

Philosophical Perspectives, 28, *Ethics*, 2014

DIVIDED WE FALL

Fission and the Failure
of Self-Interest¹

Jacob Ross
University of Southern California

Fission cases, in which one person appears to divide into two persons, pose a serious challenge to many assumptions of commonsense ethics. Chief among these assumptions is the idea that self-interest is an important source of reasons. We normally think that something like the following is true:

Self-Interest Thesis: For any person S , and any two possibilities p_1 and p_2 , S has an important kind of reason, namely self-interested reason, to prefer p_1 to p_2 whenever S should expect that S would fare better if p_1 were to obtain than if p_2 were to obtain.²

However, given certain widely-held metaphysical assumptions, the Self-Interest Thesis has highly counterintuitive implications in cases involving fission. In particular, given the assumption that a person would not survive fission, the Self-Interest Thesis appears to imply that we have an important kind of reason to avoid fission, since fission would terminate our lives and hence reduce our total welfare. And yet it seems, intuitively, that we needn't have any important kind of reason to avoid fission. For it seems that, in many cases, fission could be just as good as ordinary survival with respect to everything we should care about. And so the Self-Interest Thesis appears to get things wrong.³

Many philosophers have responded to this kind of argument by rejecting the metaphysical assumptions on which it rests. In particular, many philosophers have suggested that fission is a process that one would survive.⁴ Some have suggested that one would survive fission as *one or other* of the two individuals who emerge from the operation, and some have suggested that one would survive fission as *both* these individuals. Still others have suggested that, when fission occurs, it is undergone by two distinct persons who are initially colocated and who part ways at the time of the operation. Adopting one of these alternative views, it has been argued, will allow us to hold on to our commonsense understanding

of self-interested reasons while avoiding counterintuitive implications in fission cases.⁵

The principal aim of this paper is to argue that these and all other attempts at metaphysical solutions to the challenge raised by fission cases are doomed to failure. For I will argue that, no matter what view we may adopt concerning what would happen in cases involving fission and related processes, we will be committed to highly counterintuitive implications so long as we accept the Self-Interest Thesis. This argument will occupy the first two sections of the paper. In §1, I will focus on four prominent metaphysical views. And I will show that there is a case in which all four views have counterintuitive implications about what we have reason to prefer, assuming the Self-Interest Thesis. Then, in §2, I will generalize and argue that the Self-Interest Thesis will commit us to counterintuitive implications *regardless* of what metaphysical view we adopt.

If the argument of the first two sections is successful, then it will show that there can be no purely metaphysical solution to the clash between the Self-Interest Thesis and our intuitions about what we have reason to prefer. Thus, something must give on the ethical side: either we must reject or revise the Self-Interest Thesis, or else we must abandon some of our ethical intuitions. And doing either of these things would have significant methodological repercussions. For, historically, one of the main ways of arguing for or against theories of personal identity has been to present scenarios involving strange processes (fission, brain transplants, teletransportation, etc.), and arguing that the theories in question have right or wrong implications about what someone would have self-interested reason to prefer in these cases.⁶ Typically, these arguments implicitly assume the Self-Interest Thesis, in addition to assuming the reliability of our ethical intuitions about such strange cases. But if my argument is successful, then it follows that we must reject one or other of these assumptions, and hence that we must reject such arguments.

In the third and final section, I take up the most popular response to the ethical challenge of fission cases that involves revising our ethical views rather than our metaphysical views. This response, first advocated by Derek Parfit, is to maintain that what fundamentally matters is not identity, but rather “relation R,” which is a relation between person-stages.⁷ I argue that this view, like the commonsense view it replaces, cannot avoid counterintuitive implications in certain cases involving fission and related processes.

1. Why Four Metaphysical Views Clash with the Self-Interest Thesis

In this section, I will present four views about what would happen in cases involving fission and related processes. And I will then present a case in which

each of these views, when combined with the Self-Interest Thesis, has counterintuitive implications concerning what one would have reason to prefer. But first, some clarificatory remarks are in order.

1.1. Stage Setting

Let's start with some terminology. By an *individual* I mean a living human being consisting of an ordinary human body and brain with an ordinary human psychology. (It will be useful to have this term, in addition to the term "person," since in some cases it will be controversial how to individuate persons.)

By a *perfectly divisible* individual, I mean an individual with two unusual features. First, she is perfectly symmetrical, in the sense that her right half is a mirror image of her left half, down to the molecular level. Second, the left and right halves of her brain are redundant realizers of her mental life, in the sense that each one provides a sufficient basis for all her psychological features (mental states, dispositions, abilities, etc.). Thus, the loss of either half of her brain would make no psychological difference.

By *fission* I mean an operation in which the right and left halves of a perfectly divisible individual (call them *L* and *R*) are instantaneously separated from one another, and immediately thereafter *L* is connected to a newly created molecular duplicate of *R*, and *R* is connected to a newly created molecular duplicate of *L*, so as to form two individuals that are each qualitatively identical with the original individual. In cases of fission, I will typically use "Lefty" and "Righty" as proper names referring to individuals that emerge from the operation with the original left half, and with the original right half, respectively.

By *semi-replacement* I mean an operation in which either the left or the right half of a perfectly divisible individual is instantaneously annihilated and, immediately thereafter, a newly created molecular duplicate of the annihilated half is connected to the remaining half, so as to form one individual who is qualitatively identical with the original individual. By *left-preserving semi-replacement* I mean a semi-replacement operation in which the left half is preserved and the right half is replaced, and by *right-preserving semi-replacement* I mean such an operation in which the right half is preserved and the left half is replaced.

When I say that someone has a *torture-filled life*, I mean that, at all times at which she is conscious, she experiences nothing but intense physical pain, and her life contains none of the other things that are normally thought to make life worth living, such as achievement, interpersonal relationships, etc. Similarly, when I say that someone has a *bliss-filled life*, I mean that, at all times at which she is conscious, she experiences nothing but intense physical pleasure, and her life contains nothing else that is normally thought to make life worth living.

By the *counterpart* relation, I mean whatever relation it is that plays what we may call the *counterpart role*. For our purposes, the relevant features of this role are as follows:

- (i) For any object x , proposition p , and world w , “if p , then x would exist” is true in w whenever, in all the nearest worlds to w in which p is true, x has a counterpart.
- (ii) For any object x , proposition p , and world w , “if p , then x would *not* exist” is true in w whenever, in all the nearest worlds to w in which p is true, x has *no* counterpart.
- (iii) For any object x , proposition p , feature A , and world w , “if p , then x would be A ” is true in w whenever, in all the nearest worlds to w in which p is true, all of x ’s counterparts are A .

In adopting the language of counterparts, I don’t mean to be taking any stand on the debate between the Lewisian and Kripkean theories of modality.⁸ Thus, I want to allow for the Kripkean view which maintains that the same object can exist in many possible worlds, and that the truth-maker for counterfactual claims about a given object is *that very object*, rather than some distinct but similar object. If this Kripkean view is correct, then the identity relation is what plays the counterpart role, and so what I am calling the ‘counterpart relation’ is simply the identity relation.

Lastly, I will outline four assumptions I will be making in what follows. The first two assumptions concern what we may call *same person cases*—cases where the person at issue would exist regardless of which alternative obtains:

- (1) For any person S , and any possibilities p_1 and p_2 , if S would have a bliss-filled life regardless of which of these possibilities were to obtain, and S would experience *more bliss* given p_1 , then S would fare *better* given p_1 .
- (2) For any person S , and any possibilities p_1 and p_2 , if S would have a torture-filled life regardless of which of these possibilities were to obtain, and S would experience *more torture* given p_1 , then S would fare *worse* given p_1 .

The remaining two assumptions concern what we may call *non-identity cases*—cases where the person at issue would exist given one alternative but not the other:

- (3) It’s better for someone to have a bliss-filled life than not to exist at all. Thus, for any person S , and any possibilities p_1 and p_2 , if S would have a *bliss-filled* life given p_1 and would not exist at all given p_2 , then S would fare *better* given p_1 .
- (4) It’s worse for someone to have a torture-filled life than not to exist at all. Thus, for any person S , and any possibilities p_1 and p_2 , if S would have

a *torture-filled* life given p_1 and would not exist at all given p_2 , then S would fare *worse* given p_1 .

There are several ways in which one could object to these assumptions. First, one could maintain, as some opponents abortion and euthanasia maintain, that existing, or continuing to exist, is *always better* for someone than not existing, or than ceasing to exist.⁹ Someone who holds such a view may reject (2) or (4) or both. But while some of my arguments employ these premises, variants of these arguments could be given without them. To construct these variants, we would need to replace the comparisons I make between torture-filled lives with corresponding comparisons between bliss-filled lives.

Second, one could maintain, as Schopenhauer and his followers maintain, that existing, or continuing to exist, is *always worse* for someone than not existing, or than ceasing to exist.¹⁰ Someone who holds such a view may reject (1) or (3) or both. But while some of my arguments employ these premises, variants of these arguments could be given without them. To construct these variants, we would need to replace the comparisons I make between bliss-filled lives with corresponding comparisons between torture-filled lives.

Third, one could maintain that it makes no sense to say that someone is better off or worse off for existing.¹¹ Someone who holds this view will reject both premise (3) and premise (4). Even if we reject both these premises, we can still give versions of the arguments I will present. To do so, however, we would need to strengthen the Self-Interest Thesis, so that it states not only that we have an important kind of reason to prefer possibilities where we should expect to fare better to possibilities where we should expect to fare worse, but also that we have such a reason to prefer possibilities where we would have a bliss-filled life to possibilities where we would not exist, and that we likewise have such a reason to prefer possibilities where we would not exist to possibilities where we would have a torture-filled life.

A final objection one might give to my four assumptions is that they all assume too narrow a supervenience base for a person's welfare. For one might hold that we should adopt a broad conception of a person's welfare on which it can depend not only on how this person's life goes while she is alive, but also on what happens to others after she ceases to exist.¹² In particular, following Aristotle,¹³ one might propose a conception of a person's welfare on which it can depend on what happens posthumously to this person's *children*. Or, analogously, one might propose a conception of a person's welfare on which it can depend on what happens posthumously to her *fission products*. I will take up this kind of view in §3, where I will argue that it will not allow us to reconcile the Self-Interest Thesis with our ethical intuitions.

1.2. The Problem Case

In this section, I will present a case which, as I will later show, raises a problem for a number metaphysical views that have been proposed. However, before considering the case that will be the focus of my argument, it will be useful to begin by considering a simpler case.

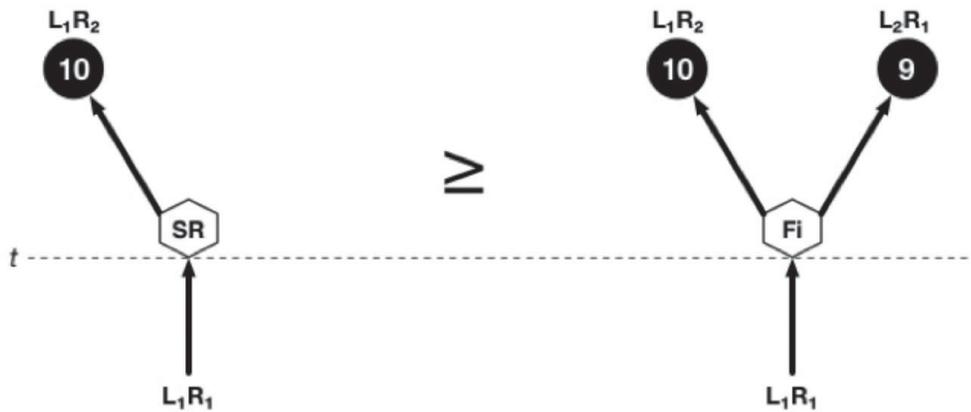
Suppose that, up until time t , a perfectly divisible person, named Clive Van Cleave, consists of a left half L_1 and a right half R_1 , and he experiences nothing but torture. Then, at t , he chooses the first of the two alternatives as follows.

Single Torture

There occurs a left-preserving semi-replacement operation in which R_1 is replaced with R_2 . Then whoever emerges from the operation experiences 10 years of torture before being destroyed.

Double Torture

There occurs a fission operation from which Lefty emerges with L_1 and R_2 and Righty emerges with L_2 and R_1 . Then Lefty experiences 10 years of torture before being destroyed, whereas Righty experiences 9 years of torture before being destroyed.

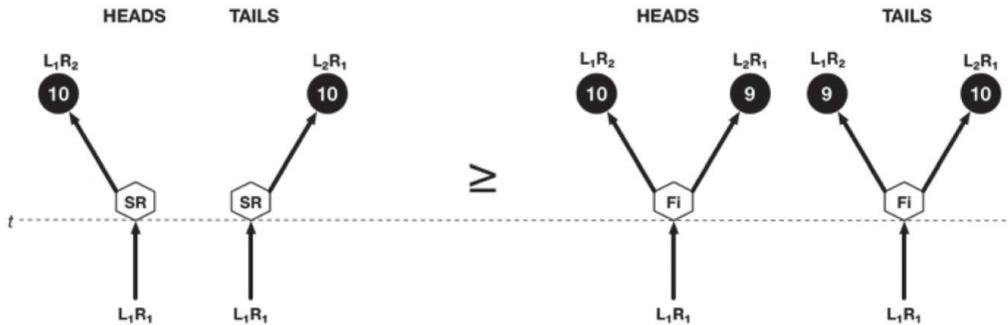


(In this and in all future diagrams, hexagons represent operations, black circles represent torture, and white circles represent bliss.) Note that these two alternatives do not differ with respect to what happens to Clive's original left half, L_1 : in both cases, L_1 ceases to be connected to R_1 and becomes connected to R_2 instead, and the resulting individual is tortured for 10 years before being destroyed. These alternatives differ only with respect to what happens to Clive's original right half, R_1 : in *Single Torture*, R_1 is simply destroyed, whereas in *Double Torture* it is connected to L_2 and the resulting individual is tortured for 9 years before being destroyed. It hardly seems that this difference could give Clive a reason to prefer *Double Torture*. And so it seems, intuitively, that Clive has no important kind of reason to prefer *Double Torture* to *Single Torture*.

With this in mind, let's turn to the case that will be our main focus. Suppose, once again, that, up until time t , Clive consists of L_1 and R_1 , and he experiences nothing but torture. But this time let's suppose that, at time t , Clive chooses the first of the two alternatives as follows.

Stochastic Single Torture

Depending on the outcome of a fair coin toss, there occurs either a left-preserving semi-replacement operation in which R_1 is replaced with R_2 (if the coin comes up heads) or a right-preserving semi-replacement operation in which L_1 is replaced with L_2 (if the coin comes up tails). Then whoever emerges from the operation is tortured for 10 years before being destroyed.



Stochastic Double Torture

There occurs a fission operation from which Lefty emerges with L_1 and R_2 and Righty emerges with L_2 and R_1 . Then, depending on the outcome of a fair coin toss, Lefty and Righty are tortured either for 10 and 9 years respectively (if the coin comes up heads) or for 9 and 10 years respectively (if the coin comes up tails).

Now suppose the coin comes up heads. On this supposition, *Stochastic Single Torture* is equivalent to *Single Torture*, and *Stochastic Double Torture* is equivalent to *Double Torture*. And, as we have seen, it seems intuitively that Clive has no important kind of reason to prefer *Double Torture* to *Single Torture*. And so it seems that, on the supposition that the coin comes up heads, Clive can have no important kind of reason to prefer *Stochastic Double Torture* to *Stochastic Single Torture*.

Next, suppose the coin comes up tails. On this supposition, *Stochastic Single Torture* is equivalent to the mirror image of *Single Torture*, and *Stochastic Double Torture* is equivalent to the mirror image of *Double Torture*. And, just as it seems, intuitively, that Clive has no important kind of reason to prefer *Double Torture* to *Single Torture*, so it seems, intuitively, that he has no important kind of reason to prefer the mirror image of the former to the mirror image of the latter. And so it seems that, on the supposition that the coin comes up tails, Clive can have no important kind of reason to prefer *Stochastic Double Torture* to *Stochastic Single Torture*.

But if Clive can have no such reason either on the supposition that the coin comes up heads or on the supposition that it comes up tails, then it follows that he can have no such reason unconditionally. Hence, if our intuition about the

first case is correct, then it seems that, at t , Clive can have no important kind of reason to prefer *Stochastic Double Torture* to *Stochastic Single Torture*.

1.3. The Failure of Four Views

In this section, I will consider four metaphysical views, and show that, if the self-interest theorist (i.e., the proponent of the Self-Interest Thesis) accepts any of these views, then her view will violate our intuition about the case just considered.

Most philosophers would agree that a person would survive semi-replacement, especially since, given the way we have defined this operation, anyone who undergoes it must be perfectly divisible, so the half of the brain that is not replaced is a sufficient basis for all the person's psychological features.¹⁴ Philosophers disagree, however, concerning what would happen to someone who underwent fission. Four of the most prominent views are as follows:

No Survivor View: Anyone who undergoes fission would not survive.¹⁵

Asymmetry View: Anyone who undergoes fission would survive as one or other of Lefty and Righty, but not both.¹⁶

Cohabitation View: Whenever fission occurs, there are two distinct persons involved all along. Prior to the operation, these two persons are collocated, and after the operation one survives as Lefty and the other as Righty.¹⁷

Double Identity View: Whenever fission occurs, only one person is present before the operation whereas two distinct persons are present after the operation, Lefty and Righty. And yet the person who is present before the operation survives as Lefty, and this person likewise survives as Righty.¹⁸

In §2, I will consider what happens if we reject the assumption that one would survive semi-replacement. But, for the time being, I will take this assumption for granted. And I will show that this view of semi-replacement, combined with any one of the four views of fission just outlined, will commit the self-interest theorist to the counterintuitive claim that Clive has an important kind of reason, namely self-interested reason, to prefer *Stochastic Double Torture* to *Stochastic Single Torture*.

To see why this is so, let's consider these four views of fission in turn, beginning with the No Survivor View. According to this view, Clive would not survive fission. Hence, if he were to choose *Stochastic Double Torture*, then he would undergo no torture whatsoever after the time of the operation. We are assuming, however, that anyone who undergoes semi-replacement survives the operation. And so it follows that, in the actual world, where Clive chooses *Stochastic Single Torture*, he will survive the operation and go on to be tortured

for 10 years before being destroyed. And so Clive should expect that he would fare better if he were to choose *Stochastic Double Torture* instead of *Stochastic Single Torture*. Hence it follows, from the Self-Interest Thesis, that he has an important kind of reason, namely self-interested reason, to prefer *Stochastic Double Torture*, contrary to intuition.

Next, consider the Asymmetry View. According to this view, if Clive were to choose *Stochastic Double Torture*, then he would survive the fission operation as either Lefty or Righty. And, regardless of whether he survived as Lefty or as Righty, a fair coin toss would determine whether he goes on to be tortured for 9 years or for 10 years. Hence, Clive's expectation for how much torture he would undergo after t if he were to choose *Stochastic Double Torture* should be 9.5 years. However, it follows from our assumptions that in the actual world, where Clive chooses *Stochastic Single Torture*, he will survive and be tortured for 10 years after t . And so, once again, he should expect that he would fare better given *Stochastic Double Torture*. Hence it follows from the Self-Interest Thesis that he has an important kind of reason, namely self-interested reason, to prefer *Stochastic Double Torture*, contrary to intuition.

Next, consider the Cohabitation View. On this view, if *Stochastic Double Torture* were chosen, then there would be exactly two persons involved, *Lefty and Righty*. And so Lefty and Righty are the only candidates for being counterparts of Clive in worlds in which *Stochastic Double Torture* is chosen. And Lefty and Righty each experience either 9 or 10 years of torture after t , depending on the outcome of a fair coin toss. Hence, in all worlds in which *Stochastic Double Torture* is chosen, every counterpart of Clive awaits an expected 9.5 years of torture after t . Thus, given the way we are understanding the counterpart relation, it follows that, if *Stochastic Double Torture* were chosen, then Clive would await an expected 9.5 years of torture.¹⁹ However, it follows from our assumptions that, in the actual world, where Clive chooses *Stochastic Single Torture*, he will survive and be tortured for 10 years after t . And so, once again, it follows that he should expect to fare better given *Stochastic Double Torture*. Hence it follows from the Self-Interest Thesis that he has an important kind of reason, namely self-interested reason, to prefer *Stochastic Double Torture*, contrary to intuition.

Last, let's consider the Double Identity View. It isn't immediately obvious what this view implies about the case we've been considering. And so it will be helpful to begin by considering some simpler cases. First, consider a case in which Clive undergoes a fission operation from which two individuals emerge, each weighing 180 lbs. According to the Double Identity view, since Clive survives as someone who weighs 180 lbs, Clive will weigh 180 lbs after the operation. Moreover, since, on this view, Clive does not survive as anyone who weighs either more or less than 180 lbs (in particular, he doesn't survive as the whole consisting of Lefty and Righty, and so he doesn't survive as anyone who weighs 360 lbs), it is not the case that, after the operation, Clive will weigh either more or less than 180 lbs. Thus, we may say that it is *unequivocally true* that, after the operation, Clive will weigh 180 lbs.

Next, consider a case in which Clive undergoes a fission operation from which two individuals emerge, each of whom goes on to be tortured for 9.5 years before being destroyed. In this case, the view implies that, since Clive survives as someone who is tortured for exactly 9.5 years after the operation, Clive will be tortured for exactly 9.5 years after the operation. And, since Clive does not survive as anyone who is tortured for either more or less than 9.5 years after the operation, it is not the case that, after the operation, Clive will be tortured for more or less than 9.5 years. And so this view implies that it is unequivocally true that, after the operation, Clive will be tortured for exactly 9.5 years.

Next, consider the following case:

Uniform Double Torture: Clive Undergoes a fission operation from which Lefty and Righty emerge. Then, depending on the outcome of a fair coin toss, either Lefty and Righty are each tortured for 10 years before being destroyed (if the coin comes up heads), or they are each tortured for 9 years before being destroyed (if the coin comes up tails).

On the Double Identity View, if the coin comes up heads, then everyone as whom Clive survives will be tortured for exactly 10 years. And so it's unequivocally true that, if the coin comes up heads, then Clive will be tortured for exactly 10 years. Similarly, if the coin comes up tails, then everyone as whom Clive survives will be tortured for exactly 9 years. And so it's unequivocally true that, if the coin comes up tails, then Clive will be tortured for exactly 9 years. Hence this view implies that it's unequivocally true that Clive will be tortured for either 10 years or 9 years after the operation, depending on whether the coin comes up heads or tails. And so this view implies that the expectation for how much torture Clive will experience after the operation is 9.5 years.

Lastly, let us compare *Uniform Double Torture* with *Stochastic Double Torture*. Note that these alternatives don't differ concerning what happens to Lefty: in either alternative, he is tortured for 10 years given heads and 9 years given tails. These alternatives do differ, however, concerning what happens to Righty: in one alternative he is tortured for 10 years given heads and 9 given tails and in the other he is tortured for 9 years given heads and 10 given tails. But since heads and tails are equally probable, these two alternatives are equivalent. And so it follows that, in *Stochastic Double Torture*, Lefty and Righty each stand the same chance of being tortured for either 9 or 10 years as they would in *Uniform Double Torture*. Hence, in *Stochastic Double Torture*, each individual as whom Clive survives stands the same chance of being tortured for either 9 or 10 years as he would in *Uniform Double Torture*. Consequently, in *Stochastic Double Torture*, it is unequivocally true that, after the operation, Clive will stand the same chance of being tortured for either 9 or 10 years as he would in *Uniform Double Torture*. And so it follows that the expectation for how much torture Clive would experience given *Stochastic Double Torture* must be 9.5 years.

Recall, however, that, on our current assumptions, Clive should expect to experience a full 10 years of torture given *Stochastic Single Torture*. And so we

arrive, once again, at the conclusion that Clive should expect to fare better given *Stochastic Double Torture* than given *Stochastic Single Torture*.²⁰ Hence it follows from the Self-Interest Thesis that he has an important kind of reason, namely self-interested reason, to prefer *Stochastic Double Torture*, contrary to intuition.

We may conclude, therefore, that it is no easy task to reconcile the Self-Interest Thesis with our intuitions about what one would have reason to prefer. For we have seen that these intuitions clash with the Self-Interest Thesis on each of the four views of fission we have considered, so long as we assume that anyone who undergoes semi-replacement would survive. I have not, however, established the more general claim that there is *no metaphysical view whatsoever* that will enable us to achieve this reconciliation—for I have not considered every possible view of fission, nor have I considered the view that one would not survive semi-replacement. Defending this more general claim will be the task of the next section.

2. Why Every Possible View Clashes with the Self-Interest Thesis

I will proceed as follows. After some clarificatory remarks in §2.1, I will go on, in §2.2, to present five metaphysical claims about cases involving fission and semi-replacement, and I will argue that the self-interest theorist will be committed to counterintuitive ethical implications if she maintains that *any* of these five propositions are *false*. Then, in §2.2, I will argue that the self-interest theorist will likewise be committed to such counterintuitive implications if she maintains that *all* of these five propositions are *true*. And so it will follow that the self-interest theorist cannot avoid counterintuitive ethical implications, since she will be committed to them regardless of what combination of truth values we assign to these five propositions. I will conclude, in §2.4, by arguing that the self-interest theorist can't solve the problems I've raised by maintaining that it's indeterminate what would happen in the cases at issue.

2.1. Some Further Stage Setting

Let's start with some terminology. By *incremental replacement* I mean an operation in which a person is unconscious for one year and, over this time, either all the cells in her left half or all the cells in her right half are replaced, one cell at a time, in rapid succession. Thus, incremental replacement can be thought of as a protracted variant of semi-replacement that extends over a one-year period of unconsciousness. We can distinguish two kinds of incremental replacement: *left-preserving incremental replacement*, in which the cells in the right half are replaced, and *right-preserving incremental replacement*, in which the cells in the left half are replaced.

For any person-halves A and B , by an AB person I mean a person whose left half is A and whose right half is B .

I will be using “choose,” but not “decide,” as a success term. Thus, as I will be using these terms, while one can *decide* upon an outcome without successfully bringing it about, one counts as *choosing* the outcome only if one’s deciding upon it results in its coming about.

In each of the examples to come, t will denote the time at which a choice is made among two or more operations, and t' will denote the time at which whoever emerges from the chosen operation ceases to exist.

Next, a note about the experiences that figure in the examples to come. In these examples, it will be important to keep the durations of torture, and similarly the durations of bliss, constant. Hence, instead of measuring amounts of torture and bliss in *years*, I will measure them abstractly in terms of *units* of torture and bliss, and I will be assuming that the experiences differ not in their duration, but rather in their intensity.

Lastly, let me remark on two ways in which my applications of the Self-Interest Thesis will differ in the present section from the previous section. In §1, the possibilities being compared were coarse-grained, in the sense that they were compatible with more than one possible world (e.g., worlds in which the coin comes up heads and worlds in which it comes up tails). Hence, in comparing someone’s welfare in different possibilities, we had to invoke *expected* levels of welfare. By contrast, in the present section, the possibilities being compared will be individual possible worlds. And, since a possible world fixes a unique outcome, how well someone should be *expected* to fare in given possible world coincides with the level of welfare this person *would have* if this world were actual. Consequently, in comparing possible worlds, we needn’t refer to expectations, and we can employ the following, simpler formulation of the Self-Interest Thesis:

Self-Interest Thesis (possible worlds formulation): For any person S and any two possible worlds w and w' , S has an important kind of reason, namely self-interested reason, to prefer w to w' whenever S would fare better if w were actual than if w' were actual.

The second way in which my applications of the Self-Interest Thesis will differ in the present section from the previous section is this. In §1, I applied the Self-Interest Thesis to pairs of possibilities that differ only with respect to how things go *after* the time of evaluation. By contrast, in the present section, I will often be applying this thesis to pairs of possibilities that differ with respect to how things go *up until* the time of evaluation. This kind of application seems legitimate, at least in some cases. For it seems that, on our commonsense view, someone can have self-interested reason to prefer one possibility to another in virtue of faring better *in the past* in one of these possibilities than in the other. For example, when someone, on her deathbed, is comparing the actual world to alternative possibilities in which her life went differently, it seems she has self-interested

reason to *wish* things had gone in ways in which she would have fared better, and to be *glad* that things didn't go in ways in which she would have fared worse.

It's unclear, however, whether, on the commonsense view, the Self-Interest Thesis can be applied to *all* pairs of possibilities that differ with respect to how things have gone prior to the time of evaluation. As a potential counterexample, consider the following pair of possibilities:²¹

- w_A : Cleo experiences no torture prior to t , and 1 unit of torture after t .
- w_B : Cleo experiences 2 units of torture prior to t , and no torture after t .

One could plausibly maintain that Cleo's concern for her *future* welfare should have priority, perhaps even lexical priority, over her concern for her *past* welfare.²² Hence, even though Cleo fares better overall in w_A than in w_B , one could plausibly deny that she has, at t , self-interested reason to prefer w_A . Hence, one could plausibly deny that the Self-Interest Thesis applies in this case. In order to exclude this kind of comparison, while at the same time allowing the retrospective comparisons that seem legitimate, I will restrict the Self-Interest Thesis as follows:

Self-Interest Thesis (restricted possible worlds formulation): For any person S and any two possible worlds w and w' , S has an important kind of reason, namely self-interested reason, to prefer w to w' whenever S would fare *better* overall, and at least as well after t , if w were actual than if w' were actual.

This restriction will be implicitly adopted in all the arguments to come. With these clarifications in place, we may now proceed with the main argument.

2.2. Five Propositions We Must Assume to Achieve the Desired Reconciliation

In this section I will argue that, in order to reconcile the Self-Interest Thesis with our ethical intuitions, we must accept five metaphysical propositions. This section is divided into five steps, one for each proposition. Each step follows the same pattern. I begin each step by presenting a case, and making an intuitive claim about what someone would have reason to prefer in this case. I then use a *reductio* argument to show that the self-interest theorist can accept this intuitive claim only if she accepts a metaphysical claim about the case in question. I then argue that, in accepting this particular claim, she will be committed to accepting a more general claim.

Step 1: Show that we must hold that anyone who chooses to undergo semi-replacement survives the operation.

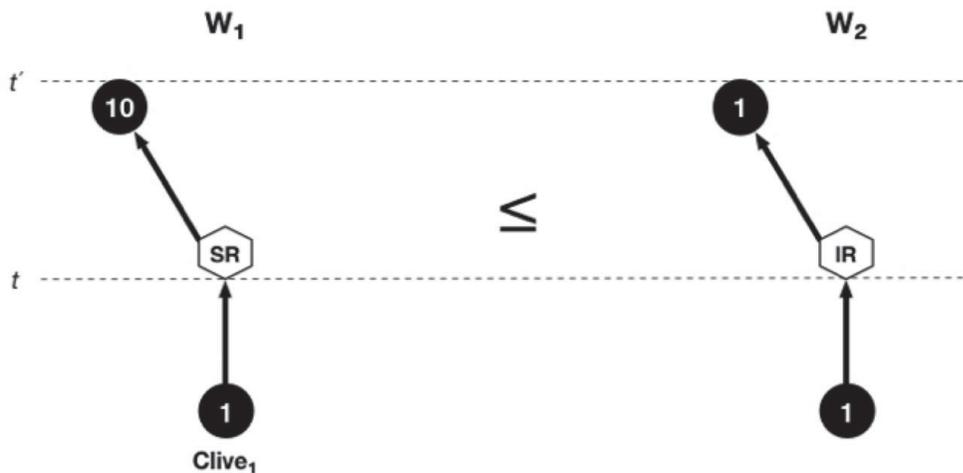
Case 1: Let w_1 be a world in which Clive₁, a perfectly divisible person consisting of L_1 and R_1 , chooses the first of the following two alternatives at t . And let w_2 be one of the nearest worlds to w_1 in which the second alternative is chosen at t instead. Assume that, in both worlds, whoever consists of L_1 and R_1 experiences nothing but 1 unit of torture prior to t . (as follows).

Left-Preserving Semi-Replacement₁

Throughout the 1 year interval after t , whoever consists of L_1 and R_1 is unconscious. During this time, all the cells in the right half of his body and brain are simultaneously replaced with their molecular duplicates. Whoever emerges from this operation goes on to experience 10 units of torture before being destroyed.

Left-Preserving Incremental Replacement₂

Throughout the 1 year interval after t , whoever consists of L_1 and R_1 is unconscious. Over this period, all the cells in the right half of his body and brain are incrementally replaced with their molecular duplicates. Whoever emerges from this operation goes on to experience 1 unit of torture before being destroyed.



Intuitive Claim 1: It is not the case that Clive₁ has, at t , an important kind of reason, namely self-interested reason, to prefer w_1 to w_2 .

Support for Intuitive Claim 1: If half of Clive₁'s cells are going to be replaced while he is unconscious, and if these cells are entirely redundant with respect to the preservation of psychological continuity (as they must be, since we have stipulated that he is perfectly divisible), then it seems he should be indifferent as to whether these cells are replaced simultaneously or incrementally. And surely he shouldn't prefer for such an operation to be followed by more torture rather than less torture. And

so it seems he couldn't have any important kind of reason, self-interested or otherwise, to prefer w_1 to w_2 .

Reductio Argument 1:

- 1A. Clive₁ does not survive the semi-replacement operation in w_1 . (Supposed for *reductio*.)
- 1B. It follows that, while Clive₁ would have a torture-filled life in either world, he would experience more torture in w_2 . For, since we're supposing that Clive₁ does not survive the operation in w_1 , he can experience no torture after t in w_1 . By contrast, he would experience 1 unit of torture after t in w_2 . For, since w_2 is one of the nearest worlds to w_1 in which the second alternative is chosen rather than the first, it should be uncontroversial that Clive₁ has a counterpart in w_2 . And since it should likewise be uncontroversial that anyone who undergoes incremental replacement survives, we may infer that any counterpart of Clive₁ in w_2 must survive the operation and go on to experience 1 unit of torture after t .
- 1C. It follows, from assumption 2 from §1.1, that Clive₁ fares better in w_1 than he would in w_2 .
- 1D. Hence it follows, from the Self-Interest Thesis, that Clive₁ has, at t , an important kind of reason, namely self-interested reason, to prefer w_1 to w_2 (contrary to intuition).

Conclusion 1: In order to reconcile the Self-Interest Thesis with our ethical intuitions, we must maintain that Clive₁ survives the semi-replacement operation in w_1 .

Generalization1: Recall that w_1 is an arbitrary world in which *Left-Preserving Semi-Replacement*₁ is chosen, and that Clive₁ is an arbitrary person who makes this choice. Hence, in order to reconcile the Self-Interest Thesis with our ethical intuitions, we must accept the following, more general claim:

X: In any world in which *Left-Preserving Semi-Replacement*₁ is chosen, whoever makes this choice survives the operation.

Moreover, the only feature of *Left-Preserving Semi-Replacement*₁ that is plausibly relevant to the truth of X is that the operation it involves consists of semi-replacement. The other features, such as the fact that the operation is *left*-preserving rather than *right*-preserving, and the amount of torture that follows the operation, are clearly irrelevant. And so we must accept the following, still more general claim:

P1: Anyone who chooses to undergo semi-replacement survives the operation.

Step 2: Show that we must hold that, whenever semi-replacement is chosen, whoever emerges from the operation already existed at the time of choice.

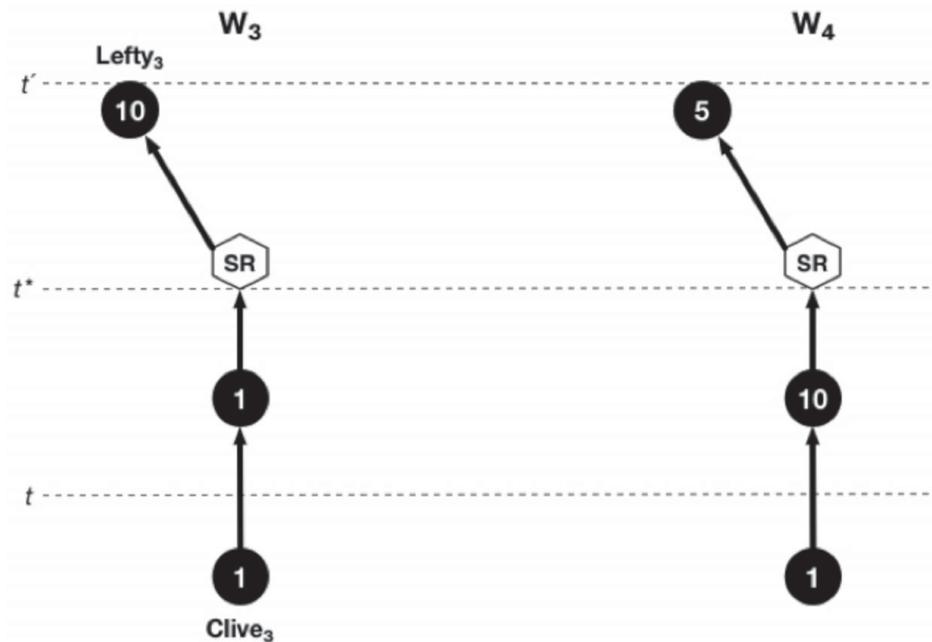
Case 2: Let w_3 be a world in which Clive₃, consisting of L_1 and R_1 , chooses the first of the following two alternatives at t . And let w_4 be one of the nearest worlds to w_3 in which the second alternative is chosen instead. Assume that, in both worlds, whoever consists of L_1 and R_1 experiences nothing but 1 unit of torture prior to t . Let Lefty₃ be someone who emerges from the operation in w_3 . (as follows).

Left-Preserving Semi-Replacement₃

After t , whoever consists of L_1 and R_1 experiences 1 unit of torture, followed by a left-preserving semi-replacement operation that occurs at t^* . Whoever emerges from this operation goes on to experience 10 units of torture before being destroyed at t' .

Left-Preserving Semi-Replacement₄

After t , whoever consists of L_1 and R_1 experiences 10 units of torture, followed by a left-preserving semi-replacement operation that occurs at t^* . Whoever emerges from this operation goes on to experience 5 units of torture before being destroyed at t' .



Intuitive Claim 2: It is not the case that, in order for everyone in w_3 to have all the preferences favored by their self-interested reasons, at t' Lefty₃ must simultaneously prefer w_3 to w_4 and prefer w_4 to w_3 .

Support for Intuitive Claim 2: It seems intuitively clear that, if everyone in w_3 had all the preferences favored by their self-interested reasons, then everyone would have consistent preferences, and so Lefty₃ would not have conflicting preferences at t' .

Additional Remark: The aim of the *reductio* argument below will be to derive the negation of Intuitive Claim 2. Note that this negation follows from the conjunction of the following two claims:

- (i) In order for everyone in w_3 to have all the preferences favored by their self-interested reasons, Lefty₃ must prefer w_3 to w_4 at t' .
- (ii) In order for everyone in w_3 to have all the preferences favored by their self-interested reasons, Lefty₃ must prefer w_4 to w_3 at t' .

In giving this *reductio* argument, it will suffice for us derive (ii), since we are already committed to (i). To see why we are committed to (i), we must focus on Clive₃, the person who chooses the operation in w_3 . It follows from P1 that Clive₃ survives the semi-replacement operation that occurs at t^* . And so in w_3 he experiences a total of 12 units of torture (1 unit + 1 unit + 10 units). This is less than the 16 units of torture he would experience in w_4 (1 unit + 10 units + 5 units). And so it follows from the Self-Interest Thesis that Clive₃ has self-interested reason to prefer w_3 to w_4 at t' . Hence, in order for everyone to have all the preferences favored by their self-interested reasons, Clive₃ must prefer w_3 to w_4 at t' .

Note, further, that Clive₃ stands in a very close relationship to Lefty₃ at t' . For, since Clive₃ survives the semi-replacement operation, he must be present at t' . And since Lefty₃ emerges from this same operation, he must likewise be present at t' . But there is only one body and brain present at t' . And so it follows that either Clive₃ and Lefty₃ are the same person, or else they are two persons sharing an ordinary human body and brain. Either way, all their mental states, including their preferences, must be identical. Hence, Clive₃ can prefer w_3 to w_4 at t' only if Lefty₃ likewise has this preference. And so it follows that, in order for everyone to have all the preferences favored by their self-interested reasons, Lefty₃ must prefer w_3 to w_4 at t' .

Hence, to derive the negation of Intuitive Claim2, all that remains is to derive (ii). And to do so, it will suffice to derive the claim that Lefty₃ has, at t' , self-interested reason to prefer w_4 to w_3 . This latter claim will be the conclusion of the following *reductio* argument.

Reductio Argument 2:

- 2A. Lefty₃ does not yet exist at t . (Supposed for *reductio*.)
- 2B. It follows that, while Lefty₃ would not experience anything but torture in either world, he experiences more torture in w_3 than he would in w_4 . For, if he doesn't yet exist at t , then he must come into existence at t^* , the time of the semi-replacement operation. And, while he experiences 10 units of torture after t^* in w_3 , his counterpart can experience no more than 5 units of torture after t^* in w_4 .
- 2C. It follows, from assumptions 2 and 4 from §1.1, that Lefty₃ fares worse in w_3 than he would in w_4 .
- 2D. Hence it follows, from the Self-Interest Thesis, that Lefty₃ has, at t' , self-interested reason to prefer w_4 to w_3 . And this implies (as we saw in the Additional Remark, above) that, in order for everyone to have all the preferences required by their self-interested reasons, at t' Lefty₃ must simultaneously prefer w_3 to w_4 and prefer w_4 to w_3 (contrary to intuition).

Conclusion 2: In order to reconcile the Self-Interest Thesis with our ethical intuitions, we must maintain that Clive₃ already exists at t in w_3 .

Generalization 2: Since w_3 is an arbitrary world in which *Left-Preserving Semi-Replacement*₃ is chosen, and Lefty₃ is an arbitrary person who emerges from this operation, we must maintain that the same holds for all such persons and all such worlds. And since the only feature of *Left-Preserving Semi-Replacement*₃ that is plausibly relevant to the claim at issue is that it involves a semi-replacement operation, we must maintain that the same holds for every world in which such an operation is chosen. And so we must accept the following:

P2: Whenever semi-replacement is chosen, whoever emerges from the operation already exists at the time of choice.

Step 3: *Show that we must hold that, whenever fission is chosen, whoever emerges with the original left half would have existed at the time of choice if left-preserving semi-replacement had been chosen instead.*

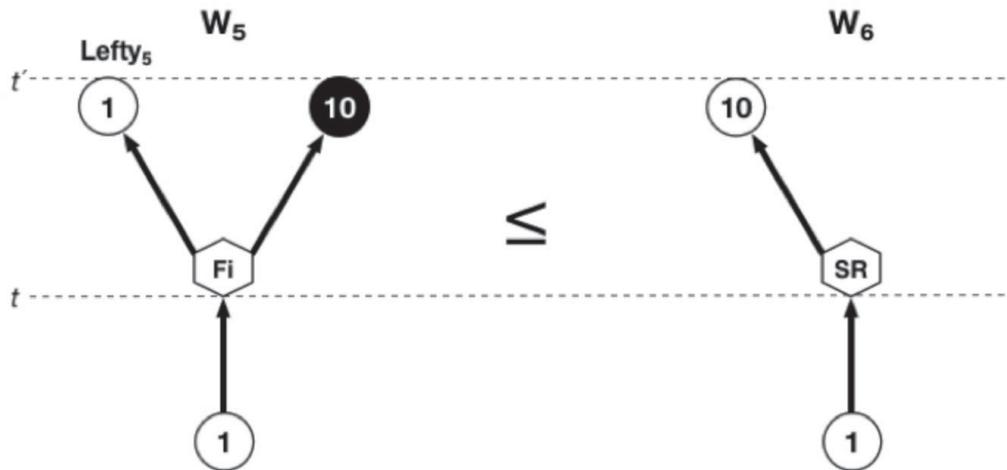
Case 3: Let w_5 be a world in which an $L_I R_I$ person chooses the first of the following two alternatives at t , and let w_6 be one of the nearest worlds to w_5 in which the second alternative is chosen instead. This time, assume that, in both worlds, whoever consists of L_I and R_I experiences nothing but 1 unit of *bliss* prior to t . Let Lefty₅ be someone who, in w_5 , emerges from the fission operation consisting of L_I and R_I .²³ (as follows).

Fission₅

Immediately after t , whoever consists of L_1 and R_1 undergoes fission. Then whoever emerges with the original left half experiences 1 unit of bliss before being destroyed, and whoever emerges with the original right half experiences 10 units of torture before being destroyed.

Left-Preserving Semi-Replacement₆

Immediately after t , whoever consists of L_1 and R_1 undergoes left-preserving semi-replacement. Then whoever emerges experiences 10 units of bliss before being destroyed.



Intuitive Claim 3: It is not the case that Lefty₅ has, at t' , an important kind of reason, namely self-interested reason, to prefer w_5 to w_6 .

Support for Intuitive Claim 3: Note that w_5 can be derived from w_6 by adding 10 units worth of torture-filled existence and decreasing the intensity of bliss-filled existence from 10 units to 1 unit. And surely no one should regard either of these differences as counting in favor of w_5 .

Reductio Argument 3:

- 3A. It is not the case that Lefty₅ has a counterpart in w_6 who is present at t . (Supposed for *reductio*.)
- 3B. It follows that, while Lefty₅ has a bliss-filled life in w_5 , he would not exist if w_6 were actual. For if he has no counterpart in w_6 who is present at t , then he has no counterpart whatsoever in w_6 . For the only plausible way in which he could have a counterpart in w_6 who is not present at t would be if this counterpart came into existence in the semi-replacement operation. And this possibility is ruled out by P2.

- 3C. It follows, from assumption 3 from §1.1, that Lefty₅ fares better in w_5 than he would in w_6 .
- 3D. Hence it follows, from the Self-Interest Thesis, that Lefty₅ has, at t' , an important kind of reason, namely self-interested reason, to prefer w_5 to w_6 (contrary to intuition).

Conclusion 3: In order to reconcile the Self-Interest Thesis with our ethical intuitions, we must maintain that Lefty₅ has a counterpart in w_6 who is present at t .

Generalization 3: Since w_5 is an arbitrary world in which *Fission*₅ is chosen at t , and since w_6 is an arbitrary nearest world to w_5 in which *Left-Preserving Semi-Replacement*₆ is chosen instead, we must maintain that the same holds for all such pairs of worlds. And since, apart from the operations involved in these alternatives, none of their other features are plausibly relevant to the claim at issue, we must maintain that the same holds for every such pair of nearest worlds in which these kinds of operations are chosen, respectively. And so we must accept the following:

- Y: For any world w in which a fission operation is chosen at a given time t , whoever emerges from this operation with the original left half has a counterpart who is present at t in all the nearest worlds to w in which a left-preserving semi-replacement operation is chosen instead.

And this entails:

- P3:** Whenever fission is chosen, whoever emerges with the original left half would have existed at the time of choice if left-preserving semi-replacement had been chosen instead.

Step 4: *Show that we must hold that, whenever fission is chosen, whoever emerges with the original left half would not have existed if right-preserving semi-replacement had been chosen instead.*

Case 4: Let w_7 be a world in which an $L_I R_I$ person chooses the first of the following two alternatives at t , and let w_8 be one of the nearest worlds to w_7 in which the second alternative is chosen instead. Assume that, in both worlds, whoever consists of L_I and R_I experiences nothing but 1 unit of torture prior

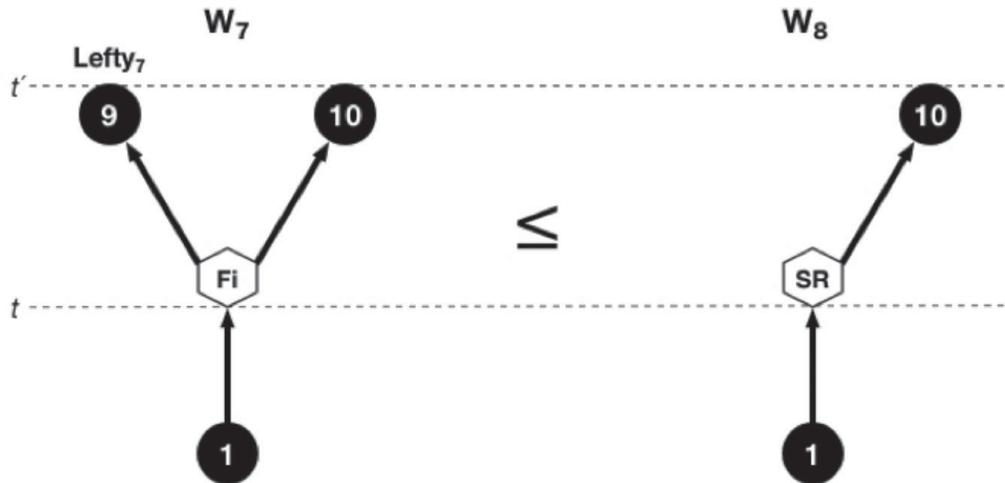
to t . Let Lefty_7 be someone who, in w_7 , emerges from the fission operation with the original left half. (as follows).

Fission₇

Immediately after t , whoever consists of L_1 and R_1 undergoes fission. Then whoever emerges with the original left half experiences 9 units of torture before being destroyed, and whoever emerges with the original right half experiences 10 units of torture before being destroyed.

Right-Preserving Semi-Replacement₈

Immediately after t , whoever consists of L_1 and R_1 undergoes right-preserving semi-replacement. Then whoever emerges experiences 10 units of torture before being destroyed.



Intuitive Claim 4: It is not the case that Lefty_7 has, at t' , an important kind of reason, namely self-interested reason, to prefer w_7 to w_8 .

Support for Intuitive Claim 4: Note w_7 can be derived from w_8 by adding 9 years of torture-filled existence. And surely no one should regard this difference as counting in favor of w_7 .

Reductio Argument 4:

- 4A. Lefty_7 has a counterpart in w_8 . (Supposed for *reductio*.)
- 4B. It follows that, while Lefty_7 would have a torture-filled life in either world, he would experience more torture in w_8 . For it follows from P1 that Lefty_7 's counterpart in w_8 can't cease to exist in the semi-replacement operation, and it follows from P2 that he can't come into

existence in this operation. And so he must be present both before and after the operation, thereby experiencing a total of 11 units of torture. By contrast, in w_7 , Lefty₇ can experience no more than 10 units of torture.

- 4C. It follows, from assumption 2 from §1.1, that Lefty₇ fares better in w_7 than he would in w_8 .
- 4D. Hence it follows, from the Self-Interest Thesis, that Lefty₇ has, at t' , an important kind of reason, namely self-interested reason, to prefer w_7 to w_8 (contrary to intuition).

Conclusion 4: In order to reconcile the Self-Interest Thesis with our ethical intuitions, we must maintain that Lefty₇ has no counterpart in w_8 .

Generalization 4: Since w_7 is an arbitrary world in which *Fission*₇ is chosen at t , and since w_8 is an arbitrary nearest world to w_7 in which *Right-Preserving Semi-Replacement*₈ is chosen instead, we must maintain that the same holds for all such pairs of worlds. And since, apart from the operations involved in these alternatives, none of their other features are plausibly relevant to the claim at issue, we must maintain that the same holds for every such pair of nearest worlds in which these kinds of operations are chosen, respectively. And so we must maintain the following:

P4: Whenever fission is chosen, whoever emerges with the original left half would not have existed if right-preserving semi-replacement had been chosen instead.

Step 5: *Show that we must hold that, whenever left-preserving semi-replacement is chosen, whoever chooses the operation would have existed at the time of choice if right-preserving semi-replacement had been chosen instead.*

Case 5: Let w_9 be a world in which Clive₉, consisting of L_1 and R_1 , chooses the first of the following two alternatives at t . And let w_{10} be one of the nearest worlds to w_9 in which the second alternative is chosen instead. Assume that, in both worlds, whoever consists of L_1 and R_1 experiences nothing but 1 unit of torture prior to t . (as follows).

Left-Preserving Semi-Replacement₉

Immediately after t , whoever consists of L_1 and R_1 undergoes left-preserving semi-replacement. Whoever emerges from the operation goes on to experience 1 unit of torture before being destroyed.

Right-Preserving Semi-Replacement₁₀

Immediately after t , whoever consists of L_1 and R_1 undergoes right-preserving semi-replacement. Whoever emerges from the operation goes on to experience 10 units of torture before being destroyed.



Intuitive Claim 5: It is not the case that Clive₉ has, at t , an important kind of reason, namely self-interested reason, to prefer w_{10} to w_9 .

Support for Intuitive Claim 5: It follows from our definition of semi-replacement that anyone who undergoes such an operation must be perfectly divisible. Hence it follows that, at t , the two halves of Clive₉ must be perfectly symmetrical. And so it seems that Clive₉ should be indifferent as to which of these halves is replaced. And surely he has no reason to prefer for semi-replacement to be followed by more torture rather than less torture.

Reductio Argument 5:

- 5A. It is not the case that Clive₉ has a counterpart in w_{10} who is present at t . (Supposed for *reductio*.)
- 5B. It follows that, while Clive₉ has a torture-filled life in w_9 , he would not exist if w_{10} were actual. For, if he has no counterpart in w_{10} who is present at t , then he has no counterpart whatsoever in w_{10} . For the only plausible way in which he could have a counterpart in w_{10} who is not present at t would be if this counterpart comes into existence

in the semi-replacement operation. And this possibility is ruled out by P2.

5C. It follows, from assumption 4 from §1.1, that Clive₉ fares worse in w_9 than he would in w_{10} .

5D. Hence it follows, from the Self-Interest Thesis, that Clive₉ has, at t , an important kind of reason, namely self-interested reason, to prefer w_{10} to w_9 (contrary to intuition).

Conclusion 5: In order to reconcile the Self-Interest Thesis with our ethical intuitions, we must maintain that Clive₉ has a counterpart in w_{10} who is present at t .

Generalization 5: Since w_9 is an arbitrary world in which *Left-Preserving Semi-Replacement*₉ is chosen at t , and since w_{10} is an arbitrary nearest world to w_9 in which *Right-Preserving Semi-Replacement*₁₀ is chosen instead, we must maintain that the same holds for all such pairs of worlds. And since, apart from the operations involved in these alternatives, none of their other features are plausibly relevant to the claim at issue, we must maintain that the same holds for every such pair of nearest worlds in which these kinds of operations are chosen, respectively. And so we must maintain the following:

P5: Whenever left-preserving semi-replacement is chosen, whoever chooses the operation would have existed at the time of choice if right-preserving semi-replacement had been chosen instead.

2.3. Why We're in Trouble

I have argued that, in order to reconcile the Self-Interest Thesis with our ethical intuitions, we must accept propositions P1 through P5. I will now argue that, if we accept all these propositions, then the Self-Interest Thesis will commit us to a further counterintuitive ethical implication. And so it follows that there is no way to reconcile the Self-Interest Thesis with our ethical intuitions.

To see why this is so, let's suppose that P1 through P5, along with the Self-Interest Thesis, are all true. And let's consider a world, w_{11} , in which an L_1R_1 person chooses the first of the following three alternatives at t . Assume that, prior to t , this person experiences nothing but 1 unit of bliss. And assume that each of these three alternatives is an available option at t , in the sense that, if it were decided upon by an L_1R_1 person at t , then it would come about as specified as follows.

Fission₁₁

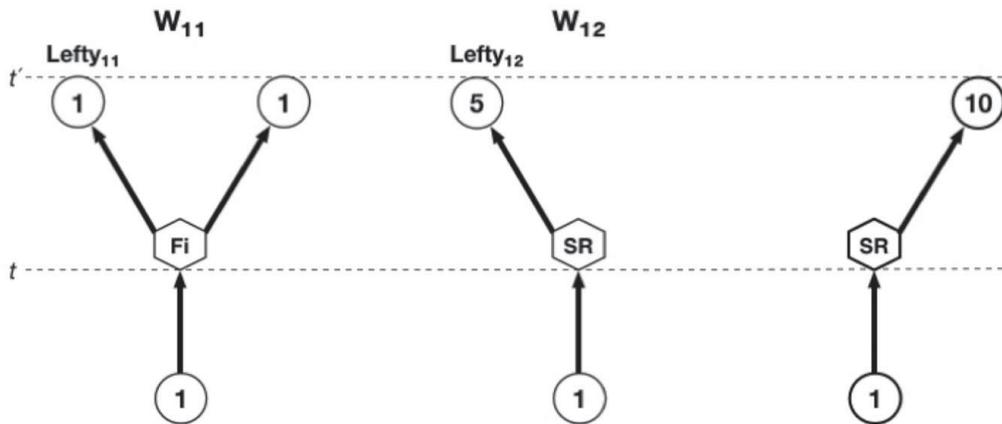
Immediately after t , whoever consists of L_1 and R_1 undergoes fission. Then each of the two individuals who emerge from the operation goes on to experience 1 unit of bliss before being destroyed.

Left-Preserving Semi-Replacement₁₂

Immediately after t , whoever consists of L_1 and R_1 undergoes left-preserving semi-replacement. Then whoever emerges from the operation goes on to experience 5 units of bliss before being destroyed.

Right-Preserving Semi-Replacement₁₃

Immediately after t , whoever consists of L_1 and R_1 undergoes right-preserving semi-replacement. Then whoever emerges from the operation goes on to experience 10 units of bliss before being destroyed.



Let $Lefty_{11}$ be someone who emerges from the fission operation in w_{11} with the original left half. On the basis of our current assumptions, I will now argue that, in w_{11} at t' , $Lefty_{11}$ has self-interested reason to wish he had chosen the second of the three alternatives, *Left-Preserving Semi-Replacement₁₂*. To see why this is so, we should first note that $Lefty_{11}$ could have chosen this second alternative at t . For there are only two ways in which he could have failed to have this option. First, *Left-Preserving Semi-Replacement₁₂* might not have been an available option at all, in the sense that it would not have come about even if it had been decided upon by an L_1R_1 person at t . But this possibility is ruled out by our description of the case. Second, *Left-Preserving Semi-Replacement₁₂* might not have been an option for $Lefty_{11}$. For it might be that, if an L_1R_1 person had chosen this alternative at t , then $Lefty_{11}$ would not have existed at the time of choice, and so the person choosing the alternative would have been someone else. But this possibility is ruled out by P3.

Furthermore, not only could $Lefty_{11}$ have chosen the second alternative, but he would have fared better had he done so. For it follows from P2 that he would survive the semi-replacement operation, and hence that he would go on to experience 5 units of bliss after t , in addition to the 1 unit of bliss prior to t , for a total of 6 units of bliss. And this clearly exceeds the amount of bliss he

experiences in w_{11} . And so it is true, in w_{11} , that Lefty₁₁ would have fared better if he had chosen *Left-Preserving Semi-Replacement*₁₂ instead of *Fission*₁₁.

What if the third alternative, *Right-Preserving Semi-Replacement*₁₃, had been chosen instead? It follows from P4 that, if this third alternative had been chosen, then Lefty₁₁ would not have existed, and thus he would not experience any bliss. And so it must be true, in w_{11} , that the possibility in which Lefty₁₁ would fare best is the one in which he would choose the second alternative, *Left-Preserving Semi-Replacement*₁₂. For, if either of the other alternatives were chosen, then either Lefty₁₁ would not fare as well, or he would not exist at all. Hence, it follows from the Self-Interest Thesis that, in w_{11} at t' , Lefty₁₁ has self-interested reason to prefer the possibility in which he chooses the second alternative at t to either of the alternative possibilities. Thus, in w_{11} at t' , Lefty₁₁ has self-interested reason to wish he had chosen *Left-Preserving Semi-Replacement*₁₂.

But there would be something very peculiar about this choice. For let us consider what would have happened if Lefty₁₁ had chosen this second alternative. To do so, let us consider an arbitrary nearest world to w_{11} (call it w_{12}) in which Lefty₁₁'s counterpart (call him Lefty₁₂) chooses *Left-Preserving Semi-Replacement*₁₂. Note that, in w_{12} at t , Lefty₁₂ *could* have chosen the third alternative, *Right-Preserving Semi-Replacement*₁₃. For, once again, there are only two ways in which he could have failed to have this option. First, it might be that the third alternative would not have come about even if it had been decided upon by an L_1R_1 person at t . But this possibility is ruled out by the similarity between w_{12} and w_{11} , and by our stipulation that, in w_{11} , the third alternative *would* have come about given such a decision. Second, it might be that if the third alternative had been chosen at t , then Lefty₁₂ would not have existed at the time of choice. But this possibility is ruled out by P5.

Furthermore, not only *could* Lefty₁₂ have chosen the third alternative, but he would have fared *better* had he done so. For it follows from P1 that he would survive the semi-replacement operation, and hence that he would go on to experience 10 units of bliss after t , in addition to the 1 unit of bliss prior to t , for a total of 11 units of bliss. And this clearly exceeds the amount of bliss he experiences in w_{12} . And so it is true, in w_{12} , that Lefty₁₂ would fare better if he were to *Right-Preserving Semi-Replacement*₁₃ instead of *Left-Preserving Semi-Replacement*₁₂. Hence it follows from the Self-Interest Thesis that, in choosing *Left-Preserving Semi-Replacement*₁₂, Lefty₁₂ acts against his self-interest.

And yet we have already seen that, in w_{11} at t' , Lefty₁₁ has self-interested reason to wish he had chosen *Left-Preserving Semi-Replacement*₁₂. And so we now arrive at the highly counterintuitive conclusion that, in w_{11} at t' , Lefty₁₁ has self-interested reason to wish he had acted against his self-interest.

Thus, if we assume that propositions P1 through P5 are all true, then the Self-Interest Thesis will commit us to a highly counterintuitive implication. And we saw, in §2.2, that it will likewise commit us to such implications if we assume that any of these propositions are false. And so there doesn't appear to be any way to reconcile the Self-Interest Thesis with our ethical intuitions.

2.4. Indeterminacy to the Rescue?

In the arguments above, I have assumed that the metaphysical claims at issue are either true or false. Thus, in each of the *reductio* arguments in §2.2, I show that the self-interest theorist will be committed to a counterintuitive implication if she maintains that a given metaphysical claim is *true*, and I infer that she can avoid such implications only by maintaining that this metaphysical claim is *false*. But the self-interest theorist might hope to avoid such implications by maintaining that the metaphysical claim in question is *indeterminate*.²⁴ Elsewhere I discuss this kind of view in greater detail.²⁵ But here I will simply indicate, very briefly, a problem that it faces. For the sake of illustration, let's focus on the metaphysical claim at issue in *Reductio* Argument 5, namely:

5A. It is not the case that Clive₉ has a counterpart in w_{10} who is present at t .

As we have seen, assuming the Self-Interest Thesis, 5A implies the following:

5D. Clive₉ has an important kind of reason, namely self-interested reason, to prefer w_{10} to w_9 .

Therefore, assuming the Self-Interest Thesis, 5D can be determinately false only if 5A is determinately false. Hence, if the self-interest theorist maintains that 5A is not determinately false, then she must maintain that 5D is likewise not determinately false. But this is itself a counterintuitive implication. For, intuitively, it seems straightforwardly false to say that Clive₉ has self-interested reason to prefer an outcome in which his *right* half is replaced followed by *more* torture to an outcome in which his *left* half is replaced followed by *less* torture. And so, intuitively, 5D seems straightforwardly false.

The same situation arises in relation to all the *reductio* arguments in §2.2. In each case, maintaining that premise A is indeterminate will commit the self-interest theorist to maintaining that D is not determinately false. And this conclusion is itself counterintuitive.

There are, of course, further moves available to the self-interest theorist who claims that the metaphysical facts are indeterminate. I discuss and criticize these moves elsewhere.²⁶

3. The Appeal to Relation R

In this section, I will consider the view that what we have special reason to care about is not our own welfare, narrowly conceived, but rather the welfare of all those person-stages that stand in a special relation (called relation R) to our current person-stage.²⁷ I will argue that such a view will not allow us to preserve

our ethical intuitions about cases involving fission and related operations. But first, some clarificatory remarks are in order.

3.1. Some Final Stage Setting

Let's start with some terminology. By *fusion* I mean an operation that proceeds as follows. Initially, there are two qualitatively identical, perfectly divisible individuals. Then the right half of one and the left half of the other are instantaneously annihilated. And then the two remaining halves are joined to form a single individual that is qualitatively identical with each of the original individuals.

By *switcheroo* I mean an operation that proceeds as follows. Initially, there are two qualitatively identical, perfectly divisible individuals. Then the left and right halves of each of these individuals are instantaneously separated from one another. And the left half of the first individual is joined with the right half of the second, while the right half of the first is joined with the left half of the second, so as to form two individuals, each of whom is qualitatively identical with each of the original individuals.

I will say that a later person-stage is *strongly connected to* an earlier person-stage just in case:

- (i) The later person-stage retains, to a sufficient degree, the psychological features of the earlier person-stage,
- (ii) What explains (i) is that the later person-stage retains neurological structures by which the psychological features in question are realized.

I will say that two person-stages are *R-related* just in case they both belong to a sequence of person-stages such that, for every pair of consecutive person-stages in this sequence, the later person-stage is strongly connected to the earlier one. Thus, relation R is like the ancestral of the *strongly connected to* relation, except that, unlike the latter, relation R is symmetric.²⁸

For any person *S* and time *t*, by *S's R-related person-stages (at t)* I mean the person-stages that are R-related to *S's* person-stage that is located at *t*.

For any person *S* and time *t*, by *S's R-welfare (at t)* I mean the overall level welfare of *S's* R-related person-stages (at *t*). (How this overall level of welfare is to be measured will be at issue in what follows.)

Lastly, for any time *t* and any two persons, *S*₁ and *S*₂, let us say that, at *t*, *S*₂ is a *continuant* of *S*₁ just in case all of *S*₂'s person-stages are R-related to *S*₁'s person-stage at *t*.

One possible response to my arguments would be to maintain that what we have special reason to care about isn't our own, narrowly construed, personal welfare *per se*—it isn't, in other words, the welfare of the person-stages belonging to our own lives *per se*. Rather, it is our R-welfare—that is, the overall welfare

of all the person-stages that are R-related to our present person-stage. Thus, we might maintain the following:

Relation R Thesis: What we normally think of as *self-interested* reasons are best understood as *quasi-self-interested reasons*, or reasons to promote one's R-welfare. For any person S , time t , and possibilities, p_1 and p_2 , at t , S has quasi-self-interested reason to prefer p_1 to p_2 just in case S should expect that S 's R-welfare would be greater if p_1 were to obtain than if p_2 were to obtain.²⁹

I will argue, however, that moving to this thesis will not allow us to preserve our ethical intuitions. In order to show this, I will begin, in §3.2, by arguing that we cannot succeed in reconciling the Relation R Thesis with our ethical intuitions unless we accept five propositions about how R-welfare is to be measured. Then in §3.3 I will argue that, even if we accept all five propositions, we cannot succeed at this reconciliation.

3.2. Five Propositions We Must Assume to Achieve the Desired Reconciliation

In ordinary cases that don't involve any unusual operations such as fission, we have the intuition that a person has self-interested reason to promote her own personal welfare. But recall that, according to the Relation R Thesis, what we normally think of as self-interested reasons are really reasons to promote one's R-welfare. Hence, in order for the relation R theorist to reconcile her view with our ethical intuitions, she must maintain the following:

A1: In ordinary cases not involving any unusual operations, a person's R-welfare coincides with her personal welfare. Thus, for any two such ordinary outcomes, a person's R-welfare is greater in the first outcome than in the second just in case her personal welfare is greater in the first outcome than in the second.

The second assumption required in order to reconcile the Relation R Thesis with our ethical intuitions is the following:

A2: For any person S and time t , in measuring S 's R-welfare at t , *post-semi-replacement torture* (that is, to torture experienced by person-stages that are separated from S 's person-stage at t by a semi-replacement operation) is not to be discounted relative to ordinary future torture.

This assumption is motivated by the same kinds of consideration that figured in Step 1 of §2.2 (where we discussed the claim that one would survive semi-replacement). For if the Relation R theorist denies A2, then she must maintain,

counterintuitively, that one could have reason to prefer *more torture* preceded by semi-replacement to *less torture* preceded by no unusual operation.

The third assumption required for the desired reconciliation is this:

A3: For any person S and time t , in measuring S 's R-welfare at t , *post-fusion torture* is not to be discounted relative to *post-semi-replacement torture*.

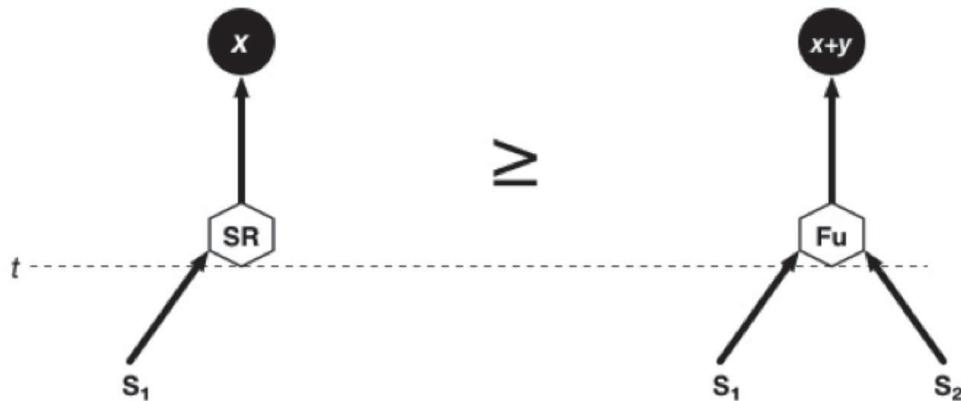
For if we accept the Relation R Thesis while denying A3, then we will get a counterintuitive result in relation to pairs of alternatives with the following form of as follows.

Solitary Torture

At t , S_1 undergoes left-preserving semi-replacement. Then whoever emerges from the operation goes on to experience x units of torture before being destroyed.

Torture Tango

At t , S_1 and her molecular duplicate S_2 undergo a fusion operation in which S_1 's original left half is connected to S_2 's original right half. Then whoever emerges from the operation goes on to experience $x + y$ units of torture before being destroyed (where $y > 0$).



It seems intuitively clear that S_1 can't have reason to prefer *Torture Tango* to *Solitary Torture*. For, in both these outcomes, S_1 's original right half is destroyed, and a duplicate of the destroyed half is connected to S_1 's original left half before the resulting whole is subjected to torture. And it hardly seems that S_1 can have reason to want this duplicate half to come from someone else (as in *Torture Tango*) rather than being newly created (as in *Solitary Torture*). And so it doesn't seem that the difference between these two operations can give S_1 reason to prefer the outcome involving more torture. But if the relation R theorist denies A2, then she will be committed to saying that, when x is sufficiently small, S_1 can have quasi-self-interested reason to prefer *Torture Tango*, since the torture it contains will have less impact on S_1 's R-welfare.

The fourth assumption required for the desired reconciliation is this:

A4: For any person S and time t , S 's R-welfare at t is never increased by the existence of person-stages experiencing nothing but torture.

If the Relation R theorist denies A4, then she must maintain, counterintuitively, that one can have quasi-self-interested reason to prefer a first outcome to a second outcome when the first can be derived from the second by the mere addition of torture-filled existence.

The fifth and final assumption required for the desired reconciliation is as follows:

A5: For any person S and time t , if S has more than one continuant at t , then *increasing* the welfare of one of these continuants while *decreasing* the welfare of the other will not increase S 's R-welfare at t , so long as the magnitude of the decrease exceeds the magnitude of the increase by a sufficient ratio.

It would be counterintuitive for the relation R theorist to deny A5. For this would require maintaining that we can have quasi-self-interested reason to make comparatively huge sacrifices to one of our continuants for the sake of comparatively tiny benefits to another such continuant.

Thus, in order to reconcile the Relation R Thesis with our ethical intuitions, we must accept A1 through A5. Let the *Standard Relation R View* (or, for brevity, the *Standard View*) be the view that combines the Relation R Thesis with these five assumptions. As we will now see, this view has a number of problematic features.

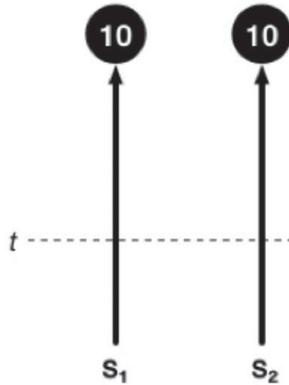
3.3. Why We're In Trouble

The first thing to note about the Standard Relation R View is that it entails the following:

SWITCHEROO WORSENS TORTURE: For sufficiently small positive values of x , two molecular duplicates, S_1 and S_2 , would each have quasi-self-interested reason to prefer *Unified Greater Torture* to *Scattered Lesser Torture*, where these outcomes are defined as follows.

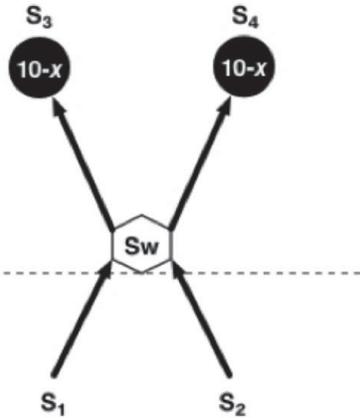
Unified Greater Torture

Beginning at t , S_1 and S_2 each experience 10 units of torture before being destroyed. Neither one undergoes any unusual operation.



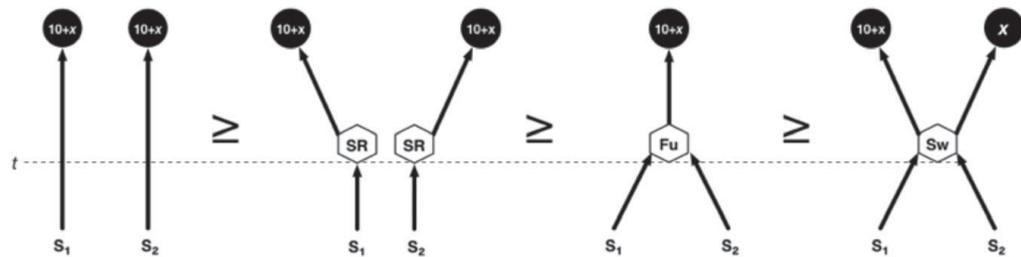
Scattered Lesser Torture

At t , S_1 and S_2 undergo a switcheroo operation from which S_3 and S_4 emerge. Then S_3 and S_4 each experience $10 - x$ units of torture before being destroyed (for some $x > 0$).



To see why this follows from the Standard View, consider the sequence of outcomes as follows.

- | O_1 | O_2 | O_3 | O_4 |
|---|---|--|--|
| Beginning at t , S_1 and S_2 each experience $10 + x$ units of torture before being destroyed. Neither one undergoes any unusual operation. | At t , S_1 and S_2 each undergo semi-replacement. Then the two individuals who emerge each experience $10 + x$ units of torture before being destroyed. | At t , S_1 and S_2 undergo fusion. Then the individual who emerges experiences $10 + x$ units of torture before being destroyed. | At t , S_1 and S_2 undergo switcheroo. Then, before being destroyed, the two individuals who emerge experience $10 + x$ units of torture and x units of torture, respectively. |



Since all these outcomes are symmetrical with respect to S_1 and S_2 , we can focus on S_1 . For any outcome O , let $W(O)$ represent S_1 's R-welfare at t in O . It follows from A1 that $W(\text{Unified Greater Torture}) > W(O_1)$. For S_1 's personal welfare is greater in the first of these outcomes than in the second, and hence, since these outcomes involve no unusual operations, it follows from A1 that S_1 's R-welfare must likewise be greater in the first outcomes. Further, it follows

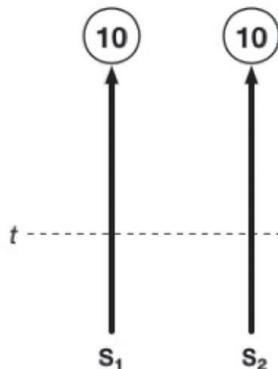
from A2 that $W(O_1) \geq W(O_2)$, since A2 implies that the post-semi-replacement torture in O_2 has no less disvalue than the ordinary torture in O_1 . It follows from A3 that $W(O_2) \geq W(O_3)$, since A3 implies that the post-fusion torture in O_3 has no less disvalue than the post-semi-replacement torture in O_2 . It follows from A4 that $W(O_3) \geq W(O_4)$, since O_4 can be derived from O_3 by the mere addition of person-stages experiencing nothing but torture. And it follows from A5 that, for sufficiently small values of x , $W(O_4) \geq W(\text{Scattered Lesser Torture})$. For, as x approaches zero, the margin by which one of the continuants is worse off in *Scattered Lesser Torture* approaches 10 units, while the margin by which the other continuant is better off in *Scattered Lesser Torture* approaches zero units. And so A5 implies that, for sufficiently small values of x , the benefit to one continuant in *Scattered Lesser Torture* can't outweigh the harm to the other. And so it follows that, when x is sufficiently small, $W(\text{Unified Greater Torture}) > W(O_1) \geq W(O_2) \geq W(O_3) \geq W(O_4) \geq W(\text{Scattered Lesser Torture})$, which entails $W(\text{Unified Greater Torture}) > W(\text{Scattered Lesser Torture})$. Thus, for sufficiently small values of x , S_1 's R-welfare at t is greater in *Unified Greater Torture* than in *Scattered Lesser Torture*. And the same goes for S_2 .³⁰

Thus, the proponent of the Standard Relation R View must accept *Switcheroo Worsens Torture*. The analogous view about bliss-filled lives is the following:

SWITCHEROO ENHANCES BLISS: For sufficiently small values of x , two molecular duplicates, S_1 and S_2 , would each have quasi-self-interested reason to prefer *Scattered Lesser Bliss* to *Unified Greater Bliss*, where these outcomes are defined as follows.

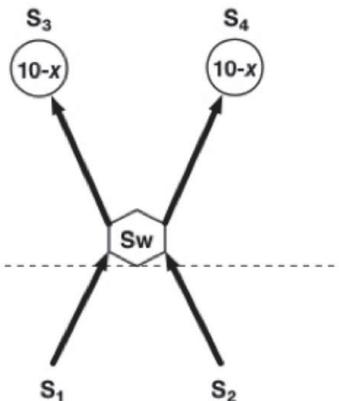
Unified Greater Bliss

Beginning at t , S_1 and S_2 experience 10 units of bliss before being destroyed. Neither one undergoes any unusual operation.



Scattered Lesser Bliss

At t , S_1 and S_2 undergo a switcheroo operation from which S_3 and S_4 emerge. Then S_3 and S_4 each experience $10 - x$ units of bliss before being destroyed (where $x > 0$).



I will now argue that the proponent of the Standard Relation R View faces a dilemma, since she will be committed to counterintuitive implications regardless of whether she accepts or rejects *Switcheroo Enhances Bliss*. Suppose, first, that she rejects this claim. If she rejects *Switcheroo Enhances Bliss* while accepting *Switcheroo Worsens Torture* (which, as we have seen, she must accept), then she must maintain that there is an important disanalogy between *post-switcheroo bliss* (that is, bliss that's separated from one's present person-stage by a switcheroo operation) and *post-switcheroo torture*. In particular, she must maintain that, in relation to someone's R-welfare at a given time, post-switcheroo bliss has less weight, relative to ordinary future bliss, than post-switcheroo torture has relative to ordinary future torture.

But such a view has the counterintuitive implication that we should evaluate a gamble differently before and after undergoing a switcheroo operation. To see why this is so, suppose that S_1 is offered a bet, and that, as soon as she makes her choice, S_1 and S_2 undergo a switcheroo operation from which S_3 and S_4 emerge. If S_1 accepts the bet, then, after the operation, a fair coin is tossed. If the coin comes up heads, then S_3 and S_4 each experience 10 units of *bliss*, but if it comes up tails, then they each experience $10 - x$ units of *torture*. (Assume the scales for bliss and torture are so defined that the value of 10 units of bliss equals the disvalue of 10 units of torture.) If post-switcheroo bliss is discounted relative to post-switcheroo torture, then, for sufficiently small values of x , S_1 will have quasi-self-interested reason to decline the gamble, since the value of the chance of bliss will be outweighed by the disvalue of the chance of torture. But if S_1 chooses to decline this gamble, then S_3 and S_4 will have quasi-self-interested reason to regret this choice. For they are not separated from the torture or the bliss by any special operation, and so, from their point of view, the value of a .5 chance of a greater amount of bliss outweighs the disvalue of a .5 chance of a lesser amount of torture. Thus, the proponent of the Standard View will be committed to saying that, at t , S_1 has quasi-self-interested reason to make a choice that is *nonratifiable*,³¹ in the sense that all the future person-stages that are R-related to the person-stage making the choice will have quasi-self-interested reason to regret this choice.

In order to avoid this result, the proponent of the Standard Relation R View must accept *Switcheroo Enhances Bliss*. But *Switcheroo Enhances Bliss* is itself highly counterintuitive. For it hardly seems that S_1 and S_2 should prefer an outcome in which they are each chopped in half, their parts are recombined, and the resulting individuals each experience less than 10 units of bliss, to an outcome in which they each remain intact and experience a full 10 units of bliss. Interchanging body parts so as to produce worse lives hardly seems like a rational choice for S_1 and S_2 to make.

What makes matters worse is that the problem iterates. Just as *Switcheroo Enhances Bliss* implies that S_1 and S_2 would have quasi-self-interested reason to trade their 10 units of bliss each for a switcheroo operation resulting in two individuals, S_3 and S_4 , who each experience *less* than 10 units of bliss, this view

likewise implies that S_3 and S_4 would in turn have quasi-self-interested reason to trade their *less than 10* units of bliss each for a switcheroo operation resulting in two individuals who each experience *still less* bliss. And this could be continued indefinitely, each time resulting in lower and lower prospects of bliss.

And there is a further problem with *Switcheroo Enhances Bliss*, that is, with the claim that, for sufficiently small values of x , S_1 and S_2 have quasi-self-interested reason to prefer *Scattered Lesser Bliss* to *Unified Greater Bliss*. For suppose they were to act on this reason and choose *Scattered Lesser Bliss*. Assuming that the two individuals who emerge from the operation, S_3 and S_4 , have counterparts in worlds in which *Unified Greater Bliss* is chosen, these individuals will have quasi-self-interested reason to regret the choice of *Scattered Lesser Bliss*. For, when S_3 and S_4 emerge from the operation, they each have $10 - x$ units of R-welfare. But any counterparts of these individuals in worlds in which *Unified Greater Bliss* is chosen would have a full 10 units of R-welfare. Hence, assuming that S_3 and S_4 have counterparts in worlds in which *Unified Greater Bliss* is chosen, they would have greater R-welfare given this choice. And so it follows that they will have quasi-self-interested reason to regret the choice of *Scattered Lesser Bliss*, in spite of the fact that this choice was favored by quasi-self-interested reasons.

We could avoid this particular instance of nonratifiability by maintaining that anyone who emerges from a switcheroo operation would never have existed if this operation had not been performed, and hence that S_3 and S_4 would not have existed if *Unified Greater Bliss* had been chosen. This view, however, will give rise to nonratifiability in relation to other choice situations, such as the choice between the alternatives as follows.

Unified Greater Torture

Beginning at t , S_1 and S_2 each experience 10 units of torture before being destroyed. Neither one undergoes any unusual operation.

Scattered Minimal Torture

At t , S_1 and S_2 undergo a switcheroo operation from which S_3 and S_4 emerge. Then S_3 and S_4 each experience 1 unit of torture before being destroyed.



Clearly, S_1 and S_2 would have quasi-self-interested reason to choose *Scattered Minimal Torture*. But if they were to make this choice, then, on the view under consideration, S_3 and S_4 would owe the existence of their torture-filled lives to the choice of *Scattered Minimal Torture*. And so they would have quasi-self-interested reason to regret this choice, in spite of the fact that it was favored by quasi-self-interested reasons.

Thus, the Relation R theorist is in a bind. For we saw in §3.2 that she will be committed to counterintuitive implications if she rejects any of the five assumptions belonging to the Standard Relation R View. And we have now seen that, if she adopts this Standard View, then she will likewise be committed to counterintuitive implications, regardless of which horn of our dilemma she may choose.

4. Conclusion

It's beginning to look like there may be no plausible way to preserve all our intuitions about what a person would have reason to prefer in cases involving fission and related operations. For we saw, in §§1–2 above, that there is no way to preserve these intuitions simply by revising our metaphysical views. And the argument of §3 demonstrates the difficulties involved in attempting to preserve these intuitions by revising the Self-Interest Thesis.

There are some who hold that we have little reason to trust our intuitions about such strange cases.³² They might claim, therefore, that, in light of the conflict between these intuitions and general ethical principles that we find compelling (such as the Self-Interest Thesis), we should reject these intuitions, while retaining all our general principles. Unfortunately, however, this is not a viable option. For each of the intuitive claims that figure in the foregoing arguments can be derived from seemingly compelling general principles. Consider, for example, the first such intuitive claim, that is, the claim that it is not the case that Clive has an important kind of reason, namely self-interested reason, to prefer *Double Torture* to *Single Torture*. This claim can be derived from the following principle:

Mere Addition Principle: For any two possibilities A and B , if A can be derived from B by the mere addition of life consisting of undeserved suffering, then no one has any important kind of reason to prefer A to B .

Thus, rejecting the intuitive claim in question would require rejecting the Mere Addition Principle. Moreover, the idea that adding undeserved suffering to the world doesn't improve things in any important respect is central to our ordinary understanding of what matters. And so the rejection of the Mere Addition Principle would involve a significant revision to our commonsense ethical worldview. And the same applies to each of the other intuitive claims that figure in the arguments I have presented. For each of these claims can likewise be

derived from seemingly compelling general principles. Hence, while it may be true that we should reject some of our intuitions about the cases in question, doing so will require us to reject or revise general principles that we find compelling.

Thus we may conclude that, even if fission poses no threat to the survival of persons, it poses a serious threat to the survival of our commonsense understanding of what matters.³³

Notes

1. I am very grateful to Kenny Easwaran, Stephen Finlay, David Manley, Jeff McMahan, Derek Parfit, Mark Schroeder, Julia Staffel, Larry Temkin, Ralph Wedgwood, J. R. Williams and Timothy Williamson for very helpful discussions and comments concerning earlier drafts of this paper. Special thanks to Lindsay McLeary for help with the diagrams. I am especially grateful to Tim Campbell and Theron Pummer for very extensive and insightful comments. My greatest debt is to Shieva Kleinschmidt for invaluable comments on countless drafts of this paper.
2. This kind of principle is defended in Sidgwick 1874, Brink 1997a, Crisp 2006, and Hare 2009.
3. See Parfit 1971, 1984, 1995. The direct target of Parfit's fission argument is not the Self-Interest Thesis *per se*, but rather the related claim that identity is what matters in survival.
4. For references to relevant works, see notes 15–18, below.
5. This motivation is at least implicit in many of the works cited in notes 15–18. For especially clear instances, see Lewis 1976 and Sider 1996.
6. This methodology traces back at least to Locke 1975 [1694] book II, chapter xvii, and has enjoyed a revival in recent decades beginning with such works as Williams 1970 and Parfit 1971. For critical discussions of this methodology, see Johnston 1987, Wilkes 1988, and Gendler 2002.
7. This kind of view is defended in Parfit 1984 and McMahan 2002. Similar, though not identical, views are defended in Jeske 1993 and Shoemaker 1999.
8. For Lewis's counterpart theory, see Lewis 1968 and 1986. For Kripke's alternative possible worlds semantics, see Kripke 1963. And for the debate between Lewis and Kripke, see Kripke 1980 and Lewis 1983.
9. This kind of view is suggested in Paterson 2010 and Keown 2012. For a critical discussions of such views, see Kagan 2012.
10. This kind of view is defended in Schopenhauer 1974 [1851] ch.xii, Mainländer 1897, and Benatar 2006.
11. This kind of view is maintained in Williams 1973, Bigelow and Pargetter 1988, and Heyd 1992. The contrary view is defended Holtug 2001, Persson 1995 and Ryberg 1995.
12. For influential defenses of this kind of view, see Feinberg 1984 and Pitcher 1984. For criticisms, see Luper 2004 and Taylor 2012.
13. See Aristotle 1999, Chap. 1.10.
14. There are, however, some philosophers who maintain that semi-replacement could not be survived. These include van Inwagen 1990 and Sidelle 2000. See Olson 1997 for discussion.

15. This view is defended in Unger 1990 and in Brink 1997b. It is also a consequence of the view defended in Nozick 1981 for the perfectly symmetrical cases we are considering. And in Parfit 1971, 1984 and 1995, it is maintained that the No Survivor View is the best view of fission cases (see, however, note 25, below).
16. This view is suggested in Chisholm 1976 and Swinburne 1984, and is defended in Johansson 2010.
17. The cohabitation theory was introduced in Lewis 1976. Similar, though not identical, views can be found in Mills 1993 and Noonan 2003.
18. The clearest example of this kind of view can be found in Sider 1996 and 2001. A similar view is defended in Gallois 1998. The views defended in Perry 1972 and Moyer 2008 are also similar in some respects, and can plausibly be read as having the same practical implications in the cases we will be considering.
19. Here I am assuming that Clive has at least one counterpart in worlds in which *Stochastic Double Torture* is chosen. But we will arrive at the same conclusion if we suppose instead that he has no counterparts in such worlds. For, on this supposition, Clive would not exist if *Stochastic Double Torture* were chosen, and so he would not experience any torture whatsoever. Thus, on either supposition, Clive would not fare as badly if *Stochastic Double Torture* were chosen. (The kind of comparison that figures in this note will be discussed in greater detail in §2.)
20. For an alternative way of thinking about the practical implications of the Double Identity View, which supports the same conclusion about the present case, see Tappenden 2011.
21. This example is adapted from Parfit 1984, §64.
22. For discussions of these issues, see Hare 2008 and Brink 2010.
23. Here, and throughout this section, I'm assuming that there *is* someone who emerges from the fission operation consisting of L_1 and R_2 , and that, after t , this person experiences only the torture belonging to the human organism consisting of L_1 and R_2 . This assumption might be denied—either by maintaining that there is *no* person who emerges from the fission operation, or else by maintaining that everyone who emerges from the operation experiences all the torture belonging to either of the two human organisms that emerge from the operation (this latter view is consistent with the multilocation account of fission defended in Dainton 1992 and Wright 2006). Denying this assumption, however, would commit the self-interest theorist to further counterintuitive implications. To see why this is so, consider two worlds, w_C and w_D , in which, at t , the following alternatives are chosen, respectively, by a human organism consisting of L_1 and R_1 : **Fission_C**: Immediately after t , the organism consisting of L_1 and R_1 undergoes a fission operation in which L_1 is connected to R_2 and R_1 is connected to L_2 . Then the two human organisms that emerge each experience 6 units of torture before being destroyed. **Fission_D**: Just like **Fission_C**, except that, after the operation, the human organism consisting of L_1 and R_2 experiences 10 units of torture before being destroyed, whereas the human organism consisting of L_2 and R_1 experiences 1 unit of torture before being destroyed. It seems intuitively clear that not everyone who emerges from the operation in w_C has self-interested reason to prefer w_D to w_C . But if *no one* emerges from the fission, or if everyone who emerges experiences all the torture belonging to

- either of the resulting human organisms, then everyone who emerges from the operation in w_C fares worse in w_C than in w_D . And so it will follow, from the Self-Interest Thesis, that everyone who emerges from the operation in w_C has self-interested reason to prefer w_D , contrary to intuition.
24. The idea that it's indeterminate what happens to someone who undergoes an operation such as fission is suggested in Parfit 1984 and defended in Johnston 1989. J. R. G. Williams discusses the practical significance of indeterminacy in a number of works, such as Williams 2014.
 25. Ross (unpublished manuscript).
 26. See Ross (unpublished manuscript). One move available to the self-interest theorist is to maintain that *indeterminacy is to be treated like risk*. I argue that such a move would result in counterintuitive implications in variants of the cases discussed in §2.2. Another move would be to maintain that one has self-interested reason to prefer one outcome to another *only when it's determinately true* that one would fare better in the first outcome. I argue that such a move would have counterintuitive implications in sorites-like cases.
 27. For references to some relevant literature, see note 7, above.
 28. The idea of relation R is borrowed from Lewis 1976 and Parfit 1984, and traces back to Russell 1956 [1918].
 29. In order to allow for the view that our concern for future person-stages should trump our concern for past person-stages, we could add the condition that S must expect that her *future* R-related person stages would fare better overall given p_1 than given p_2 , or at least that S must *not* expect that these future person-stages would fare *worse* overall given p_1 than given p_2 . These additional conditions are satisfied in all the examples I will be considering below.
 30. One might attempt to block this argument by maintaining that, in some of these outcomes, S_1 's R-welfare would be *indeterminate*. Such a response, however, would face difficulties similar to those we saw in §2.4.
 31. I borrow this term from Jeffrey 1981.
 32. See, for example, Gendler 2002.
 33. Several other very recent works lend support to the claim that fission cases have more revisionary ethical implications than had previously been recognized. Notable examples include Pummer (2014), Briggs and Nolan (forthcoming) and Campbell (n.d.).

References

- Aristotle. 1999. *Nicomachean Ethics*. Translated by Terence Irwin. Indianapolis: Hackett Publishing Company.
- Benatar, David. 2006. *Better Never to Have Been*. Oxford: Oxford University Press.
- Bigelow, John and Robert Pargetter. 1988. "Morality, Potential Persons and Abortion." *American Philosophical Quarterly* 25: 173–181.
- Briggs, Rachel and Daniel Nolan. Forthcoming. "Utility Monsters for the Fission Age." Forthcoming in *Pacific Philosophical Quarterly*.
- Brink, David. 1997a. "Rational Egoism and the Separateness of Persons." In *Reading Parfit*, edited by Jonathan Dancy, 96–134. Oxford: Blackwell.
- . 1997b. "Self-Love and Altruism." *Social Philosophy and Policy* 14: 122–157.

- . 2010. “Prospects for Temporal Neutrality.” In *The Oxford Handbook of Time*, edited by Craig Callender, 353–381.
- Campbell, Timothy. n.d. “Personal Identity and Value Aggregation.” Unpublished manuscript.
- Chisholm, Roderick. 1976. *Person and Object: A Metaphysical Study*. La Salle, IL: Open Court.
- Crisp, Roger. 2006. *Reasons and the Good*. Oxford: Oxford University Press.
- Dainton, Barry. 1992. “Time and Division.” *Ratio* 2: 102–128.
- Feinberg, Joel. 1984. *Harm to Others: The Moral Limits of the Criminal Law*. New York: Oxford University Press.
- Gallois, André. 1998. *Occasions of Identity*. Oxford: Oxford University Press.
- Gendler, Tamar S. 2002. “Personal Identity and Thought Experiments.” *Philosophical Quarterly* 52: 34–54.
- Hare, Caspar. 2008. “A Puzzle About Other-Directed Time Bias.” *Australasian Journal of Philosophy* 86: 269–277.
- . 2009. *On Myself, and Other, Less Important Subjects*. Princeton: Princeton University Press.
- Heyd, David. 1992. *Genethics: Moral Issues in the Creation of People*. Berkeley: University of California Press.
- Holtug, Nils. 2001. “On the Value of Coming into Existence.” *Journal of Ethics* 5: 361–384.
- Jeffrey, Richard. 1981. “The Logic of Decision Defended.” *Synthese* 48: 473–492.
- Jeske, Diane. 1993. “Persons, Compensation, and Utilitarianism.” *The Philosophical Review* 102: 541–575.
- Johansson, Jens. 2010. “Parfit on Fission.” *Philosophical Studies* 150: 21–35.
- Johnston, Mark. 1987. “Human Beings.” *Journal of Philosophy* 84: 59–83.
- . 1989. “Fission and the Facts.” *Philosophical Perspectives* 3: 369–97.
- Kagan, Shelly. 2012. *Death*. New Haven: Yale University Press.
- Keown, John. 2012. *The Law and Ethics of Medicine: Essays on the Inviolability of Human Life*. Oxford: Oxford University Press.
- Kripke, Saul. 1963. “Semantical Considerations on Modal Logic.” *Acta Philosophica Fennica* 16: 83–94.
- . 1980. *Naming and Necessity*. Cambridge: Harvard University Press.
- Lewis, David. 1968. “Counterpart Theory and Quantified Modal Logic.” *Journal of Philosophy* 65: 113–26.
- . 1976. “Survival and Identity.” In *The Identities of Persons*, edited by Amélie O. Rorty, 17–40. Berkeley: University of California Press.
- . 1983. “Postscripts to ‘Counterpart Theory and Quantified Modal Logic.’” In *Philosophical Papers: Volume 1*, 39–46. Oxford: Oxford University Press.
- . 1986. *On the Plurality of Worlds*. Oxford: Blackwell.
- Locke, John. 1975 [1694]. *An Essay concerning Human Understanding*, ed. Peter Niddich. Oxford: Oxford University Press.
- Luper, Steven. 2004. “Posthumous Harm.” *American Philosophical Quarterly* 41: 63–72.
- Mainländer, Philipp. 1897. *Philosophie der Erlösung*. Berlin: Verlag von Theodor Hofmann.
- McMahan, Jeff. 2002. *The Ethics of Killing: Problems at the Margins of Life*. New York: Oxford University Press.
- Mills, Eugene. 1993. “Dividing Without Reducing: Bodily Fission and Personal Identity.” *Mind* 102: 37–51.
- Moyer, Mark. 2008. “A Survival Guide to Fission.” *Philosophical Studies* 141: 299–322.
- Noonan, Harold. 2003. *Personal Identity*. 2nd ed. London: Routledge.
- Nozick, Robert. 1981. *Philosophical Explanations*. New Haven: Harvard University Press.
- Olson, Eric T. 1997. *The Human Animal: Personal Identity Without Psychology*. New York: Oxford University Press.
- Parfit, Derek. 1971. “Personal Identity.” *Philosophical Review* 80: 3–27.
- . 1984. *Reasons and Persons*. Oxford: Oxford University Press.

- . 1995. "The Unimportance of Identity." In *Identity*, edited by Henry Harris, 13–45. Oxford: Oxford University Press.
- Paterson, Craig. 2010. *The Contribution of Natural Law Theory to Moral and Legal Debate Concerning Suicide, Assisted Suicide, and Euthanasia*. Los Angeles: Viewforth Press.
- Perry, John. 1972. "Can the Self Divide?" *Journal of Philosophy* 69: 463–488.
- Persson, Ingmar. 1995. "Genetic Therapy, Identity and Person-Regarding Reasons." *Bioethics* 9: 17–25.
- Pitcher, George. 1984. "The Misfortunes of the Dead." *American Philosophical Quarterly* 21: 217–225.
- Pummer, Theron. 2014. "Does Division Multiply Desert?" *Philosophical Review* 123: 43–77.
- Ross, Jacob. Unpublished Manuscript. "Personal Identity, Indeterminacy, and Self-Interest." Available at www-bcf.usc.edu/~jacobmro/
- Russell, Bertrand. 1918. "The Philosophy of Logical Atomism." *Monist*, 28, 495–527; 29, 32–63, 190–222, 345–380; repr. in Bertrand Russell, 1956, *Logic and Knowledge*, London: Allen and Unwin.
- Ryberg, Jesper. 1995. "Do Possible People have Moral Standing?" *Danish Yearbook of Philosophy* 30: 96–118.
- Schopenhauer, Arthur. 1974 [1851]. *Parerga and Paralipomena: Short Philosophical Essays*. Translated by E. F. J. Payne. Oxford: Oxford University Press.
- Shoemaker, David W. 1999. "Selves and Moral Units." *Pacific Philosophical Quarterly* 80: 391–419.
- Sidelle, Alan. 2000. "Finding and Intrinsic Account of Identity: What Is the Source of Transplant Cases?" *Philosophy and Phenomenological Research* 61: 415–430.
- Sider, Theodore. 1996. "All the World's a Stage." *Australasian Journal of Philosophy* 74: 433–453.
- . 2001. *Four Dimensionalism*. Oxford: Oxford University Press.
- Sidgwick, Henry. 1874. *The Methods of Ethics*. London: Macmillan.
- Swinburne, Richard. 1984. "Personal Identity: The Dualist Theory." In *Personal Identity*, edited by Sydney Shoemaker and Richard Swinburne, 1–66. Oxford: Basil Blackwell.
- Tappenden, Paul. 2011. "Expectancy and Rational Action Prior to Personal Fission." *Philosophical Studies* 153: 299–306.
- Taylor, James Stacey. 2012. *Death, Posthumous Harm, and Bioethics*. London: Routledge.
- Unger, Peter. 1990. *Identity, Consciousness and Value*. New York: Oxford University Press.
- van Inwagen, Peter. 1990. *Material Beings*. Ithaca: Cornell University Press.
- Wilkes, Kathleen V. 1988. *Real People: Personal Identity without Thought Experiments*, Oxford: Clarendon Press.
- Williams, Bernard. 1970. "The Self and the Future." *Philosophical Review* 79: 161–180.
- . 1973. "The Makropulos Case: Reflections on the Tedium of Immortality." In *Problems of the Self*, 82–100. Cambridge: Cambridge University Press.
- Williams, J. Robert G. 2014. "Decision-Making Under Indeterminacy." *Philosophers' Imprint* 14: 1–34.
- Wright, John. 2006. "Personal Identity, Fission and Time Travel." *Philosophia* 34 (2):129–142.