



# A new argument for the phenomenal approach to personal persistence

Matt Duncan<sup>1</sup>

© Springer Nature B.V. 2019

**Abstract** When it comes to personal identity, two approaches have long ruled the roost. The first is the *psychological* approach, which has it that our persistence through time consists in the continuance of certain of our psychological traits, such as our memories, beliefs, desires, or personality. The second is the *biological* approach, according to which personal persistence consists in continuity in our physical or biological makeup. Amid the bipartite reign of these approaches, a third contender has emerged: the *phenomenal* approach. On this approach, personal persistence consists in continuity in phenomenal consciousness or the capacity for phenomenal consciousness. In this paper I will introduce and defend a new argument for the phenomenal approach. In the process, I will argue against the psychological and biological approaches. I will also address some lingering questions and outline further ways to develop the phenomenal approach.

**Keywords** Personal identity · Personal persistence · Persons · Identity · Experience · Phenomenal consciousness

When it comes to personal identity, two approaches have long ruled the roost. The first is the *psychological* approach, which has it that our persistence through time consists in the continuance of certain of our psychological traits, such as our memories, beliefs, desires, or personality. The second is the *biological* approach, according to which personal persistence consists in continuity in our physical or biological makeup.

---

✉ Matt Duncan  
sduncan@ric.edu

<sup>1</sup> Rhode Island College, Providence, USA

Amid the bipartite reign of these approaches, a third contender has emerged: the *phenomenal* approach. On this approach, personal persistence consists in continuity in phenomenal consciousness or the capacity for phenomenal consciousness. This approach is like the psychological approach in that it focuses on the *mental*. But it is different from historically prominent psychological theories in that it focuses on continuity in phenomenal consciousness rather than continuity in standing states such as memories, beliefs, or desires.

In this paper I will introduce and defend a new argument for the phenomenal approach. In the process, I will argue against the psychological and biological approaches. I will also address some lingering questions and outline further ways to develop the phenomenal approach. My aim is to break up this two-party system by giving new and compelling reasons to endorse the phenomenal approach to personal persistence.

## 1 The phenomenal approach

A theory of personal persistence says what it takes for a person who exists at one time to be identical to a person who exists at some other time. It does this by giving a *criterion*—i.e., metaphysically necessary and sufficient conditions—for personal identity through time.<sup>1</sup> There are many different theories of personal persistence. But the majority of them fit within one of two general approaches to personal persistence. First, there are theories whereby *psychological* continuity is necessary and sufficient for personal persistence. On this approach, a person P persists from time t to time t\*—that is, P at t is identical to P\* at t\*—if and only if P at t has the relevant psychological connections with P\* at t\*. Second, there are *biological* theories. On this approach, P persists from t to t\* if and only if P at t is physically or biologically continuous (in the relevant way) with P\* at t\*.<sup>2</sup>

The phenomenal approach to personal persistence might be perceived as a species of the psychological approach because it implies that personal persistence consists in a kind of mental continuity. But this approach is distinctive in that it zeroes in on that particular part of mentality that has a *what-it's-likeness*—namely, the phenomenal or experiential. There is something it's like to taste chocolate, smell coffee, or feel pain, for example, and this experiential what-it's-likeness is distinctive of phenomenal consciousness (from now on I'll use 'phenomenal consciousness' and 'consciousness' interchangeably). Other psychological theories say that continuity in standing states or traits, such memories, beliefs and desires, or personality traits, is what secures personal persistence. But on the phenomenal

<sup>1</sup> Here (and in what follows) I am setting aside anti-criterialism—the view that there are no criteria of personal identity (or any other object's identity) through time. I have argued against anti-criterialism elsewhere (see Duncan 2014, forthcoming). Rather than rehearse those arguments here, I will simply set anti-criterialism aside and work within the assumption that there is a criterion of personal persistence.

<sup>2</sup> Sydney Shoemaker (1985) and Harold Noonan (1989) hold versions of the psychological view. Peter van Inwagen (1990), Eric Olson (2007), and Judith Jarvis Thompson (1987) are among those who hold the biological view.

approach, a person P at time t is identical to a person P\* at time t\* if and only if there is the right sort of *phenomenal* or *experiential* continuity between P and P\*.<sup>3</sup>

What's the "right sort" of phenomenal or experiential continuity? This raises all sorts of important questions. One has to do with how complex or sophisticated one's experiences have to be for one to persist through time. Could one persist with experiences no more sophisticated than that of a cow, or a fish, or a worm? Or is more required? Is there something distinctive about *human* experience that's necessary for persistence? Some who defend the phenomenal approach say that in order for a person to persist, her experiences must be fairly sophisticated (e.g., Unger 1990). Others, such as Dainton and Bayne (2005), say that, "no cognitive sophistication is necessary for our survival, and that we could survive with a consciousness of the simplest of forms, e.g., a few basic bodily feelings" (p. 560–561). Those are just a couple of possible answers.

Another question has to do with whether phenomenal continuity *itself*—that is, P and P\* being subjects of the same uninterrupted episode or stream of experience—is necessary for personal persistence. People sleep. Sometimes people go into comas, then come out. And most philosophers agree that they can survive such periods of unconsciousness. So, rather than construing personal persistence in terms of phenomenal continuity *itself*, a more popular strategy within the phenomenal approach is to construe it in terms of some underlying feature that is *essential* for phenomenal continuity—typically, the *capacity* for consciousness.<sup>4</sup> So then the idea is that P and P\* are identical just in case they have the same capacity for consciousness.

Yet another question has to do with how to state all this in a way that is *non-circular*—that doesn't presuppose the identity of the person in question. Peter van Inwagen (1990), for example, presses this issue when he says:

To imagine whether a certain situation contains a continuous consciousness we have to find out first whether a certain situation contains a continuously existent thinker. We can't do things the other way around. We can't find out whether the situation contains a continuously existent thinker by first finding out whether it contains a continuous consciousness (p. 206).

The worry here is that any phenomenal theory of personal persistence will be *circular*, since characterizing personal persistence in terms of phenomenal continuity (or continuity in the capacity for consciousness) requires stating or presupposing *whose* consciousness it is—the very person whose persistence is in question.

Defenders of the phenomenal approach address this issue in a variety of different ways. Some do it by specifying personal persistence in terms of purely qualitative, non-identity-assuming features of streams of consciousness or capacities for

---

<sup>3</sup> Recent consciousness-based theories are somewhat scarce. Dainton and Bayne (2005) sketch some potential approaches to such a theory, and Dainton (2008) and Bayne (2010, Ch. 12) each describe their separate views. John Foster (1991, Ch. 8) and Peter Unger (1990) also propose consciousness-based theories.

<sup>4</sup> I'll say a lot more about capacities for consciousness—including what exactly they are, how they are individuated, and how they persist through time—later on, in Sect. 3.

consciousness (see, e.g., James 1952, p. 155; Foster 1991). Others address it by making reference to the *parts* (and interrelations between parts) responsible for consciousness or the capacity for consciousness (cf. Dainton and Bayne 2005, p. 565). Just as one could specify personal persistence in terms of bodily continuity in a non-identity-assuming way by talking about the continuance of bodies or bodily parts (“Not assuming *whose* parts they are!”), so too one could specify personal persistence within the phenomenal approach by talking about the continuance of minds, or brains, or capacities for consciousness (“Again, not assuming *whose* they are!”). As before, these are just some possible options.

I will say more about each of these issues or questions later on when I talk about how to further develop a theory of personal persistence within the phenomenal approach. But I won’t take any stands on them now. At this point, I don’t want to prejudge how the phenomenal approach should be pursued. For my argument to follow does not presuppose any particular version of it. And, at any rate, it’s important to appreciate that there’s wiggle room within the phenomenal approach—that there’s more than one way to pursue it.

So, with that, here’s how I will construe the phenomenal approach to personal persistence:

*The Phenomenal Approach to Personal Persistence:* Personal persistence consists in either phenomenal continuity (which may or may not have to be complex/sophisticated to some degree) or continuity in something that is essential for such phenomenal continuity, such as the capacity for consciousness. In other words, either phenomenal continuity, or continuity in something that is essential for such phenomenal continuity, is metaphysically necessary and sufficient for personal persistence.

I’m calling this an “approach” to personal persistence because, at this stage of development, it’s just a rough sketch of a theory—one that leaves plenty of room for further development (and disagreement) among its advocates. As I’ve said, a theory of personal persistence gives a criterion—metaphysically necessary and sufficient conditions—for personal identity through time. And while the phenomenal approach tells us a fair bit about what that criterion should look like, it is consistent with a variety of more specific criteria.

So while the first step (and the main aim of this paper) is to establish the phenomenal approach to personal persistence, an important next step is to generate a more developed theory of personal persistence. Later on I will offer some suggestions for how to do this. But now I’ll turn to my argument for the phenomenal approach.

## 2 A new argument for the phenomenal approach

My argument for the phenomenal approach is new—no one has introduced it before. But it’s also a relatively unexplored *kind* of argument for the phenomenal approach. Most such arguments rely on our intuitions about possible cases. We are asked to consider some episode, perhaps involving a brain transplant, silicon chips, or what

have you, then we are told to consult our intuitions about what would happen in that episode (Did she survive? Where did she go?), and then we are shown how these intuitions support the phenomenal approach. This is a standard way of reasoning about personal identity. But it's also very controversial. Some love it, some hate it. I won't take a stand on the merits of this methodology here. But my approach will be different. My argument for the phenomenal approach will be based, not on intuition, but on our direct first-personal awareness of our persistence through time. Hence, I will call it 'The Direct Argument'.

Here is the first premise of The Direct Argument:

- (1) Continuity in normal (i.e., unimpaired) consciousness is *sufficient* for personal persistence.

The more formal gloss on (1) is: Necessarily, if person P at time t is phenomenally continuous with—i.e., is the subject of the same unimpaired phenomenal stream as—person P\* at time t\*, then P and P\* are identical. So if I, for example, am P, then as long as *this* particular episode of consciousness, which I am currently the subject of, goes on unimpaired and uninterrupted, then I'll survive—I'll keep existing as P.

The above premise again raises the question of how to construe this continuity in consciousness in a way that is non-circular—i.e., that doesn't presuppose that P and P\* are the same person. Again, I'll have more to say about this in the next section. For now, set this issue to one side and just note the options mentioned in the previous section. In other words, for now, consider (1) to be a claim concerning continuity in certain purely qualitative features of streams of consciousness or else continuity in the parts (and interrelations between those parts) that are directly responsible for that continuity in consciousness.

Now, as I've said, most advocates of the phenomenal approach defend (1) by appealing to our intuitions about possible cases. The idea is, no matter what happens to P—even if she's diced up, deformed, transported, or transformed—as long as her phenomenal stream persists, so will she. Why this verdict? Because, it's intuitive.

My defense of (1) is different. I've laid out my argument for (1) in detail elsewhere (Duncan 2015a), and it doesn't rely on intuition—it relies instead on direct first-personal evidence of our persistence through time. Since my goal in this paper is to move *beyond* (1)—to develop it into an argument for the phenomenal approach to personal persistence—I'll limit myself at present to a short rendition of my argument for (1).

It goes as follows. Whenever I (for instance) am the subject of a phenomenal experience—the thought ' $2 + 2 = 4$ ', for example—I can be absolutely certain that that experience exists and that I am its subject.<sup>5</sup> I could be wrong about all sorts of

---

<sup>5</sup> Here, for various reasons, my example is a *thought*. But some (e.g., Carruthers 2011; Prinz 2011) deny that thoughts are conscious. These philosophers should feel free to substitute my talk of thoughts with talk of inner speech or some other conscious event. Doing so will not affect my point. Also, here I do not mean to be making any substantive assumptions about the individuation of thoughts or any other conscious event—about, for example, whether ' $2 + 2 = 4$ ' should be understood as a single thought or

things, but I couldn't be wrong that *this* thought that I'm thinking right now is *my* thought.<sup>6</sup> So right now I can be certain that I exist. So I exist. But now notice: Thinking takes time. Some thoughts are short. But even ' $2 + 2 = 4$ ' takes at least a few milliseconds to think. So the fact that I can be certain that this thought exists and that I am its thinker actually implies that I can be certain that I *persist*. So I persist. Specifically, I persist as long as it takes to think ' $2 + 2 = 4$ '. And this very same evidence is available to anyone who is the subject of *any* brief, uninterrupted and unimpaired experience. So, whenever anyone is the subject of such an experience, she can be certain that she persists. In which case she persists. Therefore, from any one moment to the next, continuity in unimpaired consciousness is sufficient for personal persistence. So, for example, if person  $P_1$  at time  $t_1$  is phenomenally continuous with person  $P_2$  a few milliseconds later at time  $t_2$ , then  $P_1$  and  $P_2$  are identical. And if  $P_2$  is phenomenally continuous with person  $P_3$  a few milliseconds later at time  $t_3$ , then  $P_2$  and  $P_3$  are identical. The same goes on down the line as long as there is phenomenal continuity. And so, by the transitivity of identity,  $P_1$  is identical to  $P_2, P_3 \dots P_n$ , etc., so long as their phenomenal continuity isn't interrupted or impaired. Therefore, continuity in unimpaired consciousness is sufficient for personal persistence. In other words, (1) is true.

In order to deny this result, one would have to say that it's possible that it's not really *you* who thinks ' $2 + 2 = 4$ ' in the above case. That's a bad option to be left with. To see this, assume for a moment that you are the person present at the end of the thought. One would have to say that there is some possible case *exactly* qualitatively like you thinking ' $2 + 2 = 4$ ', in which it seems to you as if you are thinking ' $2 + 2 = 4$ ', but in which you didn't think that thought—that it was perhaps someone else who thought it (or at least part of it). But that's false. First of all, it smacks of a dubious distinction between your thinking ' $2 + 2 = 4$ ' and it merely *seeming* to you as if you are thinking ' $2 + 2 = 4$ '. There is no such distinction. If it seems to you that you're thinking ' $2 + 2 = 4$ ', then you *are* thinking ' $2 + 2 = 4$ '. It's like pain in this way. At least in clear cases of pain, your seeming to be in pain *just is* your being in pain. The same goes for thinking simple thoughts like ' $2 + 2 = 4$ '. There is nothing more to your thinking ' $2 + 2 = 4$ ' than its *appearing* in your conscious experience.

This point gains even more force when we consider the specific kind of mistake that you would have to be making here. In the case we are considering, you wouldn't necessarily be wrong to believe that ' $2 + 2 = 4$ ' was thought. Rather, it would be in believing that *you* thought it. But this is not a mistake that you can

---

Footnote 5 continued

rather a series of shorter thoughts. ' $2 + 2 = 4$ ' is just an arbitrary example. Feel free to think up your own example, if you like.

<sup>6</sup> This point may be seen as an invocation of Descartes' Cogito. But one needn't accept Descartes' argument, or various claims associated with it, in order to accept that I can be certain that *this* thought is *my* thought. Many philosophers, only some of whom are sympathetic to Cartesian doctrines, accept that we are "immune to errors of (self-)misidentification" (see below for more on this). This claim should also be kept separate from claims about self-knowledge *in general*, or in other domains. To say that I can be certain that this thought is my thought does not entail that our powers of introspection are immune from error *in general* or even that they are particularly reliable.

make. You are, as they (i.e., philosophers) say, *immune* from such errors.<sup>7</sup> If you know that a thought is thought, and you judge on the basis of the way things seem to you that *you* are thinking it, then you are right, you *are* thinking it. So the notion that you could be wrong to believe that you are thinking ‘ $2 + 2 = 4$ ’ simply doesn’t gain any traction. And since the same goes for any brief, uninterrupted and unimpaired thought or conscious episode, we can conclude that, from any given moment to the next, such phenomenal continuity is sufficient for personal persistence. That is, we can conclude that (1) is true.

Notice that this argument for (1) does not rely on intuitions about possible cases. It may very well *explain* those intuitions. That is, our direct awareness of our persistence through time may explain why we find it intuitive that people persist if their phenomenal streams persist. Yet the argument does not rely on those intuitions.

In fact, it can *replace* those intuitions in the realm of possible cases in order to defeat competing theories of personal persistence. This includes various versions of the psychological and biological approaches. For instance, take the most prominent version of the biological approach: animalism. Animalism entails the following:

(A) Necessarily, if person P at time t is identical to person P\* at time t\*, then P and P\* are the same human animal.

With the aid of possible cases, my argument for (1) can be used to show that (A) is false, and thus, that animalism is false. To see this, start by thinking ‘ $2 + 2 = 4$ ’. Close your eyes, clear your mind, and just think ‘ $2 + 2 = 4$ ’. Now suppose that, in the few milliseconds it takes you to think ‘ $2 + 2 = 4$ ’, aliens painlessly destroy your body. The only part of you they don’t destroy is your cerebrum. But, on the bright side, the aliens manage to sustain the normal functioning of your cerebrum. So there’s no detectible phenomenal disturbance. In fact, you’re so absorbed in your thought that you completely fail to notice what the aliens have done.

Your direct first-personal evidence concerning your thinking is the same as in any other case of thinking, ‘ $2 + 2 = 4$ ’. So you can be absolutely certain that you thought that thought. So you thought it. Which means you persisted through the time it took you to think it. But (A) implies the opposite. The person sitting in your seat prior to the alien insurrection is not the same human animal as the subsequent cerebrum. Cerebra aren’t animals, after all.<sup>8</sup> So if (A) is true, you didn’t persist. But you did! You know it. How? Because you know you thought ‘ $2 + 2 = 4$ ’. And in

<sup>7</sup> See, for example, Shoemaker (1968), O’Brien (2007), Evans (2001), Howell (2006), and Gertler (2011, p. 215–217). I take the claim that we are immune from the sort of errors mentioned above to be uncontroversial. It is controversial which cases are to count. But the case that I have described should be safe by anyone’s standards. And even if it is not—if, for example, you think that ‘ $2 + 2 = 4$ ’ is *too long* of a thought to be immune from error through misidentification—then feel free to just pick a different, shorter thought. Maybe just think about something for as quickly as you can—for a few milliseconds, perhaps. It’s doubtful that any thought of which we are aware, or that’s ever been used as an example to illustrate immunity from error, is instantaneous (see Duncan (2015a)). So some such thought will do.

<sup>8</sup> Olson (2007, p. 41) says that detached cerebra aren’t even *organisms*, let alone animals. However, some philosophers who might be called animalists say that cerebra are (or at least can be) organisms if they are separated from the body (see van Inwagen 1990). These philosophers might say that you *do* persist from t to t\*. So the present argument does not apply to those versions of animalism.

order to think that thought, you had to take the time to think it, which is more than the time it took the aliens to do their deed. So, contra (A), you persisted. Thus, (A) is false, and so is animalism.

This line of reasoning also fells various versions of the psychological approach, including the most prominent among them: the memory theory. This theory entails:

(M) Necessarily, if person P at time t is identical to person P\* at time t\*, then P\* has at least some first-personal memories of events occurring to P.

Now suppose that a mad neuroscientist erases all of your first-personal memories, including both your explicit and tacit first-personal memories, while you think '2 + 2 = 4'. (M) entails that you do not persist. But you do. Thus, (M) is false, and so is the memory theory.

There are other actual and potential versions of the psychological approach that face the same fate. Some say that people persist in virtue of continuity in other psychological states or processes, such as beliefs, desires, or personality traits. But my argument rules out these views. For my argument shows that you can persist through the loss of *any* part, state, or process that is inessential to thinking '2 + 2 = 4'. You can lose beliefs, desires, personality traits, and a wide variety of other psychological states and processes without thereby losing your ability to think. So my argument rules out any version of the psychological approach featuring such states or processes.

There's a lot more to say about my arguments for (1) and against (A), (M), and other theories. But I must move on, lest I remain moored to this spot for too long. For, again, my goal here is to move *beyond* (1) in defense of the phenomenal approach (for a more detailed defense of the above arguments, see Duncan 2015a).

So back to The Direct Argument. Here's its second premise:

- (2) Given (1), we should accept the phenomenal approach. That is, given (1), we should accept that either phenomenal continuity (which, again, may or may not have to be sophisticated to some degree) or continuity in something that is essential for phenomenal continuity (e.g., the capacity for consciousness) is metaphysically necessary and sufficient for personal persistence.

Premise (2) does not say that the phenomenal approach follows *logically* from (1). It only says that (1) gives us sufficient reason to accept the phenomenal approach. My argument for (2) will be an argument *by elimination*. Here's a rough outline of it: Any potential theory personal persistence will either be *simple*—stating just one condition for personal persistence (e.g., biological continuity)—or it will be *complex*—stating multiple conditions meant to be either individually necessary and jointly sufficient (i.e., a conjunctive theory) or individually sufficient and jointly necessary (i.e., a disjunctive theory) for personal persistence. Those are all of the possible options. (Most, if not all, actual theories of personal persistence are simple theories. But it's best to cover all the bases.) And, as it turns out, any such theory not within the phenomenal approach will either be inconsistent with (1) or will be unacceptable for other reasons. Thus, given (1), we will have sufficient reason to accept the phenomenal approach to personal persistence.



Let me explain further. First, take some condition for personal persistence, C. By (1), unimpaired phenomenal continuity is sufficient for personal persistence. So unless the obtaining of C is essential for such phenomenal continuity, a person can persist without C obtaining. In which case C isn't necessary for personal persistence. In which case C, by itself, isn't the correct theory of personal persistence. Thus, we can conclude that, given (1), if C isn't essential for phenomenal continuity—in other words, if C is inconsistent with the phenomenal approach—then C is not, by itself, the correct theory of personal persistence.

Suppose C is just one part of a complex theory of personal persistence. Specifically, suppose it is part of a *conjunctive* theory of personal persistence—a theory whereby person P at time t is identical to person P\* at time t\* if and only if both C *and* some other condition D obtain. By (1), if the obtaining of either C or D is inessential for phenomenal continuity, then a person can persist without the conjunction of C and D obtaining. So if any conjunctive theory of personal persistence contains a condition that is inessential for phenomenal continuity, then it is not the correct theory of personal persistence. So we can rule it out.

Now suppose C is part of a *disjunctive* theory of personal persistence—a theory whereby person P at time t is identical to person P\* at time t\* if and only if either C *or* D obtains. Since phenomenal continuity is sufficient for personal persistence, a disjunctive theory of personal persistence must contain phenomenal continuity as a disjunct. But it is consistent with (1) that it also contains another disjunct that has nothing to do with phenomenal continuity. So a disjunctive theory that isn't *fully* within the phenomenal approach is consistent with (1).

But, as we will see, disjunctive theories of personal persistence are implausible. Plus, no one (who I know of) defends one.<sup>9</sup> That could be for all sorts of reasons—reasons that speak to the implausibility just mentioned. It could be because disjunctive theories of personal persistence are counterintuitive (why would we have multiple kinds of persistence conditions?). Or because they are less simple. Or because there is no good motivation for adopting one. Or because philosophers are averse to disjunctive theories *in general*.

Or it could be because disjunctive theories face a special threat of contradiction. Here's an example to illustrate: Suppose that for C to obtain between x and y is for x and y to have the same *brain*, and for D to obtain between x and y is for x and y to have the same *right arm*. Now consider Peter, whose right arm is severed and then attached to a different body—to someone we might call 'Quinton'—shortly thereafter. So we have Peter who exists at t, Quinton who exists at t\*, and then we have the disarmed person—call him 'Peter\*'—who exists at t\*. D obtains between Peter and Quinton, since they have the same right arm. C obtains between Peter and Peter\*, since they have the same brain. Thus, according to the disjunctive theory in

<sup>9</sup> Of course, any theory of personal persistence can be *stated* in disjunctive form. A psychological continuity theory, for instance, can be stated as a disjunction of all the *determinate* mental continuities that make for psychological continuity. But the kind of theory we want to consider here is one with multiple sufficient (and only *jointly* necessary) conditions that are divergent enough from each other that the theory isn't better stated in terms of a single, non-disjunctive necessary and sufficient condition. And, again, no one that I know of defends such a theory.

question, Quinton is identical to Peter, Peter is identical to Peter\*, and so by the transitivity of identity, Quinton is identical to Peter\*. But Quinton is *not* identical to Peter\*. So we get a contradiction.

For a more tendentious example, we might suppose that for C to obtain between x and y is for x and y to be *psychologically continuous*, and that for D to obtain between x and y is for x and y to be *biologically continuous*. Then we might imagine a case in which some person P's psychology is transferred to a new body so that C obtains between P and one person, and D obtains between P and another person. We would get the same result: a contradiction.

These are just two examples. In fact, *any* disjunctive theory with disjuncts that can possibly diverge in this way will lead to contradiction. So what a potential disjunctive theorist would need to do is come up with conditions that do not lead to contradiction when disjoined. Of course, she would also have to make sure that her disjunction is a plausible theory of personal persistence. This is a tall order—one I doubt can be filled. Thus, since disjunctive theories of personal persistence face serious problems, and they lack defenders, I will set them aside.

Where does that leave us with (2)? Well, any simple (or conjunctive) theory of personal persistence inconsistent with the phenomenal approach is false. And the other option—a disjunctive theory—is very unappealing. So we should accept (2). That is, given (1), we should accept that phenomenal continuity, or else in continuity in something that is essential for phenomenal continuity (e.g., the capacity for consciousness), is necessary and sufficient for personal persistence. That is, given (1), we should accept the phenomenal approach.

So if (1) is true, we should accept the phenomenal approach. (1) *is* true, as I've argued. So we should accept the phenomenal approach. Thus, with (1) and (2), we arrive at my main conclusion—what I aim to establish.

So here's my argument—The Direct Argument—in full:

- (1) Continuity in normal (i.e., unimpaired) consciousness is *sufficient* for personal persistence.
- (2) Given (1), we should accept the phenomenal approach. That is, given (1), we should accept that either phenomenal continuity or continuity in something that is essential for phenomenal continuity (e.g., the capacity for consciousness) is metaphysically necessary and sufficient for personal persistence.
- (3) Therefore, we should accept the phenomenal approach (1, 2).

The above constitutes an argument—indeed, a new argument—for the phenomenal approach to personal persistence. Let's review its motivations. (1) is based on direct, first-personal awareness of one's experiences plus the observation that experiences (e.g., the thought '2 + 2 = 4') are temporally extended. (2) is the result of some straightforward reasoning about our options in light of (1). And (3) follows from (1) and (2).

None of these premises depend on intuitions about possible cases. They aren't claims that require you to think through puzzles or weigh controversial theses. They're truths that you can discover right here and now. Right now, you can attend to your thoughts and discover that you've persisted through their thinking. Right

now, you can reason about your options in light of (1), and discover that (2) is true. And right now, you can put this all together and discover that you should accept the phenomenal approach to personal persistence.

### 3 Toward a theory of personal persistence

My primary objective in this paper is to lay out a new argument for the phenomenal approach to personal persistence. Since I have now done that with The Direct Argument, I have satisfied my primary objective in this paper.

However, I want to say a bit more. For I want to give us a more specific direction—a path toward a *criterion* of personal persistence. So, in this section, I will address some lingering questions—including those raised in the first section—and offer various suggestions for how to further develop the phenomenal approach.

Now, one question that I won't discuss at length here is the question of how complex or sophisticated one's experiences have to be (or have to be capable of being) in order for one to persist. As I mentioned earlier, some defenders of the phenomenal approach think that a person's experiences must remain fairly sophisticated in order to persist (e.g., Unger 1990); others say we can persist with very basic experiences (e.g., Dainton and Bayne 2005). I am inclined to think that the truth lies somewhere in between. Elsewhere I have argued that human consciousness includes a certain kind of *self-awareness* or *self-experience* (see Duncan 2015a, b, 2018, forthcoming). And my view is that, in order for a person to persist through time, her experiences needn't be particularly complex or sophisticated, but they must contain (or be capable of containing) such self-experiences.<sup>10</sup> However, I won't pursue this point further here. In what follows I will talk about phenomenal continuity but leave open how complex or sophisticated that continuity needs to be for personal persistence.

Another question—which I mentioned earlier, in Sect. 1—is whether it's phenomenal continuity *itself*, or rather continuity in something that is essential to phenomenal continuity, that is necessary and sufficient for personal persistence. I suggest the latter. For I suggest that phenomenal continuity isn't necessary for personal persistence. Here's why: I took a nap today and made it through. I'll leave you to construct your own proof (though this paper may aid you in this proof in

---

<sup>10</sup> This view is not only intuitively plausible (or so I say), it may also soothe other intuitions that may at first seem to be at odds with the phenomenal approach. For example, some philosophers say that we people are more than just *experiencers*—thinkers, feelers, and perceivers. They say we are also *agents* (Velleman 2006); or we are essentially capable of *reflection* and *self-evaluation* (Taylor 1976); or we are essentially crafters of *self-narratives* (Schechtman 2011; MacIntyre 1984); or we are essentially bearers of *moral responsibility* (Locke 1689/1975). I believe that these philosophers are mistaken—for some of the reasons I gave in the previous section against animalism and various psychological continuity theories, I don't think we are any of the above things *essentially*. However, given my view that personal persistence requires the continued capacity for self-experience, I can help explain why the above traits may *seem* to be essential to us. I think the above traits are *connected* to our essence in that the capacities for action, reflection, self-evaluation, and self-narration are grounded in our capacity to undergo experiences that contain self-experiences. They depend on the kind of self-awareness afforded to us by our distinctive form of consciousness.

various ways). The point is, people can persist through periods of unconsciousness. So phenomenal continuity is not necessary for personal persistence.

Unlike with the premises in The Direct Argument, I don't have much more to say to support the claim that we can survive sleep, besides that it's obvious, that nearly everyone agrees that it's obvious, that it's a crucial to our reasoning about who we are, what are responsible for, what it's rational for us to look forward to or regret ... things like that. Maybe this counts as or involves intuitions (I don't claim that the suggested developments in this section are completely intuition-free.). Be that as it may, I do think we can survive sleep. So I suggest that phenomenal continuity itself is not necessary for personal persistence.

From this, together with (3) above, it follows that something essential to phenomenal continuity (e.g., the capacity for phenomenal consciousness), but not phenomenal continuity itself, is necessary and sufficient for personal persistence. Thus, given (3) plus the plausible claim that we can survive sleep, I conclude that something essential to phenomenal continuity is necessary and sufficient for personal persistence.

To be more specific, I suggest that it's continuity in the *capacity* for phenomenal consciousness that is necessary and sufficient for personal persistence. Here's how I understand what a capacity for consciousness is:

*Capacity for Consciousness:* Person P has the capacity for consciousness at time t if and only if, at t, P's parts and their interrelations are such that either P is consciousness or P would be so if P were to be appropriately stimulated.

A capacity for consciousness is a dispositional feature, the exercise of which results in consciousness. I assume that people have this capacity in virtue of the nature of certain of their parts—such as parts of the brain, perhaps—and the interrelations among those parts.<sup>11</sup> And I assume that capacities for consciousness are individuated by those parts. I have a capacity for consciousness in virtue of having certain parts, and you have your own capacity for consciousness in virtue of having different parts. If we are wholly physical things, then it's likely that we each have the capacity for consciousness in virtue of how the parts of our brains are constituted and interrelated. But let me set aside issues of underlying ontology for now. I'll return to them later. The point here is just that, whatever the relevant parts are, they serve to individuate our capacities for consciousness so that, although we sometimes talk about capacities as something we share, we can also talk, as I will do in what follows, about *individual* capacities for consciousness—i.e., each person's distinct capacity for consciousness as individuated by the particular parts underlying them.

If at any given time a person's capacity for consciousness is being exercised, then that person is conscious. But a person need not be conscious at a given time in order

---

<sup>11</sup> I say "such as parts of the brain, perhaps," but my claim that people have the capacity for consciousness in virtue of the nature of certain of their parts is compatible with wide range of views—even, for example, substance dualism. Most substance dualists say that minds/souls are *simple* (i.e., they lack proper parts). But then the above claim would just be that people have the capacity for consciousness in virtue of the nature of their *one* part—their simple immaterial soul.

to have the capacity for consciousness. I, for example, had the capacity for consciousness throughout the night last night even though, at times, I was fast asleep. This is because my parts were such that I *could have* been consciousness—I was *poised* for it. If I would've been appropriately stimulated—if an alarm would have gone off, for example—then I would have been conscious.

To *appropriately stimulate* someone is to do whatever it takes to get that person conscious. But there are limits. We don't want a notion of appropriate stimulation so broad that *any* intervention, no matter how extreme, counts. Perhaps a lifeless lump of tissue could be modified, rearranged, and enhanced so that it has experiences. But we wouldn't want to say that, prior to this intervention, the lump of tissue had the capacity for consciousness (cf. Dainton 2008, §10.7). On the other hand, we don't want a notion of appropriate stimulation so narrow that only gentle taps and pleading voices count. After all, heavy sleepers have the capacity for consciousness.

So we need to draw the line somewhere. Here's my suggestion: Appropriate stimulation may include any intervention that does not add to or subtract from the sum of the token parts directly responsible for one's consciousness, and does not alter the structural arrangement of those parts except for minor alterations that are mere bi-products of stimulation (such as small changes in neuronal arrangement that inevitably result from any stimulation).<sup>12</sup> Thus, alarm clocks, vigorous shaking, splashes of cold water, and CPR all count as appropriate stimulation, but frontal lobe transplants do not (if, that is, the frontal lobe is part of what is directly responsible for consciousness). This is somewhat vague. But I doubt there is any way for me to completely avoid vagueness here without saying something implausible or speculative.<sup>13</sup>

So a person has the capacity for consciousness at any given time if and only if, at that time, she is conscious or would be so if she were appropriately stimulated. Now we need to determine what makes for *continuity* (or sameness) in a person's capacity for consciousness *across time*. Here's what I suggest makes for sameness of capacities for consciousness at *adjacent* times (which I will call 'CC-Connectedness'):

*CC-Connectedness*: Person P's capacity for consciousness C at time t is identical to person P\*'s capacity for consciousness C\* at temporally adjacent time t\* if and only if P and P\* would be phenomenally continuous if P and P\* were appropriately stimulated.

---

<sup>12</sup> I've added "directly responsible" here because removing a person's arm (for example) would change her structural features, but not in any way that's relevant to her capacity for consciousness.

<sup>13</sup> This is not to say that there is no fact of the matter as to whether an intervention counts as appropriate stimulation. I think there is always a fact of the matter as to whether something is a person, as to whether it is conscious, and as to whether something has the capacity for consciousness. Thus, I think there is always a fact of the matter as to whether an intervention merely stimulates (rather than creates) a capacity for consciousness, and thereby counts as appropriate stimulation. Still, there will be borderline cases. But that's true of almost any theory of personal persistence (Consider: being caught up in a life, being an animal, having the same mental states or characteristics as, and various other conditions that feature in theories of personal persistence, all admit of borderline cases).

On my account, sameness of capacities for consciousness at adjacent times is a matter of whether P and P\* are, or would be (if appropriately stimulated), *phenomenally continuous*—that is, subjects of the same, uninterrupted phenomenal stream. By design, this is a *minimal* account of CC-Connectedness, in that it leaves open a number of important issues, including, most pressingly, further specifics about what it takes for P and P\* to be phenomenally continuous.

As an example, some philosophers claim that the persistence of any entity or process requires a certain kind of cross-temporal *causal dependence*. In the present case, the claim would be that to count as the same episode of consciousness, P\*'s thinking must causally depend (in the relevant way) on P's thinking—it can't *just* be that they occur in succession.<sup>14</sup> This causal condition is not explicitly stated in my characterization of CC-Connectedness, but it may be adopted as part of what it takes for P and P\* to be phenomenally continuous.

It may also be that in order for P and P\* to be phenomenally continuous there must be a certain degree of continuity (or overlap) in the parts and interrelations between parts responsible for their consciousness. If sometime between t and t\*, P's cerebrum is swapped out for a new cerebrum, which is then P\*'s cerebrum, then although P and P\* may each be capable of having experiences, and may be causally related in some ways, one might reasonably contend that P is *not* phenomenally continuous with P\* (and thus is not CC-Connected with P\*), because P and P\* simply do not share enough parts. This raises the question of how much overlap *would be* enough. There's room for disagreement here. But one plausible way to draw the line is to say that P and P\* are phenomenally continuous only if P and P\* would still be conscious (given appropriate stimulation) if their non-overlapping parts and interrelations were removed—that is, if we subtracted their differences. So if P and P\* have different cerebra, then they are not phenomenally continuous, because if we subtracted their differences—i.e., their cerebra—they would not have the capacity for consciousness (assuming consciousness requires a cerebrum). Yet various smaller changes would be all right—they wouldn't rule out phenomenal continuity. If P and P\* differ only with respect to one atom, or one neuron, or even various parts of the brain, they may nonetheless be phenomenally continuous, since subtracting these differences would not thereby render P and P\* incapable of being conscious. This suggestion respects the idea that major discontinuities in one's parts disrupt continuity in one's capacity for consciousness (and thus CC-Connectedness), while also allowing that our capacities for consciousness can (and do) persist through some smaller changes.

This elaboration also allows us to deal with the circularity worry mentioned earlier. Again, the worry is that any phenomenal theory of personal persistence will be *circular*, since characterizing personal persistence in terms of phenomenal continuity requires stating or presupposing *whose* consciousness it is—the very

<sup>14</sup> See Lewis (1976, p. 17) and Shoemaker (1985) for discussion of this causal dependence. Dainton and Bayne (2005) contend that this sort of dependence should not be construed as *causal*—that it can be based on a kind of *experiential* dependence. It's not clear to me why the experiential dependence that Dainton and Bayne describe shouldn't count as a kind of causal dependence, broadly construed. But not much turns on this point.

person whose persistence is in question. But now, with the above elaborations in tow, there's a way around this worry. Recall something I said earlier: People have the capacity for consciousness in virtue of the nature of certain of their parts and the interrelations among those parts. The same, no doubt, is true for phenomenal continuity: Continuity in certain of my parts and their interrelations is responsible for my phenomenal continuity. If we make this explicit—if we characterize continuity in the capacity for consciousness in terms of our parts and their interrelations—we can avoid invoking the identities of any person in question. We can talk of person P at time t and person P\* at time t\*, and, without presupposing whether they are the same person, ask whether they are CC-Continuous in virtue of whether the parts that make up their capacities for consciousness are continuous in the way described above.

These are just a few ways to elaborate on what it takes to be phenomenally continuous. Further attention to this issue is needed, as that will help add flesh to the bones of my characterization of CC-Connectedness, which, again, is a characterization of what makes for sameness in the capacity for consciousness at adjacent times.

But now what about *non*-adjacent times? My capacity for consciousness is the same such capacity that I had five minutes ago, yesterday, and indeed, 20 years ago. It's like my body in this regard. My body has undergone all sorts of changes over the years, but it has remained *numerically* the same. Likewise, although the parts and interrelations between parts that are responsible for my capacity for consciousness have changed over the years, I've maintained the same such capacity. This, in itself, is neither a surprise nor a problem. Indeed, it follows from my characterization of CC-Connectedness. If person P<sub>1</sub> at time t<sub>1</sub> is CC-Connected to person P<sub>2</sub> time t<sub>2</sub> (which is adjacent to t<sub>1</sub>), who is CC-Connected to person P<sub>3</sub> at time t<sub>3</sub> (which is adjacent to t<sub>2</sub>), and so on, then by the transitivity of identity, P<sub>1</sub>'s capacity for consciousness is identical to P<sub>2</sub>'s capacity for consciousness as well as to P<sub>3</sub>'s capacity for consciousness as well as to any person's capacity for consciousness that is in an unbroken chain of CC-Connections with P<sub>1</sub>.<sup>15</sup> Thus, so long as there is an unbroken chain of CC-Connections between my 8-year-old self and me (and there is), we have the same capacity for consciousness. Again, this is neither a surprise nor a problem. However, it does raise the question of how to *talk* about sameness in capacities for consciousness over extended periods of time. We cannot say that two people have the same capacity for consciousness if and only if they are CC-Connected. For I am not CC-Connected with my 8-year-old self. So what we should say is that two people have the same capacity for consciousness if and only if there is an unbroken chain of CC-Connections connecting them across

<sup>15</sup> Four-dimensionalists may prefer to talk about persistence in terms of a series of non-identical *object stages* that are related in some way other than by strict identity. If this is one's view, then invoking the transitivity of identity won't work here. But one can get the same result by characterizing sameness in capacities for consciousness over any period of time in terms of the *ancestral* of CC-Connectedness, whereby person stage P has the same capacity for thinking as person stage P\* if and only if P is CC-Connected to P\*, or P is CC-Connected to a person stage who is CC-Connected to P\*, or P is CC-Connected to a person stage who is CC-Connected to a person stage who is CC-Connected to P\*, and so on.

time. Call this ‘CC-Continuity’. Now we can say that, for *any two times*  $t$  and  $t^*$ , person  $P$  at  $t$  has the same capacity for consciousness as person  $P^*$  at  $t^*$  if and only if  $P$  and  $P^*$  are CC-Continuous. So the parts and interrelations between parts that are responsible for a person’s capacity for consciousness may change over time, even substantially, while remaining the very same such capacity, just so long as that change is gradual enough that this person never lacks CC-Connectedness from one moment to the next.

So sameness in the capacity for consciousness is CC-Continuity. Thus, when I say that personal persistence consists in continuity in the capacity for consciousness—that continuity in this capacity is necessary and sufficient for personal persistence—I am saying that personal persistence consists in CC-Continuity. That’s how I believe the phenomenal approach to personal persistence should be developed. It can be put this way:

Necessarily, person  $P$  at time  $t$  is identical to person  $P^*$  at time  $t^*$  if and only if  $P$  has the same capacity for consciousness as—i.e., is CC-Continuous with— $P^*$ .<sup>16</sup>

Now, thus far I’ve described, in rather abstract terms, what the capacity for consciousness is. And I have described, also in rather abstract terms, what makes for continuity in this capacity. I have not, however, described which specific structures—which parts of the brain, mind, or soul—are necessary for that capacity. That is, I have not described the underlying *ontology* responsible for the capacity for consciousness. And so I have not taken any stand on issues regarding the underlying ontology of personal persistence.

*This is precisely as it should be.* For we are ignorant of what’s required, ontologically, for the capacity for consciousness. We are ignorant of what kind of stuff could give rise to consciousness and could thus ensure one’s persistence. I, for instance, have no idea which parts or interrelations between parts—whether in the brain, or in some immaterial mental substance, or whatever—are necessary for consciousness. For I have no idea what it depends on, ontologically. And it would be foolish of me to pretend otherwise. For this reason, I believe that the correct position with respect to our underlying ontology is *agnosticism*.

You might think I am overstating our ignorance. You might think that it’s obvious that the *brain* is responsible for consciousness. After all, there are strong

---

<sup>16</sup> Notice that this view rules out “gappy” existence; that is, it implies that people cannot come back into existence after having gone out of existence. I embrace this implication. But it does yield some (apparently) counterintuitive results. For example, suppose that, at time  $t$ , Aurora suffers a traumatic brain injury and loses the capacity for consciousness. But then suppose the brain heals so that, by time  $t^*$ , a person—call her ‘Aurora\*’—replete with the capacity for consciousness, wakes up as if from the dead. Aurora and Aurora\* are not CC-continuous. So, at stated, my view entails that they are not identical. But is that right? For some, this may seem like the wrong result. So they may wish to adjust my view as follows. The first step is to add a clause to my characterization of CC-Connectedness saying that CC-Connectedness may obtain between people at non-adjacent times if some to-be-specified causal dependence and/or overlap in parts connects their capacities for consciousness. Then the next step is to say that a chain of CC-Connections needn’t be *unbroken* in order for CC-Continuity to obtain. So Aurora and Aurora\*, for example, may be CC-Continuous. Gappy existence is thus tolerated. I leave further details to believers in gappy existence.



correlations between neural activity (or inactivity, as in the case of sleep or brain damage) and consciousness. So you might think that this settles the matter.

But it doesn't. And here's why. First, that there are correlations between consciousness and neural activity does not imply that the former *is* the latter, or even that the former is *constituted* by the latter. These correlations are consistent with a wide range of ontological views, even including substance dualism.

And even if the correlations between consciousness and neural activity did give us reason to believe that neural activity constitutes our consciousness, or is directly responsible for it, this still wouldn't settle the matter. For even if neural activity is responsible for our consciousness, it may still be *possible* for experiences to occur in some other, non-biological medium. It might be possible for a person's brain to be gradually replaced with silicon chips without her ever losing her capacity for consciousness. So even if we grant that the brain is what's directly responsible for our consciousness, this doesn't settle what's necessary, ontologically, for the capacity for consciousness. Thus, it doesn't settle what's necessary, ontologically, for personal persistence.

And even if we suppose that consciousness is necessarily biological, or that it is necessarily instantiated in the brain, still, that does not settle the issue. For it doesn't settle which *specific* structures or neural configurations are necessary for the capacity for consciousness. Is the whole brain required? Or is it just the cerebrum? Or do we only need certain portions of the brain? Could the biological material that makes up the brain support consciousness in another arrangement? Or is there something special—and indeed *essential*—about the current layout of the brain?

We have no idea. Perhaps someday we will. But today is not that day. And we shouldn't pretend otherwise. Given our current state of ignorance, we should not take a stand on the ontology of the capacity for consciousness. Nor should we take a stand on the ontology of personal persistence. We should remain agnostic.

So where does that leave us? My suggested development of the phenomenal approach to personal persistence leaves us with an (abstract) characterization of the capacity for consciousness and continuity in that capacity. This is a significant advance, though of course I've not answered every question in need of an answer. I certainly won't pretend to do that here. There are some loose ends—that is, questions not yet settled by theory or other theories within the phenomenal approach—that are longstanding issues, problems, or puzzles that beleaguer most (sometimes *all*) theories of personal persistence, but that I believe adopters of the phenomenal approach are in a unique position to address.<sup>17</sup> The *puzzle of fission*, which is generated by possible cases in which a person appears to be split amoeba-like into two distinct people, is one example. I believe that defenders of the

---

<sup>17</sup> There are other issues that defenders of the phenomenal approach are in a unique position to address, but that have already been addressed. For example, Dainton and Bayne (2005) show how the phenomenal approach makes sense of certain puzzling (and paradoxical) *mind transfer* cases, such as those introduced by Bernard Williams (1973).

phenomenal approach can appeal to the essential unity of consciousness to argue that fission is, in fact, impossible.<sup>18</sup> But defense of that claim is for another day.

As with almost any approach to personal persistence vying for legitimacy, the phenomenal approach requires further development. New implications and objections will inevitably arise. My goal here has not been to refute all competing approaches, nor has it been to show that the phenomenal approach is completely satisfying in every way. My goal has been to lay out a new argument for the phenomenal approach and to offer some developments of it so as to deal with the major obstructions brought out here and elsewhere. Whether the psychological/biological two-party system will ever be broken up remains to be seen. I'll be happy to have challenged it, and perhaps to have swayed some of those who were previously undecided.

**Acknowledgements** Thanks to Ross Cameron, Nina Emery, Mark Fiocco, Brie Gertler, Harold Langsam, Trenton Merricks, Paul Nedelisky, Adam Pautz, Nick Rimmel, Tomoji Shogenji, Jack Spencer, and an anonymous referee for helpful comments and/or conversations on various versions of this paper.

## References

- Bayne, T. (2010). *The unity of consciousness*. Oxford: Oxford University Press.
- Bayne, T., & Chalmers, D. (2003). What is the unity of consciousness? In A. Cleeremans (Ed.), *The unity of consciousness: Binding, integration, and dissociation* (pp. 23–58). Oxford: Oxford University Press.
- Carruthers, P. (2011). *The opacity of mind*. Oxford: Clarendon Press.
- Dainton, B. (2008). *The phenomenal self*. Oxford: Oxford University Press.
- Dainton, B., & Bayne, T. (2005). Consciousness as a guide to personal persistence. *Australasian Journal of Philosophy*, 89(4), 549–571.
- Duncan, M. (2014). A challenge to anti-criterialism. *Erkenntnis*, 79(2), 283–296.
- Duncan, M. (forthcoming). A renewed challenge to anti-criterialism. *Erkenntnis*.
- Duncan, M. (2015a). I think, therefore I persist. *Australasian Journal of Philosophy*, 93(4), 740–756.
- Duncan, M. (2015b). We are acquainted with ourselves. *Philosophical Studies*, 172(9), 2531–2549.
- Duncan, M. (2018). Subjectivity as self-acquaintance. *Journal of Consciousness Studies*, 25(3–4), 88–111.
- Duncan, M. (forthcoming). The self shows up in experience. *Review of Philosophy and Psychology*. <https://doi.org/10.1007/s13164-017-0355-2>
- Evans, G. (2001). Self-identification. In A. Brook & R. DeVidi (Eds.), *Self-reference and self-awareness* (pp. 95–142). Philadelphia: John Benjamins Publishing Company.
- Foster, J. (1991). *The immaterial self*. London: Routledge.
- Gertler, B. (2011). *Self-knowledge*. New York: Routledge.
- Howell, R. (2006). Self-knowledge and self-reference. *Philosophy and Phenomenological Research*, 72(1), 44–70.
- James, W. (1952). *The principles of psychology*. Chicago: William Benton.
- Lewis, D. (1976). Survival and identity. In A. O. Rorty (Ed.), *The identities of persons* (pp. 17–40). Los Angeles: University of California Press.
- Locke, J. (1689/1975). An essay concerning human understanding. In P. H. Nidditch (Ed.), Oxford: Oxford University Press. <https://philpapers.org/rec/NIDJLA-2>.
- MacIntyre, A. (1984). *After virtue* (2nd ed.). Notre Dame: University of Notre Dame Press.
- Noonan, H. (1989). *Personal identity*. New York: Routledge.
- O'Brien, L. (2007). *Self-knowing agents*. Oxford: Oxford University Press.

<sup>18</sup> See Bayne (2010) and Bayne and Chalmers (2003) for good discussions of what the unity of consciousness is.

- Olson, E. (2007). *What are we?*. Oxford: Oxford University Press.
- Prinz, J. (2011). The sensory basis of cognitive phenomenology. In T. Bayne & M. Montague (Eds.), *Cognitive phenomenology* (pp. 174–196). Oxford: Oxford University Press.
- Schechtman, M. (2011). The narrative self. In S. Gallagher (Ed.), *The oxford handbook of the self* (pp. 394–418). Oxford: Oxford University Press.
- Shoemaker, S. (1968). Self-reference and self-awareness. *Journal of Philosophy*, 65(19), 555–567.
- Shoemaker, S. (1985). Personal identity: A materialist account. In S. Shoemaker & R. Swinburne (Eds.), *Personal identity* (pp. 67–133). Oxford: Basil Blackwell.
- Taylor, C. (1976). Responsibility for Self. In A. O. Rorty (Ed.), *The identities of persons* (pp. 281–300). Los Angeles: University of California Press.
- Thompson, J. J. (1987). Ruminations on an account of personal identity. In J. J. Thompson (Ed.), *On being and saying: Essays for Richard Cartwright* (pp. 215–240). Cambridge: The MIT Press.
- Unger, P. (1990). *Identity, consciousness and value*. Oxford: Oxford University Press.
- Van Inwagen, P. (1990). *Material beings*. Ithaca, NY: Cornell University Press.
- Velleman, J. D. (2006). The self as narrator. In *Self to self: Selected essays*. Cambridge University Press, pp. 203–223. <https://philpapers.org/rec/VELSTS-6>.
- Williams, B. (1973). *Problems of the self*. Cambridge: Cambridge University Press.

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.