

Morality Grounds Personal Identity

Bradley Monton

There is a connection between moral facts and personal identity facts: morality grounds personal identity. If, for example, old Sally enters a teletransporter, and new Sally emerges, the fundamental question to ask is: is new Sally morally responsible for actions (and omissions) of old Sally? If the moral facts are such that she is morally responsible, then Sally persisted through the teletransporter event, and if not, Sally ceased to exist.

Keywords personal identity, meta-personal identity, grounding, mereological composition, person, conventionalism

1. Morality and Metaidentity

This is an essay in *meta-personal identity*. Just as, in ethics, we distinguish between applied ethics, normative ethics, and metaethics, so we can make similar distinctions in the field of personal identity:

Applied personal identity: did this particular person who just entered this particular teletransporter just die, or does she continue to exist, over at the exit end of the teletransporter?

Personal identity: Is it just psychological continuity and connectedness that matters for persistence of persons, or is it just physical continuity, or is it some combination thereof, or is it persistence of souls, or is it something else?

Meta-personal identity: Is there an objective fact of the matter about whether a person does or does not persist, or is it somehow subjective or relative or conventional? If there are objective facts about persistence of persons, what (if anything) *grounds* those facts? If there are no such objective facts, how should we understand everyday utterances that seem to make presuppositions about personal identity (as in: “Fred went home, but he’ll be back”)?

“Meta-personal identity” doesn’t roll off the tongue, alas, so I’ll call this “metaidentity” for short. We have to remember, though, that we’re not talking about the standard sort of identity (as in: Superman is Clark Kent). Instead, we’re talking about *personal identity*, which (by my lights, at least) isn’t really about *identity* at all. What personal identity is about is *persistence conditions*: under what conditions does this person survive, and hence persist through time, and under what conditions does this person cease to exist, as a person?

Suppose that Sally undergoes a standard personal identity thought-experiment: brain transplant, teletransportation, memory erasure,

ship of Theseus with neurons, what have you. Call “old Sally” the person before she undergoes the experiment, and “new Sally” the person who emerges. (Later we’ll take up cases where more than one person emerges.) Here are two types of questions we can ask:

- (1) Is new Sally the same person as old Sally? In other words, did Sally survive the experiment, or did she cease to exist, as a person?
- (2) Suppose that old Sally did something morally wrong just before the experiment. Is new Sally morally responsible for what old Sally did? (Or, suppose old Sally did something morally right just before the experiment; is new Sally morally commendable?)

Part of the point of this paper is to argue that there is a close connection between these two questions. But which way does the connection go? One could maintain that the personal identity question is the fundamental one, and that the answer to that question will guide us in attributing moral responsibility. This is the dominant view in the literature. Here for example, is David Shoemaker endorsing that view:

In order to attribute moral responsibility to someone for an act, we must be able to determine that that person is the same person as the person who performed the act. ... What is needed first is a plausible metaphysical account of persons and personal identity to which an ethical theory might then conform and apply. (Shoemaker 1999, 183)

For Shoemaker, the metaphysical account of personal identity is, conceptually, what comes first - the personal identity facts then get inputted into the ethical theory.

On the other hand, one could maintain that the moral responsibility question is the fundamental one, and the answer to that question grounds

the facts about personal identity (where the grounding relation is understood in the now-standard way elucidated by for example Gideon Rosen (2010)). This is the view I hold; this is what I'll be arguing for in this paper.

2. Three Assumptions, and an Argument

Let's start with three big assumptions. These are assumptions that I believe, but I won't argue for them here. In this section I'll give my argument for the claim that morality grounds personal identity using these assumptions, and later I'll talk about how my position would change when these assumptions are dropped.

The first assumption is that there are objective moral truths, independent of what we believe about morality. Though it doesn't really matter for the purposes of this paper, the sort of moral realism I endorse is non-natural and non-theistic - the moral truths are not reducible to natural facts and properties, and are not true as a result of the existence (and perhaps choices) of some divine being. Instead, moral truths have the status of other objective truths that are (contentiously) non-natural: modal truths, mathematical truths, and logical truths. This sort of moral realism has been argued for well by for example Michael Huemer (2005) and Erik Wielenberg (2009).

The second and third assumptions naturally go together. The second assumption is that the perdurance theory of persistence is true: objects persist by having temporal parts. A persisting object (in standard spacetime, at least) is a four-dimensional spacetime worm, and this four-dimensional object is composed of instantaneous temporal parts, where each instantaneous temporal part is the object at a particular time

(modulo time travel stories). (Some philosophers distinguish between the “stage view” and the “worm view”, where the stage view holds that a person is a person-stage, existing at a particular time, while the worm view holds that a person is the whole temporally extended spacetime worm. I don’t think there’s a metaphysical difference between these two views; sometimes, we refer to person-stages (“Jill is sitting”), while other times, we refer to the whole worm (“Jack had a wonderful life”).)

The third assumption is that the doctrine of unrestricted mereological composition is true. For every collection of objects, there is a whole that is composed of those objects as parts. So, for example, there is the 24-hour-long four-dimensional spacetime worm of Obama on July 1 2011, which can be combined with the 24-hour-long four-dimensional spacetime worm of Cheney on January 1 2012, to form a particular object, Obacheny, spread in space and time in a complex, disjoint way.

Given the latter two assumptions, one might think that there is no interesting debate about personal identity to be had. When faced with a personal identity thought experiment, there is an object that exists both before and after the experiment (the mereological sum of old Sally and new Sally, for example), and there is an object that ceases to exist at the time of the experiment (old Sally, for example), and there are an infinite number of other objects (the mereological sum of old Sally, the first two minutes of new Sally, and Obacheny, for example). All these objects exist, according to the doctrine of unrestricted mereological composition, and so it’s prima facie unclear how to answer the question of whether Sally persists through the experiment - which of the infinite number of mereological sums of objects are we talking about when we talk about Sally? I maintain that there is a definite answer, and the objective moral facts determine what the definite answer is.

I'll give my whole argument now, and then provide some elaboration. The first step in my argument is to point out that we're talking about *personal* identity, so we're talking about the persistence of *persons*. Suppose that John goes into a permanent vegetative state. John persists as a biological organism, but John (arguably) ceases to exist as a person. Suppose that John later dies and is buried. We make everyday utterances like "John is buried here", and I (perhaps controversially) believe that such utterances are literally true (on a certain interpretation, at least). John does not persist as a person, and does not persist as a biological organism, but does persist as a mereological sum of particles that composed the biological organism pre-death. Similarly, "we are stardust" is also a literally true claim, given a certain interpretation. There are of course ways of interpreting the claim such that it comes out false, but a charitable interpretation picks up on the fact that the utterer is talking, not about the persistence of persons, or of biological organisms, but the persistence of constituting atoms. Since these atoms were formed by fusion reactions in stars, then there's a sense in which it's true that our current temporal parts are mereological sums of stardust.

But even if you disagree with that particular controversial claim, my overall point is that, when focusing on the question of personal identity, we are interested in the persistence conditions for persons, not biological organisms or atoms. But what is a person?

This brings me to the second step in my argument: I hold that the concept of a person is a *moral* concept. First, a preliminary point: we all agree that persons have moral worth, but some non-persons have moral worth as well - it's wrong to torture a cow, even though a cow is not a person. But my main point is that persons, in addition to having moral worth, are (at least rudimentary) *moral agents*. Persons are morally

responsible for the bad things they do, and morally commendable for the good things they do, and this is part of what is involved in being a person – it is to be the sort of agent to whom moral responsibility and commendability is to be attributed.

Without moral facts, it's unclear what does ground personal identity facts. Consider a full specification of what happens in some personal identity thought-experiment. We can specify where every particle goes, we can specify what the psychological traits are of every temporal part of persons involved, and we can specify how souls are persisting, if indeed souls exist. Given all these facts, which mereological sums of temporal parts constitute a single persisting person? One still has to know (for example) that the persistence of consciousness is what establishes the persistence of a person (as for example Locke thought), or that the persistence of the soul establishes the persistence of the person (Locke famously thought otherwise). But what grounds those facts?

Now, it could just be a matter of our individual or communal practices, or our arbitrary semantic conventions, that “person” refers to a certain type of entity. I’ll consider (and reject) those lines of thought in the next section. But my position is that the fact that some temporal parts of entities have objective moral responsibility relations to other temporal parts enables us to pick out a certain privileged mereological sum of temporal parts, and that is the entity that should be thought of as a persisting person.

That, in a nutshell, is my argument. But before taking up the alternative lines of thought in the next section, there are four points I want to make.

First, I’ve said that I’m interested in the moral responsibility relation, and that this grounds personal identity. Specifically, what I’m interested

in is the ancestral of the moral responsibility relation. Just as there are legal statues of limitations, so there may well be moral ones. Eighty-year-old Sarah is not morally responsible for three-year-old Sarah's gum-stealing, but that is not a problem, on my view. As long as temporal part $\#n$ of Sarah is morally responsible for something that temporal part $\#n - 1$ did, and temporal part $\#n - 1$ of Sarah is morally responsible for something that temporal part $\#n - 2$ did, and so on, then all these temporal parts are parts of the same person.

Here's my second point. I take it that one can be morally responsible, not just for bad things one does, but for good things as well. The personal identity relation isn't just grounded in the morally wrong things people do. Even if a person never did anything morally wrong, the person is nevertheless morally commendable for good things she did in the immediate past, and that chain of moral commendability is enough to ground personal identity. One might ask: what if a person has a morally neutral day, such that she doesn't do anything morally bad or good - how can my view make sense of her continuance as a person throughout that day? I hold that persons never have morally neutral days; they are constantly doing morally good (or bad) things. The reason is that one is (at least minimally) morally commendable for *omitting* to do bad actions. So, a temporal part of a person that sits quietly has omitted from kicking anyone in the shin, and hence immediately subsequent temporal parts of the person are (at least minimally) morally commendable for that omission.

My third point is that there's an epistemological issue lurking in this whole discussion, which I want to acknowledge just to set aside. How is it that we can *know* whether a person persists through time? In other words, how can we know the answer to these moral questions?

I don't have any special insight to offer regarding how to answer these questions. In many situations, the personal identity facts are just obvious to us: it's obvious that the temporal part of Obama who got elected president in 2008 is part of the same person as the temporal part of Obama who got reelected in 2012. (And it's also the case, once one thinks about it, that these obvious-seeming claims could be false, due to skeptical worries. Perhaps Obama was abducted by aliens in 2010, and was replaced by a qualitative duplicate.) Similarly, some moral facts are just obvious to us (barring philosophical worries about the status of moral truths, an issue I'll come back to below). So in some cases, it's obvious that a person persists, just as it's obvious that a person is morally responsible for something she did.

But there are other cases where it's not at all obvious whether a person persists, just as it's not at all obvious that a current temporal part of a person is morally responsible for something that a previous temporal part of a person did. In these situations, there's nothing more we can do than to use our standard methods: we do careful philosophical reasoning, we try to reflect in a rational, unbiased way regarding how things morally seem to us, and so on. My point is that there is an objective fact of the matter about whether a temporal part of a person is morally responsible for something that another temporal part did, and that's enough to metaphysically establish whether those temporal parts are parts of the same person, even in situations where it's epistemically hard for us to figure out which moral facts, and hence which personal identity facts, hold.

My fourth and final point addresses the question: to what extent is my position, that morality grounds personal identity, original? Well, I see foreshadowing in Locke, but my reading of Locke is controversial. (I

have a lot to say about different ways of interpreting Locke on metaidentity, and how my theory of metaidentity relates, but that is best saved for another paper.) The only philosophical work I've found that comes somewhat close to my view is a paper by Eric Wiland (2000). Wiland argues that our theory of moral responsibility should influence our theory of personal identity; to that extent we are in agreement. But the details of his view, and his corresponding argument, are quite different. For example, he gives a partial characterization of the personal identity relation, by giving two necessary criteria for whether a person persists. One is that, for Y to be the same person as X, Y must be bodily continuous with X. I wouldn't want to build bodily continuity in as a necessary criterion (though it may end up being true, because of the objective moral facts being what they are). The other necessary criterion that he gives for Y to be the same person as X is that "Y is quasi-responsible for some action of X's" (2000, 84). Unfortunately, Wiland does not give a precise definition of quasi-responsibility, and his discussion introducing the concept leaves me nonplussed.

3. Conventionalism and the Concept of a Person

I maintain that facts about persistence of persons are grounded in objective moral facts. But a contrasting popular view in the metaidentity literature is *conventionalism*: facts about persistence of persons just hold as a result of decisions that are made by an individual or a community. Consider, for example, David Braddon-Mitchell and Caroline West's endorsement of conventionalism:

To survive, on our understanding, is to preserve whatever property a person's

(or perhaps community's) person-directed practices are organized around.
(Braddon-Mitchell and West 2001, 61)

For conventionalists, whether a person persists depends on what person-directed practices are followed by a person (or the community the person is a part of). So whether, for example, you survive teletransportation depends on whether you (or your community) holds funerals for people who enter teletransportation machines, or treats such machines as a normal mode of transportation.

I believe that conventionalism holds for everyday created objects, like ships. The question of whether a ship persists is, indeed, purely grounded in our practices that reveal what it takes to be a persisting ship.

But are persons relevantly like ships? I maintain that conventionalism for persons is a highly implausible position, precisely because we do not take facts about our survival to depend on what practices we choose to adopt. Facts about our survival are, intuitively at least, objective facts. A virtue of my view is that it can accommodate that. Given that the concept of a person is a moral concept, and that there are objective moral truths, then I maintain that the most plausible metaphysical view is that there are objective facts of the matter about whether a person persists, and whether a person persists is determined by the objective moral facts.

To further develop my point that the concept of a person is a moral concept, let's bring in a distinction that Matti Eklund makes in his underappreciated 2004 paper on personal identity. Eklund distinguishes between the "moral question" and the "semantic question" of personal identity. The moral question is:

what is the nature of the entities we should focus our prudential concerns and ascriptions of responsibility around? (Eklund 2004, 489)

The semantic question is:

what is the nature of the entities that 'person' is true of? (Eklund 2004, 489)

Regarding the moral question, Eklund recognizes that, in principle, the entities we should focus our prudential concerns around could be different from the entities we should focus our ascriptions of responsibility around. I maintain that ascriptions of responsibility are the focus of personal identity, and these ascriptions of responsibility should track the objective moral truths. Our prudential concerns have no such objective focus. (For example, I could be more prudentially concerned about my friend than myself when I risk my life to save her, and in choosing to be prudentially concerned in that way, I could be doing so in such a way that I don't violate any objective dictates of prudence.) Thus, the aspect of Eklund's moral question that I'm most interested in is:

what is the nature of the entities we should focus our ascriptions of responsibility around?

Since we would want our ascriptions of responsibility to track the objective moral truths about responsibility, this question has the desired focus, linking the concept of a person to the objective moral truths.

But what about the semantic question - how is it that we actually use the concept 'person'? Eklund argues that it's not clear whether the answers to the moral question and the semantic question are the same. But even if they aren't the same, Eklund points out that the semantic question ends up looking rather unimportant. His paper ends with the following:

Suppose we have answered the moral question: we have figured out what person-like entities (if any) we should structure our ... ascriptions of responsibility around. ... Still, we have not yet answered the semantic question: in order to do so we must take a close look at what our conception of persons is. However, how important is what is missing? ... All we would not know is which entities happen to be picked out by a particular concept of ours. The significance of this question seems to pale in comparison to the others. (Eklund 2004, 507-8)

I agree with Eklund that, if the moral question potentially has a different answer than the semantic question, then the moral question is the important one. In fact, the semantic question looks so unimportant when divorced from the moral question that I maintain that it is a mistake to so divorce it. Of course, different people understand a concept like 'person' in different ways, and there is debate about what exactly the concept amounts to. But a way to help resolve this debate is to make clear that what's implicit in our concept of a person are attributions of moral responsibility, and hence, the answer to the semantic question is the answer to the moral question.

That concludes my discussion of the semantic question and the moral question of personal identity. But there's one more thing I want to say about Eklund. Eklund argues that our concept of a person is indeterminate, because "our conception of what persons are is not such as to decide what to say with respect to some of the problem cases discussed in the literature" (Eklund 2004, 490). I would argue instead that the apparent indeterminacy arises, not because we don't have a full account of the concept of a person, but because we don't have agreed-upon answers to some of the problematic moral responsibility questions that can be asked in the context of a personal identity problem

case. The concept of a person tracks moral responsibility; it's because there is no agreement on whether, for example, new Sally emerging from the teletransporter is morally responsible for something that old Sally did that one might mistakenly think there is indeterminacy in the concept of a person.

I have one final point to make about the concept of a person, before moving on. There are various definitions of the concept of a person given in the literature (in the context of the abortion debate, for example); how does my moral-responsibility-based conception of a person relate to those definitions? Consider, for example, Michael Tooley's account:

Something is a person if and only if it is a continuing subject of experiences and other mental states that can envisage a future for itself and that can have desires about its own future states. (Tooley 1979, 91)

Definitions like this are compatible with my account of persons; they provide an answer to the moral question "what is the nature of the entities we should focus our ascriptions of responsibility around?" Based on Tooley's definition, the answer would be:

The entities we should focus our ascriptions of responsibility around are the continuing subjects of experiences and other mental states that can envisage futures for themselves and that can have desires about their own future states.

Given that there are objective moral truths, and that our ascriptions of responsibility should track the objective moral truths, there is an objective fact of the matter regarding whether this answer is the right one. If the objective moral ascriptions of responsibility apply to all and only the continuing subjects of experiences and other mental states that can

envisage futures for themselves and that can have desires about their own future states, then Tooley's account would be a correct account of the concept of a person. Thus, an account of personhood like Tooley's is not incompatible with mine; we're just engaging in different levels of analysis.

4. Dropping the Assumptions

My argument above was based on three assumptions; let's now consider what happens if we drop the assumptions. I'll start with the first assumption, that there are objective moral truths. There are three standard moral anti-realist positions one could endorse, relativism, nihilism, and non-cognitivism; I'll talk about how my metaidentity position would fare on each one.

According to moral relativism, moral facts are relative to an individual, or community (or something along those lines). Given my position that the moral facts establish the personal identity facts, it would follow that personal identity facts are similarly relative to an individual or community - the conventionalist position discussed above. Those who are inclined toward moral relativism may well be satisfied with this corresponding conventionalism in personal identity situations. Metaidentity conventionalism is typically argued for without bringing in moral relativism; I've presented a new line of argument for metaidentity conventionalism that moral relativists could give.

According to moral nihilism, all positive moral claims are false (because they are attributing moral properties, and moral properties don't exist). If we take this view that there are no positive moral facts, and apply it to my position that morality grounds personal identity, then we

get the view that there are no positive personal identity facts. The claim that Sally survives the teletransporter is false – and indeed, all claims that a person persists are false.

According to non-cognitivism, moral utterances don't have truth-conditions – they are, for example, like expressions of approval or disapproval, or expressions of emotion. One way to apply that to personal identity is as follows. Suppose that Fred utters: “Sally survives the teletransporter”. This is Fred expressing his desire to treat new Sally as if she is the same person as old Sally. Fred could be an enlightened non-cognitivist, and endorse the connection between morality and personal identity, and hence Fred could realize that there is no fact of the matter regarding whether Sally survives. In uttering “Sally survives”, Fred is simply expressing his desire to treat new Sally the same sort of way he treated old Sally. (For example, if Fred was married to old Sally when she entered the teletransporter, he is expressing his desire to continue to treat new Sally as his partner.)

To sum up: because of the link I'm endorsing between morality and personal identity, the different metaethical positions correspond to different metaidentity positions. All these metaidentity positions are *prima facie* live options, though my preferred one is the one that is based on moral realism.

Let's now turn to the second and third assumptions: that the perdurance theory of persistence is true, and that the doctrine of unrestricted mereological composition is true. Since these two are linked, let's drop them together: suppose that the endurance theory of persistence is true (an object is wholly present at every time that it exists), and that collections of objects don't always compose a whole.

The most radical way to give up unrestricted mereological composition

is to be a mereological nihilist. In that situation, persons wouldn't persist at all - unless they are metaphysical simples (a view defended by for example David Barnett (2013)). Let's suppose that the correct replacement for unrestricted mereological composition is such that it does allow persons to persist. But what grounds the fact that some particular person, Sally, continues to exist?

My answer, recall, is that all collections of temporal parts exist, and one of them is picked out as morally significant, and is identified as the persisting person. But the endurance theorist has to say that it's a metaphysical fact of the matter that a particular thing, Sally, continues to exist. It's unclear to me what (if anything) grounds those metaphysical facts. If nothing grounds those metaphysical facts, how do we have epistemic access to them? (We arguably have epistemic access to the moral facts via our faculty of rational intuition - do we have access to the ungrounded metaphysical facts about endurance via similar means?) And if something does ground those metaphysical facts about endurance, what is the ground? The endurance theorist is welcome to take on board my answer, that the moral facts are the ground. But if the endurance theorist does not make that move, then I find it mysterious what (if anything) is the ground for these metaphysical facts about whether a person continues to be wholly present, and how we could have access to these metaphysical facts.

Some endurance theorists argue that there is no interesting debate about personal identity to be had once one assumes the truth of the perdurance theory and unrestricted mereological composition. Eric Olson, for example, holds that "if material objects are temporally extended, then there are no substantive metaphysical problems about our identity through time, but only semantic questions about how the *language* of personal

identity works” (Olson 1997, 5). (Later in his book, Olson says that another assumption is needed to get his conclusion that personal identity questions become semantic questions; this assumption is something like the assumption of unrestricted mereological composition: “that every matter-filled region of spacetime contains an object” with the extension of that region (Olson 1997, 162).)

Let’s consider Olson’s claim that, given my perdurance and unrestricted mereological composition assumptions, personal identity questions are just semantic questions. Olson’s thought is that, if all the mereological sums of temporal parts are real, then it’s just a semantic question of which ones count as being the referent of our concept ‘person’. But what Olson is missing is that the objective moral facts can be brought in to ground our ascriptions of which mereological sums count as persisting people. It’s true that there’s a residual semantic question, but that holds for all metaphysics. If we redefine “God” as “love”, then I’m no longer an atheist, because I believe in love. But the concept of God is a metaphysically important concept, as is the concept of a person. Because the concept is metaphysically important, then picking out which entities count as persisting people is metaphysically important; it’s not just a semantic issue. And my key point is that one can hold that the concept of a person is metaphysically important even if one endorses the perdurance theory and unrestricted mereological composition, by recognizing the connection between personhood and the objective moral truths. (If moral anti-realism is true, then while the connection between personhood and morality is still there, the concept of a person is presumably less metaphysically important, because morality itself doesn’t have an objective metaphysical status.)

5. Four Objections

I'll close out this paper by considering four objections to my argument.

5.1 Objection #1: Fetuses are People Too

Some philosophers hold that a (human) fetus is a person. But fetuses aren't morally responsible (setting implausible claims about original sin aside). Hence, a fetus is a persisting person without having moral responsibility relations between its temporal parts, and hence, my view of persistence of persons is false.

It's true that my theory of metaidentity is incompatible with some other views about what counts as a person and who is morally responsible. But that is not problematic, because those other views are false. A fetus is not a person, nor are recently born infants. Providing an argument for this is beyond the scope of this paper, but briefly, my reason is as follows: a fetus is not cognitively sophisticated enough to be morally responsible for its actions, and hence is not a person. But I'd also be happy to follow for example Tooley's reasoning: a fetus is not a continuing subject of experiences and other mental states that can envisage a future for itself and that can have desires about its own future states, and hence a fetus is not a person. Assuming that Tooley's account of personhood is correct, then once an entity becomes a continuing subject of experiences and other mental states that can envision a future for itself and that can have desires about its own future states, it follows (by my lights) that this entity is morally responsible for its actions (in at least a rudimentary way), and hence it follows that it is a person.

Note that it does not follow from what I've said that abortion or

infanticide is morally permissible. Just as we have a prima facie moral obligation not to kill cats, even though cats are not people, so we may have a prima facie moral obligation not to kill non-person humans.

Given that fetuses are not persons, and older children are, how does the transition happen? Is it an all-at-once shift, or are there degrees of personhood? My view is as follows: if the agent goes from being not morally responsible to morally responsible (even if morally responsible just to a small degree), then that's enough for the agent to go from being a non-person to a person. But if there is somehow proto-moral responsibility (where it's not the case that the agent is simply not morally responsible, but it's not the case that the agent is definitely morally responsible), then there would be intermediate degrees of personhood as well, corresponding to these proto-moral responsibility states.

5.2 Objection #2: Personhood is More Fundamental than Moral Responsibility

Some might argue that a person can persist through time, even if there are no moral responsibility relations that hold between the temporal parts of the person over that time interval. Consider a person, Ally, who is by herself, on a desert island, and will continue to be by herself for the rest of her life. This is a special island such that it (and the surrounding waters) contain no sentient creatures other than her. These objectors could maintain Ally can't do anything morally wrong or right, because there are no sentient creatures for her to do those morally wrong or right actions towards. Ally is morally neutral, and yet we all agree that she continues to persist as a person. These objectors would conclude that personhood is more fundamental than moral responsibility.

I'll present five possible ways of replying to this objection. First,

theists could hold that Ally is always in a relationship with God, and hence can always do morally wrong or right by God. Second, one could hold that Ally can still harm herself, and that in itself is morally wrong. (Similarly, Ally can refrain from harming herself, and that is morally commendable.) Third, one could endorse something like virtue ethics, which holds that Ally's developing a good moral character is itself morally commendable, even if there is no one (other than Ally) to benefit from her having that good moral character. So, if a temporal part of Ally cultivates the desire to commit genocide, then the genocide-desiring immediately subsequent temporal parts of Ally are morally responsible for having cultivated that desire. Similarly, if a temporal part of Ally refrains from cultivating that desire, then the immediately subsequent temporal parts are morally commendable for having refrained. This establishes the appropriate chain of moral responsibility and commendability relations, and that grounds Ally's persistence as a person through time.

The second and third answers are my favorites, but if you don't like them, here are two more. One could appeal to counterfactuals: Ally persists because, had other people been around, she would have behaved in a morally accountable way toward them. Or, one could appeal to capacities: Ally persists because she has the capacity for being morally responsible. By my lights, those last two options bring in too much metaphysical baggage, but if you're not happy with the other answers, feel free to consider taking on the baggage.

5.3 Objection #3: What About Sleep? What About Comas? What About Brainwashing?

Some might argue that a person can't be morally responsible for anything

while she is asleep, but she clearly does persist as a person while she is asleep. Thus, my grounding of personal identity in morality fails.

The objector here is forgetting that one can be morally responsible not just for actions, but also for omissions. Omissions are key for the sleep situation. Bob is currently asleep, but he could have woken up just before now and done a morally wrong act; Bob is morally commendable for omitting to wake up and do something morally wrong.

But what about comas? Can Bob persist as a person through a coma? I hold that, if Bob is in a coma, Bob is not morally commendable for omitting to wake up, since he is incapable of doing so. Thus, when Bob is in a coma, he continues to exist as a human organism, but not as a person.

Suppose that Bob's coma is temporary, and he eventually recovers from it. After the coma, Bob continues to exist as a person. Should we be bothered by Bob's discontinuous existence here? I will argue that we shouldn't, for two reasons.

First, we have some continuity, because Bob continues to exist as a human organism throughout the coma, even though Bob as a person goes out of existence and later comes back into existence. (The referent of "Bob" changes depending on context - sometimes it refers to the person, sometimes it refers to the human organism, sometimes it refers to stardust.)

Second, metaphysicians are already familiar with discontinuous existence, when for example a person enters a time machine in the year 2015 and instantaneously (from the standpoint of her personal time) and discontinuously appears in 1815. Of course, time travel is controversial, but the arguments that are used to justify the continued existence of a person over a discontinuous time travel jump are good arguments, and

can carry over to justify the continued existence of Bob as a person. (Such arguments are given by for example David Lewis (1976a).)

For the final case where a person arguably continues to exist without being morally responsible, consider brainwashing. Suppose Sally is given a drug that gives her the desire to kill Fred - Sally is a person, but is arguably not morally responsible for killing Fred. Does this mean that Sally persists as a person without being morally responsible? No, because while Sally is brainwashed to kill Fred, Sally isn't brainwashed to (say) omit from kicking John. So when Sally does omit from kicking John, she's morally commendable for that action (even though she's brainwashed). It's only if the brainwashing controlled all of Sally's actions (and omissions) that she wouldn't be morally responsible - and if Sally is under that sort of complete control, then, indeed, she is no longer a person.

5.4 Objection #4: What About Fission?

So far, I've talked about personal identity thought-experiments as if they involve a temporal part of a person pre-experiment, and a temporal part of a person post-experiment, and we simply have to figure out whether those are temporal parts of the same person. But what about fission? Call "old Katie" the temporal part of a person that enters the fission machine, and "Katie1" and "Katie2" the temporal parts that emerge. How can my grounding of personal identity in terms of moral responsibility make sense of this scenario?

My answer is: unproblematically. Suppose first that the objective moral facts are such that Katie1 and Katie2 are *not* morally responsible for anything old Katie did. It would follow, based on my metaidentity theory, that the person which included the old Katie temporal part has ceased to

exist, and two new people have come into existence, with qualitative similarities to the person who ceased to exist.

Now suppose that the objective moral facts are such that Katie1 and Katie2 *are* morally responsible for something that old Katie did. It follows, by my analysis, that Katie1 and Katie2 are parts of the same person as old Katie. How can one make sense of this? On the perdurance theory of persistence, at least, this is unproblematic, as Lewis (1976b) has shown. There is a persisting person that the temporal parts old Katie and Katie1 are a part of, and there is a different persisting person that the temporal parts old Katie and Katie2 are a part of. The four-dimensional spacetime worms of these persisting people overlap, sharing some temporal parts in common.

But what if (as many believe) proponents of the endurance theory cannot make sense of a person persisting through fission? If the objective moral facts are such that Katie1 and Katie2 are both morally responsible for something that old Katie did, then by my lights it follows that Katie did persist through fission, and hence the endurance theory is false. In this way, we can look to morality not simply to provide answers to questions about personal identity, but to provide answers to other perennial questions in metaphysics too.¹⁾

1) For helpful comments, thanks to anonymous reviewers for this journal, and audiences at University of Sydney, the 2012 Australasian Association of Philosophy conference in Wellington, and University of Colorado Boulder. Also thanks to Rebecca Chan, Matti Eklund, Nicole Hassoun, Ryan Jenkins, Brian Kierland, David Kovacs, Kristie Miller, Chad Mohler, Denis Robinson, and David Shoemaker.

References

- Barnett, David (2013) "You Are Simple", in Georg Gasser and Matthias Stefan (eds.), *Personal Identity: Complex or Simple?*, Cambridge: Cambridge University Press.
- Braddon-Mitchell, David, and Caroline West (2001) "Temporal Phase Pluralism", *Philosophy and Phenomenological Research* 62, pp. 59-83.
- Eklund, Matti (2004) "Personal Identity, Concerns, and Indeterminacy", *The Monist* 87, pp. 489-511.
- Huemer, Michael (2005) *Ethical Intuitionism*, New York: Palgrave MacMillan.
- Lewis, David (1976a) "The Paradoxes of Time Travel", *American Philosophical Quarterly* 13, pp. 145-152.
- Lewis, David (1976b) "Survival and Identity", in Rorty, Amelie O. (ed.), *The Identities of Persons*, Berkeley: University of California Press, pp. 17-40.
- Locke, John (1694) *An Essay Concerning Human Understanding*, Second Edition.
- Olson, Eric (1997) *The Human Animal: Personal Identity Without Psychology*, Oxford: Oxford University Press.
- Rosen, Gideon (2010) "Metaphysical Dependence: Grounding and Reduction", in Bob Hale and Aviv Hoffmann (eds.), *Modality: Metaphysics, Logic, and Epistemology*, Oxford: Oxford University Press, pp. 109-135.
- Shoemaker, David (1999) "Utilitarianism and Personal Identity", *Journal of Value Inquiry* 33, pp. 183-199.
- Tooley, Michael (1979) "Decisions to Terminate Life and the Concept of

a Person”, in John Ladd (ed.), *Ethical Issues Related to Life and Death*, Oxford: Oxford University Press, pp. 62-92.

Wielenberg, Erik (2009) “In Defense of Non-Natural, Non-Theistic Moral Realism”, *Faith and Philosophy* 26, pp. 23-41.

Wiland, Eric (2000) “Personal Identity and Quasi-Responsibility”, in T. van den Beld (ed.), *Moral Responsibility and Ontology*, Dordrecht: Kluwer Academic, pp. 77-87.

Philosophy Department, Univ. of Colorado at Boulder,
bradley.monton@colorado.edu