—the diachronic self. Descartes, in arguing for his dualist view, focuses on synchronous identity.

²⁰ John Locke, *Essay Concerning Human Understanding*

The prince and cobbler case presents several issues for Descartes' view. First, it is hard to even imagine what a transfer of mental substance would entail. Connecting two completely different substances seems impossible and the appeal to divine intervention leaves a lot to be desired already, and the idea of severing and then reconnecting a soul to a different body just deepens these concerns. However, maybe a transfer of substance is not Descartes' only possible reply. Rather, it does not seem inconsistent with Descartes' view to answer that there was no switch of mental substance. This of course requires a slight rewrite of the situation, which has already described the soul of the prince shifting to the cobbler, whose soul has left. Instead, imagining a case in which one individual has lost all of their psychological characteristics and completely assumed the characteristics of another, Descartes might try to respond that the characteristics of the original mental substance of the cobbler have changed but that it is the same substance. This, too, is unsatisfactory.

Regarding the person as assuming the identity of the prince, Locke's prince-cobbler example illustrates his rejection of the bodily criterion of personal identity. Rather, Locke endorses the psychological criterion of personal identity. According to Locke's version of this view, it is the continuity of a person's consciousness that matters for personal identity, not an irreducible mental substance. To discuss a person's identity, by this view, the subject must have some sort of persistence through time. As this holds for Parfit as well, I will discuss the idea of a persisting self and the related concept of four-dimensionalism shortly.

Parfit

Opposed to the dualism of Descartes, Derek Parfit argues for a reductionist view of personal identity.²¹ He describes a reductionist view of personal identity as one that claims that the fact of a person's identity over time just involves "the holding of certain more particular facts...these facts can be described in an *impersonal* way".²² A version of this view may only consider psychological aspects, such as "thoughts, desires, intentions, and other mental states".²³ Parfit endorses this version of reductionism—psychological reductionism—and argues that the continuity and connectedness of a person's psychology is what ultimately matters in discussing the self. Parfit defines both psychological connectedness and psychological continuity. Person A is psychologically connected to Person B if A holds direct psychological connections to B, and psychologically continuous with Person B if A holds overlapping chains of direct relations.²⁴ Psychological connectedness and/or psychological continuity is referred to by Parfit as Relation R.²⁵ His version of the reductionist view is less narrow than Locke's as it encompasses more than just memory in considering psychological continuity; Parfit holds a broader view in which other characteristics, like personality traits, intentions, and beliefs, can constitute psychological continuity.

Importantly, seeing as the word 'identity' has already been mentioned several times, Parfit actually argues that our psychological continuity/connectedness is what is important rather than one's identity. Though there are difficulties for Parfit's view as well, considering them is beyond

²¹ This discussion of Parfit will be based on two of his works, "Personal Identity," *The Philosophical Review* 80 (1971) pp. 3-27 and *Reasons and Persons* (Oxford: Oxford University Press, 1984) and there is significant overlap. His essay was published earlier, and its contents can be seen throughout the later book, especially in chapter 12, titled "Why Our Identity is not What Matters".

²² Derek Parfit, *Reasons and Persons*, 210 ; his emphasis.

²³ *Ibid*

²⁴ *Ibid*, 206

²⁵ *Ibid*, 262

the scope of this project. In explanation and defense of his view, Parfit offers thought experiments of fission and fusion.²⁶

In the division cases, Parfit describes brain transplants. He proposes that most would agree that if a person's brain, I will call this the brain of Person A, were transplanted and correctly wired into a brainless body so that the new body had all of the psychological traits (including the memories) of Person A, the 'new' individual would be Person A. Parfit then proposes a case where each hemisphere of Person A's brain is successfully transplanted into a separate body. Each new individual possesses the same psychological traits as Person A, and so of the other individual—at least initially. He suggests three possibilities for Person A: a) they do not survive, b) they survive as one or the other of the two, and c) they survive as both individuals.²⁷ He argues that (a) seems too similar to the first transplant case to claim one as a success and the other as a failure, and for (b), distinguishing one of the surgeries from the other when they are the same psychologically is incoherent; the decision would be arbitrary. The last option, (c), precludes survival based on identity because the two resulting people cannot both share an identity as identity is transitive. Instead, Parfit argues that judging the case in terms of continuity and connectedness, without any corresponding talk of identity, allows for the answer that Person A does not survive as either person has the same connection to both as she would to her future self had she survived. Asking who Person A is is an empty question. Psychological continuity does not require a one-one relation.²⁸

In a case of fusion, the brains of two people are connected. Consider a situation in which the bodies of each individual in a couple are gravely injured in a car accident, but portions of their brains can be salvaged and then fused together. The resulting brain, a combination of one half of

²⁶ Derek Parfit, "Personal Identity," 4-5

²⁷ Parfit, *Reasons and Persons* (1984), 256 ; Derek Parfit, "Personal Identity," *The Philosophical Review* 80 (1971) pp. 3-27

²⁸ Parfit, "Personal Identity", 10

each person's individual brain, is stuck into a new body. The result of this is imagined to be distinct from that of the fission case. In the fission case, the person's characteristics are replicated and unchanged, while in the fusion case the two people's characteristics are brought together to form something new. For example, my roommate and I both love cats, so brought together this characteristic would be compatible and persist if our brains were fused. However, my roommate also loves eating microwaveable breakfast sandwiches and I hate them, so, brought together, by Parfit's view, these characteristics would cancel each other out and the resulting feeling toward those deplorable sandwiches from our fused brain would be something like neutrality.

Given these changes in psychological characteristics, Parfit suggests that anyone endorsing the idea that such a transformation would really be death is less unreasonable than the person endorsing that view for the fission case. However, he still disagrees. He first replies that the altering of some of our characteristics hardly seems to constitute our death—in fact, many people actively want their characteristics to change. For example, people seek out hypnotism in an effort to change the unhealthy characteristic of loving smoking cigarettes. Additionally, Parfit suggests similarities between the fusion case and getting married, stating that there can be a choice in the partner for fusion and that each has about as much of what he calls “intentional control” over what the person with the fused brain does as the control each part of a married couple has over their joint actions.²⁹ Thus, Parfit takes the fusion case to show that psychological continuity allows for degrees, whereas survival defined by identity does not.

The ideas of psychological continuity and connectedness are key to Parfit's argument, and he endorses the idea of quasi-memories to make them plausible in cases of branching. He puts forth a definition of quasi memories, or q-memories: one is q-remembering an experience if (1)

²⁹ Ibid, 19.

they have a belief about a past experience that is apparently a memory, (2) that experience did occur for someone, and (3) in the right kind of way, the memory is causally dependent on that real past experience.³⁰ Parfit argues that such memory does not require identity, because even though our q-memories are typically of our own experiences, there is nothing that requires them to be. He takes these quasi scenarios to hold for other direct psychological connections as well, including intentions and character traits.

This notion of psychological continuity is related to the concept of people being composed of temporal parts. Similar to Locke, it is important to Parfit's view that objects can exist through time, as the reductionist view concerns personal identity over time. Four-dimensionalism holds that, like the dimensions of space, time is a dimension. According to this view, things have temporal parts the same way that they have spatial ones: a spatial part of my car is its tire while a temporal part could be it being in my garage yesterday. This applies to humans as well, meaning persons exist through time as temporal parts—sometimes called person stages.³¹ Although diving too deep into metaphysical concepts of time would be tangential to my argument in this paper, it is important to mention that some conceptions of time would seem to be inconsistent with the notion of person stages.

Presentism holds that only the present moment exists, and that what is the present moment changes.³² The idea of there being any temporal parts at different times, such as a future self, is incompatible with such a view. Eternalism, holding that past, present, and future all exist and that

³⁰ Ibid, 15 ; references Shoemaker's "quasi-memory" and Penelhum's "retroognition"; there are those who will argue that quasi-memories are not possible and so cannot assure that psychological continuity does not presuppose personal identity, but evaluating such arguments is beyond the scope of this paper.

³¹ Thomas Sider writes in detail about four-dimensionalism, *Four-Dimensionalism: An Ontology of Persistence and Time* (Clarendon Press, 2013)

³² Kristie Miller, "Presentism, Eternalism, and the Growing Block" *A Companion to the Philosophy of Time* (2013)

whatever moment is ‘present’ does not change, is more compatible with four-dimensionalism.³³ The concept of person stages is important in my forthcoming discussion of duties to future selves. Without such temporal parts, the notion of having a duty to a future self would be incoherent—the self now would be indistinguishable from any other self later or prior, like in Descartes’ account.

Parfit does not argue for the same conclusion that I do. Instead, Parfit takes his psychological reductionism to a utilitarian end of universalizability. I wish to discuss this more in a later section of this paper, but for now, I want to go forward with a psychological reductionist view of personal identity. Unlike Parfit, I will argue that there can still be a prioritization of duties to certain others, including to future selves.³⁴

III

Duties to Future Selves

With a Rossian view of duty and a psychological reductionist account of the self, I want to argue that a person can have duties to a future self. There exist not just subjective agent-relative reasons to act for a future self, but also objective agent-relative prima facie duties to do so. These, similar to circumstances in other relationships, are grounded in one’s relation to their future self.

Other Relationships

In order to get an idea of a person’s relationship to their future self, I want to discuss other relationships and later compare them with the relationship to a future self. Other situations where it seems that we, at least intuitively, have obligations grounded in the special relationships we have to other people include when these people are family and friends. To start, there is the relationship

³³ Ibid ; Diving more thoroughly into issues of time/ontology has been undertaken in other works, including Miller’s, and falls outside of the scope of this paper.

between a parent and their child. Unlike other animals, like snakes, who have no parental instincts and provide no parental care, humans as a species generally provide significant parental care for many years. Parents, as seen across nations in societal norms and law, are expected to provide for their child's basic needs, including their safety, nutrition, and shelter.

There are clear biological links involved in the relationship between some parents and children—adoptive and stepparents do not have such links. Some might want to argue that this biological connection is relevant to the future self as well, but this is unnecessary and potentially problematic. First, issues with the biological approach have already been addressed above in discussing Parfit. Additionally, adoptive and stepparents are seen as having the same duties as biological parents where the biological parents of an adopted child have none. It is the relationship, rather than any biology, that seems to ground obligations—assuming the role of parents means assuming certain duties.

Additionally, a parent's duties to their child exist regardless of their feelings toward them. Although I am not a parent myself, it is my understanding that it is common for parents to not always like their children. Whether it is because of the newfound, rebellious attitude of a teenager, the tantrums of a toddler, or seemingly indefinitely many other circumstances, one may find themselves disliking their kid. Regardless, a parent is not justified in failing to fulfill their responsibilities to their child anytime they dislike them. While many parents do have subjective agent-relative reasons to help their child, I will argue that there are also objective agent-relative duties parents have to their children.

Relatedly, there are other familial relationships as well that, although different from the role of a parent, seem to impose duties on those in them. There are quite a few types of familial relationships, falling into the categories of either intergenerational or intragenerational—by this I

mean those between, say, an aunt and a niece and between siblings, respectively. Examining each of these different relationships to evaluate nuances in duties would be interesting but is ultimately beyond the scope of my current paper.³⁵ For the sake of brevity, sibling relationships are familial ones that do not seem to entail all of the same duties as a parental role, but still seem to ground prima facie duties for each sibling. Having grown up with a twin sister, I am familiar with this situation. There were countless times growing up where we did things for each other that we would not have done for someone who was not our sibling—she sat through way more soccer games and I sat through way more dance recitals than either of us would have liked.

Then there are one's relationships with friends. The term 'friend' is used in many different ways—from describing the people who know and care the most about someone, to casual acquaintances, and even to those who are nearly strangers. In talking about relationships with friends, I am referring to the friends with whom someone has a close personal connection that includes shared experiences, interests, and values.³⁶ When a friend asks you for a favor, even when it is something you would not regularly agree to do, you feel obliged to help them. There of course generally exist subjective agent-relative reasons behind helping a friend with, say, moving out of their apartment—you care about them and want them to be helped. However, these positions appear to ground duties as well.

We seem to have duties to our family and friends that generally take precedent over duties to others—they provide an additional prima facie duty in weighing one's all things considered duty. Ross gets at this point, as I have mentioned previously, in his critique of utilitarianism. He

³⁵ A number of significant relationships will have to go unexamined in this paper for the sake of time, including Ross's own examples of spouses/partners.

³⁶ I borrow this description from Jennifer Whiting, "Friends and Future Selves," *The Philosophical Review* 95 (1986) 558 ; Whiting argues that concern for our future selves is a component of psychological continuity, and she argues the reverse of what I am arguing, taking our reasons to care about ourselves to show that we have reasons to care for our friends.

argues that utilitarianism fails to consider how personal relationships factor into determining right action. Positing that personal relationships are the foundation of a prima facie duty, his view is an agent-relative one. We do generally care about friends and family members and from this concern have reasons to do certain things for them—we have subjective agent-relative reasons to act—but I also want to argue that we have objective agent-relative duties to them.

To a Future Self

A relationship to a future self is not exactly the same as a parent's relationship to their child. A parent's prima facie duties to their child are often the most demanding in any relationship, and such a dynamic does not seem to hold for the relationship between a present and future self. Although there are similarities in that a present self may be required to sacrifice, to a degree, for a future self, many people would agree that a parent could morally be required to sacrifice anything if their child's need was great enough, and that is not consistent with the relationship between a present and future self.

Other familial relationships, and those with friends, often create less demanding duties than duties grounded in a parental role. Familial relationships are more automatic than those with friends, and this is seen in the relationship with the future self as well—a person does not have control over standing in a causal relationship with their future selves.³⁷ One has more choice in who their friends are, as we are able to choose, to an extent, if we engage with certain people enough to form close friendships with them. On the topic of engagement, it looks different in the case of a relationship to a future self than it does with the other interpersonal relationship examples.

³⁷ I recognize that individuals may choose how much they engage with their family, but my point here is they cannot choose who all stands in familial relations—i.e. father, mother, brother, sister, aunt, uncle, cousin, grandmother, grandfather—to them.

A person cannot have a conversation with their future self because a future self cannot directly reply. Instead, the interaction is purely the present self's causal relation to the future self.³⁸

We have shared desires, values, and ends from shared experiences and interactions with our friends, and with our future selves from a somewhat different causal relationship, as I have described above. Relation R—continuity/connectedness—as detailed by Parfit, exists in degrees, and, as such, an individual shares more continuity with some people than with others, including oneself. Holding only psychological continuity, as defined by Parfit, to ground obligations could significantly limit one's obligations to a future self because the psychological continuity with a far enough removed future self could be so little so as to ground only weak obligations, like those had to near strangers. A suggestion here could be that there might be as much range in relationships to a future self as there is in one's relationships with friends and family. Just as you have a closer relationship with one friend than with another, or one sibling than with another, so too it seems that there would be differing degrees of closeness between temporal parts of oneself. This suggestion would seem to entail that my relationship to myself tomorrow grounds stronger prima facie duties than my relationship to myself in 10 years. However, I think this idea misrepresents the relationship in which we stand to our friends and future selves. The causal relationship in which we stand to our future selves grounds objective agent-relative duties, which exist regardless of our connection to the ends themselves.³⁹

I will borrow Kant's examples to illustrate cases in which one may have such duties to a future self. Kant's first case is of suicide. He argues that suicide is always in violation of the

³⁸ The importance of such causal relations to objective agent-relative duties is discussed in Jeske, 2001.

³⁹ Jennifer Whiting, "Friends and Future Selves," (1986) 558 gives this description (common experiences, desires, interests, and values and interactions causally affect these desires, interests, and values), based on Parfit's view ; David Brink, "Rational Egoism, Self, and Others," In O'Flanagan and A Rorty (eds), *Identity, Character, and Morality* (Cambridge: MIT Press, 1990) argues that our relationships to our future selves and others are grounded in causal interactions and subsequent changes of beliefs and desires, and from this endorses a rational egoist perspective.

categorical imperative and as such is always morally wrong. Considering suicide by Ross's prima facie duties, there are conflicts. Take for example someone who is terminally ill with an excruciatingly painful disease. That individual's relation to their future self is the foundation of a prima facie duty. There is added weight, then, to any prima facie duties stemming from this relation. Considering that all future selves would also suffer from extreme pain, not killing oneself would mean a future self suffering. As such, the situation involves the prima facie duty of non-maleficence, which requires one not to hurt others.⁴⁰

However, there are other relationships that are prima facie duties, including relationships with friends and family. The duty of non-maleficence applies to them as well because losing their friend/relative would hurt them. Killing oneself also eliminates the ability to fulfill any other prima facie duties, such as those of fidelity and gratitude. A Rossian approach does not provide an ultimate formulation for how to weigh prima facie duties against each other. It relies on the idea that one's all things considered duty in a situation will be knowable by intuition. Still, the prima facie duty of non-maleficence to a future self at least factors into the balancing of prima facie duties for determining an all things considered duty.

The case might be made simpler by imagining an individual with such a chronic condition who has no family or friends. There is a duty of self-improvement that may still weigh against an all things considered duty to suicide—killing themselves would eliminate the possibility of subsequent selves and any resulting improvements.⁴¹ However, one may respond in this case that severe enough chronic pain and a soon approaching death could have eliminated the possibility for

⁴⁰ I find the duties of non-maleficence and beneficence difficult to differentiate, including in this situation. To avoid hurting someone seems to entail an amount of happiness/benefit to them and vice versa.

⁴¹ The phrase 'to commit suicide' is sometimes deemed judgmental or normative because 'commit' is used in connection with crime. While using suicide alone as a verb may read as a mistake at first, it is an intentional writing choice to avoid bringing in any such connotations.

improvement already. While killing themselves would violate a prima facie duty of non-maleficence to their present self, I believe it is possible in this case to make an argument that the all things considered duty is to suicide because of a heavier balance of duties of beneficence/non-maleficence to the self—both future and present as the pain is ongoing potentially without any redeeming quality of life.

Kant's second example is of a lying promise. I will focus simply on the idea of making a promise to oneself, as the notion of truly lying to oneself seems nonsensical.⁴² "I promise myself _____" is a common enough phrase. It is often connected with some sort of idea of self-improvement, like "I promise myself to eat healthier." The idea of having a prima facie duty of fidelity—to keep promises—to a future self might face some criticism. For example, one may question to whom the promise is really being made. Promises are typically made between two or more existing individuals, such as someone promising their parents that they will be home for dinner. As a future self is, by definition, not present, it may seem like the promise is actually being made to the present self. One might additionally argue we do not make promises that bind others—if I promise my parents that my sister will be home for dinner, a prima facie duty of fidelity is not placed on my sister. However, this seems to show a difference between intrapersonal and interpersonal cases. When I promise my parents that I will be home for dinner, I am binding a future self. We would not find it acceptable if someone stood up in court and argued that a past self committed the crime and, because they are a different temporal part than that person, their current self should not be the one to be punished. As such, it seems that a past temporal part could place prima facie duties of fidelity on subsequent selves to even more removed selves which could ground all things considered duties.

⁴² An interpersonal promise may generally be thought to be made between two or more present people

Kant's third example, as previously discussed, is of developing talents. If a person has a natural disposition to play the piano well, they have a duty to develop that talent, according to Kant. Though not as absolute, this is seen in Ross's view as well in the prima facie duty of self-improvement. A duty of self-improvement would be to the benefit of a later person stage, as the work the current self puts into practicing the piano will only come to fruition at a later time. A person at T_1 practices so that their self at T_2 has even further developed skills. In terms of a duty to a future self, the prima facie duty involved may be that of beneficence, as future selves can benefit from the developed skills.

Another consideration is one's prima facie duty to their future self's health. Eating well and exercising now is shown to make you healthier later in life, discounting unrelated conditions like genetic pre-dispositions to illness. Smoking, on the other hand, is proven to have extremely negative effects on one's health, including causing lung cancer and death. A person's choices regarding health then have the potential to hurt or bring happiness to a future self. If a person smokes a pack of cigarettes each day, their future self is likely to suffer as a direct result. Given the special relationship with a future self, one has especially weighted prima facie duties to them, and an all things considered duty of non-maleficence to a future self could outweigh prima facie duties of beneficence to a current self.

To a Past Self

In addition to having prima facie duties to future selves, I want to briefly discuss the notion of having prima facie duties to a past self—or a future self having prima duties to a present or past self. In considering the different prima facie duties, the prima facie duty of gratitude comes to mind as one to a past self. This would apply if a former self sacrificed or otherwise did a favor for which a later self would be grateful. The case of a past self quitting smoking might constitute

grounds for a later self having a prima facie duty of gratitude, as might a past self's self-improvement efforts. The later self would be unable to say thank you to the past self, as might be required of a duty of gratitude in another circumstance. However, perhaps this duty could be fulfilled in another sort of expression of gratitude, such as honoring a desire of that past self.

IV

Objections

While I have attempted to put together a plausible theory, I realize there could be plenty of room for critique. Ross and Parfit's theories each attract their own objections, and because I have endorsed portions of each of their views, my own is not immune to all of those critiques, either. In the following section, I will consider a few objections to the theory I have formed regarding duties to future selves.

One objection to consider relates the last section to the first, and that is to argue that psychological reductionism actually supports utilitarianism and accepting it while promoting agent-relativity is inconsistent. Parfit, for example, takes his own theory to support utilitarianism. Person stages are distinct and there can be the same, if not more, psychological continuity between persons than between person stages of an individual, according to Parfit. He argues that identity is not what matters and that continuity is what is important. Arguing that you can be as connected to a different temporal part of yourself as to another person, Parfit claims that the distinction between 'you' and anyone else does not matter and thus arrives at a utilitarian universalizability. I do not agree with this jump.⁴³ From my understanding, as I have put it in this paper, Parfit argues that

⁴³ Both Diane Jeske, in "Persons, Compensation, and Utilitarianism," *The Philosophical Review* 102 (1993) pp. 541-575 and Jennifer Whiting, in "Friends and Future Selves," (1986) pp. 547-580 argue, among others, against psychological continuity leading to utilitarianism.

Relation R comes in degrees. How can such a utilitarian end be required when the relation is stronger or weaker based on those involved?⁴⁴ Instead, it seems that agent-relativity follows from psychological reductionism, but that the ‘self’ is less distinguished from other close relationships—such as those to family and friends.

Another objection could be that duties to future selves are really just prudential concerns. That is to say that rather than there being prima facie duties to future selves—which is a moral notion—a person with this objection would be arguing that people want themselves to do well and be happy in the future and act only from this perspective rather than out of any obligation. This objection seems to come down to the distinction between objective and subjective agent-relative reasons.⁴⁵ While I agree that there are subjective agent-relative reasons at play because people generally care about and want the best for themselves, claiming them as the only factor fails to grasp the relationship one has to one’s future self. The relationship one has to one’s future self grounds objective agent-relative prima facie duties, similar to how such prima facie duties exist to family and friends.

The final objection I will consider is that any duty to oneself, including to a future self, is waivable and so constitutes no significant obligation.⁴⁶ In order to properly consider this objection, it is important to understand what is being said. By suggesting that a duty is waivable, it means that the actor who has a duty can be released from it by the object of that duty. For example, if I promised an old friend that I would go to dinner with them, I may no longer have a duty to follow

⁴⁴ This is again based on the notion that utilitarianism can only be agent neutral.

⁴⁵ For more discussion of duties to the self and prudence, see Alison Hills, “Duties and Duties to the Self,” *American Philosophical Quarterly* 40 (2003) pp. 131-142.

⁴⁶ Alison Hills, “Duties and Duties to the Self,” (2003) also discusses the concern of waivability in regard to duties to the self.

through on that promise if they waive it, telling me that it is alright or that they would not be able to make it to dinner anyway. This objection brings up a number of interesting things to consider.

A factor to consider in this objection is how duties can be waived.⁴⁷ For example, imagine a situation in which a person's friend is dying in a hospital. If the non-hospitalized person hates hospitals, can the dying friend waive their duty to visit them? It seems there would be a critical difference between the dying person waiving their friend's obligations out of their own feeling of wanting to be alone versus only acting to spare their friend having to do something they disliked, despite wanting company. In the former situation, it seems that a *prima facie* duty of beneficence to the dying friend was waived but remained in the latter. Additionally, some duties have been argued as un-waivable, such as those involved in the right not to be tortured.

A case between different temporal parts provides further complexities than an interpersonal one. These complexities seem to stem in part from the level of interaction one can have with a temporally distinct self. To have a duty to a past self is similar to having a duty to a deceased or otherwise inaccessible person—they can no longer directly be affected by the duty being fulfilled. However, we still generally seem to hold that we should meet duties to those who are gone. For example, if you promised a friend you would take care of their dog after they died, your duty to care for the dog remains after the friend has passed away. Similarly, who would be waiving a duty to a future self? A future self does not exist at the same time as a present one, and so it seems they would be unable to excuse a present self from an obligation.

I recognize that this might not be a very satisfying response to this objection. Additionally, I recognize that other objections could be raised to my view—I have elected to only address a few of them here.

⁴⁷ Going further into the idea of duties being waived could lead to examining the notion of unwaivable ones. Detailing these and corresponding theories of duties/rights is beyond the scope of this paper.

V

Conclusion

In the previous sections, I hope to have shown that having objective agent-relative duties to future selves is plausible. The theory I have detailed is an agent-relative deontological one that endorses a Rossian account of duties wherein multiple prima facie duties factor into determining one's all things considered duty. I have also endorsed a psychological reductionist view of personal identity. However, having accepted this view, and still believing us to have duties to our future selves, I must ground such obligations in more than continuity. I have argued that the special relationships a person has to their friends and family are similar to their relationship with their future self, and that these relationships ground objective agent-relative duties. There are more complexities to be explored than a paper of this length could allow, but I think the subject warrants further consideration.⁴⁸

⁴⁸ It would not have been possible for me to complete this thesis without many people. I have been lucky to receive incredibly informative, supportive, and entertaining mentorship from Professor Diane Jeske and Professor Richard Fumerton throughout this project, and I am beyond thankful to them. To my parents, sister, family, and friends, thank you for always encouraging and helping me. To Molly, thank you for inspiring me, and for your love, support, and humor. To Drake, you have made this world a better place and I can only hope to bring a fraction of the kindness, intelligence, and strength to it that you have. I would not have made it through a pandemic-ridden senior year without Christine, Sam, Veronica, and Fish. Thank you to my Aunt Laura for preceding me in studying philosophy and for no doubt inspiring my interest. In addition to Professor Jeske and Fumerton, I would also like to thank Professor Katarina Perovic, Professor Jovana Davidovic, Professor David Cunning, Professor Carrie Swanson, Professor Gregory Landini, and Professor David Stern for making my time studying philosophy at the University of Iowa such engaging and fulfilling one.

Works Cited:

Brink, David O. "Rational Egoism, Self, and Others," In O'Flanagan and A Rorty (eds), *Identity, Character, and Morality* (Cambridge: MIT Press, 1990)

Foot, Philippa. "The Problem of abortion and the doctrine of the double effect" *Oxford Review* 5 (1967) pp.5–15

Harman, Gilbert. *The Nature of Morality: An Introduction to Ethics* (New York: Oxford University Press, 1977)

Jeske, Diane. "Persons, Compensation, and Utilitarianism," *The Philosophical Review* 102 (1993) pp. 541-575

---Jeske, Diane. "Friendship and Reasons of Intimacy," *Philosophy and Phenomenological Research* 63, no. 2, (2001) pp. 329-346

Kant, Immanuel. *Groundwork of the Metaphysics of Morals*. 1785. Trans. Jonathan Bennett (Early Modern Texts, 2005)

Mill, John Stewart. *Utilitarianism*. 1863. Trans. Jonathan Bennet (Early Modern Texts, 2005)

Miller, Kristie. "Presentism, Eternalism, and the Growing Block" *A Companion to the Philosophy of Time* (Wiley Books, 2013)

Parfit, Derek. "Personal Identity," *The Philosophical Review* 80 (1971) pp. 3-27

--- Parfit, Derek. *Reasons and Persons* (Oxford: Oxford University Press, 1984)

Ross, W. D. *What Makes Right Acts Right* (Oxford: Oxford University Press, 1930)

Sider, Thomas. *Four-Dimensionalism: An Ontology of Persistence and Time* (Clarendon Press, 2013)

Whiting, Jennifer. "Friends and Future Selves," *The Philosophical Review* 95 (1986) pp. 547-580