

Paper prepared for
The 6th Euroacademia International Conference
Identities and Identifications: Politicized Uses of Collective Identities

Florence, Italy
22 – 23 June 2017

This paper is a draft
Please do not cite or circulate

Self-made?

The limits of pure internalism for moral personhood

Simon Cushing
Department of Philosophy
University of Michigan-Flint
Flint, MI 48502-1950
simoncu@umflint.edu
(810) 875-5507

Abstract:

In the debate over free will, metaphysical libertarians have argued that only they can account for the intuition we have that it is not enough that I do what I want for my action to be free, but that I must be the shaper of the self that has the desires that motivate my choice. That is, unless I have, in Robert Kane's terminology, performed some *self-forming* actions, my behavior is ultimately unfree and as such ultimately removed from the sphere of moral praise or blame just as much as the movements of a programmed robot would be. One could call this a kind of internalism about the source of (at least part of) the self.

A similar internalism about the self is evident in a recent response to challenges to the Lockean tradition of personal identity. Marya Schechtman has argued that, in an important sense, we must be *self-constituting* in order for our persistence through time to carry the normative weight in matters of moral responsibility and prudential concern that we intuitively take it to.

Both of these philosophical projects, while emerging from very different bodies of philosophical literature, share the paradoxical idea that the self must be in some sense self-producing if we are to be beings with moral powers and the locus of moral evaluations. In this paper I explain the concerns that drive each author into apparent paradox, but argue that while the internalist view that results most centrally captures "what matters" (to use Derek Parfit's phrase), neither theorist has delivered on their promise, and they have not given accounts whereby the self truly structures itself. I conclude that this is not to be regretted, because the worries that motivated their attempts do not require some kind of bootstrapping.

Key Words:

Person, free will, responsibility, Kane, Schechtman

In a 1998 paper Eric Olson argued that "There is No Problem of the Self," because "the matters discussed under the heading of 'self' are *so* various that no one can seriously say that they are all about some one thing, the self."¹ The two writers I will focus on appear to give support for his claim. Robert Kane and Marya Schechtman are engaged in different debates, but what they share in common is the idea that a prerequisite of moral agency is that some core aspect of the agent be either "self-formed" (Kane) or "self-constituted" (Schechtman). Kane argues that only a will that is self-formed can be genuinely free. Schechtman sees self-constitution as what explains "what makes persons suitable targets of forensic judgments"² (like blame and praise). However, the upshot of both approaches is that, for us to be morally responsible beings, we must in some way have had a direct role in creating our own selves. This very notion, of course, is paradoxical. In the words of Susan Wolf "literal self-creation is not just empirically, but logically, impossible."³ Both Kane and Schechtman attempt to find ways around the paradox and are ingenious in their efforts, as we shall see. I argue, however, that what each ends up arguing for falls short of deserving the name. Finally, I conclude that self-creation is not just unattainable, it is not worth seeking.

Imagine if you will that a prominent philosopher is giving a speech to a huge rapturous audience, as they do, and I step from the shadows and brutally gun him down. Factors that could lessen or outright undermine my responsibility for this action include: insanity, hallucinations (I thought he was a giant lizard creature about to feast on us all), or direct control over my mind or body by sinister anti-philosophy forces. However, if I both wanted to kill him and rationally planned to do so, then clearly I am responsible, and my act was freely done. What this appears to illustrate is that for an act of mine to be free it must be that I alone was in control of that act, or that it is my self or character that is the sufficient cause of it. Such control requires that the universe behaves in a lawlike, predictable manner.

Classic "compatibilists" like Locke, Hume and Mill, claimed that considerations like this show that *determinism*, the assertion that every event is necessitated by prior events in accordance with the laws of nature, is not only *compatible* with the kind of free will associated with moral responsibility, but *a fortiori*, that you could not have such responsibility *without* determinism. In opposition to this view, are the metaphysical *libertarians*, who insist that true freedom cannot be determined, because for an agent's performance of an action to be truly free, it must be that she

could have done otherwise. Compatibilists, however, argue that we don't want our actions to be that unpredictable, because the absence of control is not freedom but randomness. To illustrate this compatibilist point, imagine there was a particularly savage serial killer (the Deranged Disemboweler) on the loose in your city. Hearing of this killer's gruesome crimes, your mother calls you up and asks, in all earnestness, "Are you the Deranged Disemboweler?" When you respond in outrage, she seems surprised and says, "But don't you want true freedom? And that requires that all options be open to you, no matter what your character."

Control over my actions is clearly necessary for free action. Is it sufficient, however? What if, instead of intervening between character and action, those aforementioned philosophobes have instead played the long game and have molded my character so that I am such that I only want what they want. Killing the philosopher is thus an act that results from my actual character and supposedly meets the classic compatibilist criteria for being freely done. But it seems wrong to hold me responsible for an act that results from desires and beliefs I had no say in acquiring. But then, if *that* is a barrier to freedom, then it looks like none of our actions are free, because surely our characters or selves are ultimately entirely the result of forces outside of our control: if not a sinister scientist, then a combination of nature and nurture, genetics and environment. As Peter van Inwagen, who calls a version of this problem "the consequence argument," puts it, if determinism is true than none of our actions are "up to us".

It is to address both compatibilist and libertarian objections that Robert Kane introduces his notion of "self-forming actions," so called because "they would be the actions in our lives by which we *form* our character and motives... and make ourselves into the kinds of person we are".⁴ Kane sees these actions as providing two vital elements to a freedom worth the name. First (compatibilist), they provide ultimate responsibility by making the agent herself the cause of the nature that produces her actions. Her character may even *determine* her actions: Kane agrees with Daniel Dennett's example that Martin Luther's statement, "Here I stand; I can do no other"⁵ can be taken literally without it robbing his action of moral worth. Second (libertarian), they address the consequence argument, because SFAs are *undetermined* actions because:

if there were no such undetermined SFAs in our lifetimes, there would have been nothing we could have ever voluntarily done to make ourselves different than we are—a condition that I think is inconsistent with our having the kind of responsibility for being what we are which genuine free will requires.⁶

Thus Kane means "self-forming" in two ways: the self is both the active former and the passive formee—hence the apparent paradox and the need for explication. Kane's account rests heavily on what he calls "plural voluntary control." In a nutshell this consists of the following claims:

1. An element of randomness in a person's action does not necessarily undermine its attributability to an agent for purposes of praise or blame. For example, if I have a genuinely random tremor in my arm that makes it uncertain whether or not I will succeed in hitting the famous philosopher when I shoot, if I *do* succeed, I am fully responsible.
2. Randomness is a necessary feature of SFAs lest they simply be determined by the universe and not "up to" the agent.
3. However, the randomness in SFAs will never result in outrageously out-of-character actions like deranged disemboweling, because the set of possible actions from which one will be randomly selected only includes actions the agent would endorse.
4. Moreover, the randomness is *as a result of the conflict* among the competing motivations within the agent's will, and thus is not an element introduced from outside the agent.

Consider a person who facing a dilemma of whether to lie or tell the truth:

There is a tension and uncertainty in our minds at such times of inner conflict... in short, a kind of stirring up of chaos in the brain that makes it sensitive to microindeterminacies at the neuronal level. As a result, the uncertainty and inner tension we feel at such soul-searching moments of self-formation is reflected in the indeterminacy of our neural processes themselves. What is experienced phenomenologically as uncertainty corresponds physically to the opening of a window of opportunity that temporarily screens off complete determination by the past.⁷

According to Kane, if this person ends up telling the truth, it will be a self-forming action that can lead to a virtuous character down the line that will determine her to tell the truth in all cases.

Kane's view is certainly inventive, but it falls short of true self-formation for a number of reasons. The most fundamental problem is that self-formation does not necessarily even occur. There are two aspects to my complaint here. First, it has to be the case that the self that is acting *already has* the aspect that the agent acts on. So, for Sennett's truth teller, it has to be that she already is inclined to truth-telling. So that aspect of her character is already in place before any self-forming act can occur, and thus of necessity is not her creation. Now, in response Kane could say that the "formation" is like that done by a sculptor, who chisels away extraneous bits of an already-existing lump of stone. But then the second aspect to my complaint is that what he describes does not necessarily permanently close the door to more lying. If telling the truth *just once* creates the permanent habit of veracity, then that itself is because

of a contingent factor of our truth-teller, that she is instantly impressionable. Another individual, whose nature is such that he is far less quickly set in his ways, could experience *another* undetermined SFA the very next choice he makes and this time choose to lie. So, ironically, this person would have twice the number of self-forming acts and yet have an unsettled character at the end of it.

A further qualm I have with calling these events self-forming acts is that that gives the impression that they are intentional (something that is compounded by the use of phrases like “soul-searching”). But it is clear that whether one lies or tells the truth is not really “up to” the agent. In cases like these, which of the competing tendencies in a person’s self wins out is as a result of a genuinely random event (assuming the view of quantum mechanics that is non-deterministic), and thus open to the classic compatibilist complaint of arbitrariness. Now, Kane has set up the case so that this only happens when the agent is genuinely indifferent between the outcomes (which leads to a third complaint which I shall get to in a second), but what that means is that the best you can say is that I intend a disjunction: I intend to *tell the truth or lie*, and I am content that randomness settles which one results. But these are supposed to be important, character-forming events, where the path not taken is closed *permanently*, and clearly I did not intend that, as it *was* a path that I was as invested in as the one chance sent me down.

The third complaint is that self-forming acts are likely to be rare, because they require genuine indifference between genuinely morally weighty options. (I do not think randomness settling my choice between broccoli and cauliflower as a side dish would count as a Kaneian SFA.) But that has the paradoxical implication that the kind of person most likely to have more SFAs is someone indifferent between virtues and their corresponding vices. Imagine somebody whose life was filled with nothing but such SFAs: I think we would call such a person aimless rather than admire her for the amount of genuine free self-formation she indulged in. Finally, imagine two such beings in parallel possible worlds. In one world, randomness dictates that the virtuous path is taken every time. In the other, the vicious path always wins out. As a result, in the first world a pure self is formed, while in the second a corrupt one. But can we truly say that the saint of world one deserves credit for her saintly self, and the sinner of world two deserves blame? It seems clear to me that calling either’s self “self-formed” is profoundly disingenuous.

Let us now turn to Schechtman, who arrives at her notion of “self-constitution” by a different route, the debate over personal identity dating back to the work of John Locke.

When I was a teenager, going to high school in the southwestern English city of Exeter, I ran into a person whom I took to be somebody with whom I’d been friends with as a young child in London, whom I hadn’t seen for many years. Embarrassment ensued, because, as it turned out, he did not share this view. But how to tell which of us was right?

In one of the most pored-over discussions in all of early modern philosophy, Locke distinguishes between criteria of identity for “the man” (by which he meant individual human beings) and “the person”. The former has the same identity conditions as any organism, something that biology can settle. However, whether or not this human being remains the same *person* is a distinct matter. Locke offers the following definition of person:

a thinking intelligent being, that has reason and reflection, and can consider itself as itself, the same thinking thing, in different times and places; which it does only by that consciousness which is inseparable from thinking... and as far as this consciousness can be extended backwards to any past action or thought, so far reaches the identity of that person.⁸

In theory it is possible that I was correct that the Exeter teen was the same human as the London child, but he was correct that he was not the same person. If Exeter-teen had no consciousness of being a child in London who was my friend, then they were two different people, even if they were housed in the same human.

But surely, it can’t be that we’re *both* right – either he was my friend or he wasn’t. Which takes precedence, the human or the person? One clear reason that Locke appears to favor person over human is that, according to him “person” is a *forensic* term “appropriating actions and their merit; and so belongs only to intelligent agents, capable of a law, and happiness, and misery”.⁹ That is, for matters concerning responsibility, blame, praise, regret, rational anticipation and self-concern, it is sameness of person that settles the issue. So, say, the London child made me a solemn promise and the question is whether the Exeter teen should be held to that promise, then him being the same human is irrelevant. Locke’s preference was very influential over the literature that followed, culminating in contemporary neo-Lockeans who are typically grouped together under the heading “psychological continuity theorists” because their answer as to whether or not Exeter teen is the same person as London child is that it depends on whether there is a chain of psychological connections of sufficient quantity and quality traceable back from the former to the latter.¹⁰

All psychological continuity theorists offer a *reductive* theory of personal identity, in that they claim that there is nothing about individuals at different times being one and the same person beyond their having the right kinds of psychological relations. Will those relations alone support the forensic weight that Locke intended personhood to bear? All sorts of ingenious thought experiments have been used to show that the theory cannot, in fact, bear this weight. Source of a number of these is Derek Parfit, and I will paraphrase his “branch line case” as follows. Imagine that we live in a world where the transporters of *Star Trek* fame actually exist. You can go into a booth and press a button and your body is systematically disintegrated particle by particle, while simultaneously the information about each particle is analyzed and transmitted to a booth in your chosen destination where an identical body is built up according to the transmitted blueprint. The being who steps out of the destination booth is completely indistinguishable from the original, down to the smallest physical particle, and down to the most trivial personal and psychological characteristic. Is it you? It certainly thinks so, and by a pure psychological criterion it would certainly appear to be, because it has as complete a psychological connection to you as you do to yourself a second ago. In any case, let us assume that people use these transporters all the time to get around and regard them as great modern conveniences. However, now imagine that you decide that Florence is boring and you would like to see what all the fuss is about Flint, Michigan. You get into the transporter booth as usual, type in the code for Flint, and press the “TRANSMIT” button. However, this time nothing appears to happen. Then, with horror, you see on the scanner screen a view of the inside of the booth in Flint and see that *there you are*, whistling happily and just about to exit into the Michigan sunshine. You grab the handle to get out of the Florence booth, but it’s locked. Then the screen flashes the following message: “disintegration unit temporarily disabled. Error fixed, coming back online in 3, 2, 1...” At this moment you realize the horrible truth: you have *never* survived transport. The real you died on using the transporter for the first time, and what survived was an imposter, who merely *believed* they were you, just like the imposter now exploring the multitudinous delights of Flint, MI, while you face your imminent destruction. The moral of the story appears to be: even 100% psychological continuity is not identity.

Further problems can also be raised: suppose the original is *not* disintegrated, and there are now two of you. Or, suppose the original is disintegrated, but copies are sent to several cities at once. As each of these rapidly acquires psychological differences, and plainly are not one and the same person, how can they all be identical to the original? But saying that only one of them is appears arbitrary, so must we say that none is, even though all have the requisite psychological continuity?

There are ways to adapt a psychological continuity account of identity to save it from these problems: you can stipulate that psychological continuity is not sufficient, and that there must also be non-duplication, and that there must be some continuity of matter between the organism currently housing the person and the organism that did so in the previous instant. However, as Parfit himself argued, these seem arbitrary. Adding to the arbitrariness, as “psychological continuity” is a matter of degree but identity is not, it looks like we have to draw a line somewhere and say “if Exeter teen has x number of psychological connections with London child, then they are one and the same person, but if Exeter teen has $x-1$ connections, they are two different people,” where this could be a difference of one memory. Such considerations led Parfit to contend that “identity doesn’t matter” and what matters is simply psychological connectedness. Instead of caring about whether or not there will be a future person who is identical with me, I should think in terms of degrees of survival, and care about future person (or persons) to varying degrees depending on the number and degree of psychological connections.

This is too radical, however. It seems to require that, should I manage to break out of the transporter booth in Florence, I should care roughly the same about the survival of the clone in Flint as I do about the “original” in Florence, and I just wouldn’t. So critics of Parfit can either take the position that he is right that identity doesn’t matter, but go further and say that psychological continuity doesn’t matter either, or go the other way and argue that the fact that psychological continuity doesn’t capture identity is a knock on psychological continuity rather than identity. This is the view that Marya Schechtman takes. Her diagnosis of what is wrong with psychological continuity theories is their *reductionism*, which entails that “they frame the question of personal identity as a question about the reidentification of an object”.

This requires that they define identity in terms of relations between time-slices, which in turn requires thinking of our psychological lives as something that can intelligibly be described in terms of momentary units, distorting the very features that make the psychological approach appealing in the first place.¹¹

No theory that is a response to a reidentification question can capture the forensic features that Locke correctly asserted as being what matters in identity, Schechtman charges. A satisfactory theory will instead begin by addressing the *characterization* question. The difference between the two is that the former is the question an amnesiac would ask, whereas the latter is the question someone wishing to settle an identity crisis would ask. That is, I answer the characterization question about myself by delineating the limits of my *self*. What is me and what is not me? An account that can settle this question will make the reidentification question moot.

It is not because I am like someone who took an action or worked some number of hours that I am responsible for that action or entitled to compensation; it is because the experiencing subject suffering the consequences (or enjoying the

rewards) is the same subject who took the relevant action. It is not because someone in the future will be like me that I care in a particular way about her experiences, but because I expect to experience them myself.¹²

Schechtman's candidate of an account that can do this she calls the *Narrative Self-Constitution View*. The way this view avoids the failings of the psychological continuity views is to give the person an active role in piecing together the elements of her self.

According to the narrative self-constitution view, the difference between persons and other individuals... lies in how they organize their experience, and hence their lives. At the core of this view is the assertion that individuals constitute themselves as persons by coming to think of themselves as persisting subjects who have had experience in the past and will continue to have experience in the future, taking certain experiences as theirs. Some, but not all, individuals weave stories of their lives, and it is their doing so which makes them persons. On this view a person's identity (in the sense at issue in the characterization question) is constituted by the content of her self-narrative, and the traits, actions, and experiences included in it are, by virtue of that inclusion, hers.¹³

Schechtman sees in her view a return to two aspects of Locke's account of person that have been left behind in the development of neo-Lockean psychological continuity theories. First, Locke is clear that the *same consciousness* has to exist not just across my various parts *right now* (so that my finger is part of me but my clothes are not) but across my entire history.¹⁴ Only a view that is *diachronically holistic* can capture why I care more about an old, future version of me that has abandoned many of my current values than I do about someone other than me who shares all of them. Now, psychological continuity theorists *thought* they were honoring Locke's view by focusing on memory, but somehow along the way they lost the idea of a single pre-existing conscious subject being a prerequisite for the uniting of various parts into a whole.

Second, Schechtman would note that Locke requires that a person have the ability to "consider itself as itself...in different times and places." This requirement motivates the *self-constitution* element of her view. Putting the self-constitution together with the diachronic holism and adding the extra element that the self that is created must have a narrative structure, and you get her view. In sum:

The formation of an identity-constituting narrative creates a single, temporally extended subject of experience, and any two actions or experiences attributed to the same person by this view are necessarily attributable to the same subject of experience.¹⁵

While most commentators have focused on the narrative element of the view, in this paper I am concerned with the *self-constitution* part. As with Kane's "self-forming," that very phrase smacks of paradox, as if the author of a book were herself a fictional character within the book.

I put this to Schechtman in an interview and she said that she liked Christine Korsgaard's response to such puzzlement, which was to say that one might just as well look at a giraffe eating and metabolizing and ask bemusedly, "how is that giraffe creating itself?"¹⁶ In Schechtman's and Korsgaard's terms "organisms are self-constituting beings" and as nobody questions this, nobody should question the idea of persons being the same. However, this is disingenuous. Schechtman's self-constitution seems to entail the subject of experience being created by itself acting on its own experiences. Thus we have the question of how the experiences have come into being if they are required to make the thing that *has* experiences. Surely this is different from a giraffe eating food that it does not itself create, whereby the giraffe's internal structure is such that the nutrition extracted is turned into cells that become part of the giraffe. But this is *self-perpetuation* rather than *self-constitution*. The giraffe is created by other giraffes.¹⁷

Let us press this giraffe analogy, however. Just as parts of the giraffe work to make other parts of the giraffe using non-giraffe fuel, maybe we can think of an experiencing self that is provided with perceptions by an external world impinging on her senses which instantly become infused with her sense of her place in the world, creating memories that are, as they are stored, part of a narrative. This doesn't quite accommodate future experiences, but when I envision things happening to me in the future, I don't see that as an event that might as well happen to a stranger, I put myself in the shoes of the experiencer and thereby make it matter to me in the way that Locke requires.

Notice that on this picture we have two separate entities: we have what we might call *the self*, which is the subject of experiences, and which exists in the moving now,¹⁸ but we also have *the person*, which is the four-dimensional object that the self constructs. It is as if the self is the spider and the person is its web. However, this has lost the crucial appeal of *self-constitution*. Recall that the point of Schechtman's account was to solve the problem of the psychological continuity theories that they did not explain why (say) London child should identify with Exeter teen, or Exeter teen feel responsible for the actions of London child. If the self is a separate entity from the person, then surely one can ask "which is more essentially *me*?" By Lockean standards it would have to be the self, as this is the conscious entity. But then one can object, as Galen Strawson has to Schechtman, that it is possible to be an *episodic*, a being who lives in the moment and does not identify with the experiences of prior or future selves.¹⁹ If Strawson is right then it looks like one does not need a Schechtmanian "person" for the forensic qualities Locke attributed to it,

because he also argues that episodics can be just as responsible and ethical as “diachronics,” or those beings who view themselves as extended in time beyond their current direct consciousness.²⁰

If we are going to avoid separating the momentary self from the diachronic person, there are two possibilities, where we reduce one into the other. Let us take first the view that it is the person that is primary and the self that is only a *slice* of that being. This view might be compared with that defended by the philosopher David Lewis, who argues that we are 4-dimensional creatures that extend not just in space but temporally from our beginnings (whether that be conception or the start of brainwaves or whatever) to our end (whether that be cortical death or decomposition or whatever), so our conscious self at any moment is just a temporal part of our greater, diachronic whole.²¹ Something like this seems to follow from this passage in Schechtman’s first book:

Concern for the future...is an ongoing, active orientation that creates a kind of experience that is not present without it. The subject worrying about his future is a narrative self and not some particular moment of this self, so the effects of self-concern do not consist only in the fact that at one moment (or even at each moment) a particular *anticipated* future changes a person’s present. Instead, the formation of a narrative brings into being a temporally extended subject who has this concern for her whole self. By the time someone is in the position to worry about the future he is already more than a momentary creature.²²

However, taking this picture too literally is hard: we are not beings that are conscious in the same way of all the events in our history. We experience time as a flow. Whether or not the past and future exist in the same way as the present, as so-called B-series theorists maintain, or only the present is real, as the A-series view holds, it is impossible to deny that we only have sensory access to the present. Moreover, Schechtman herself, in a recent talk, acknowledges that “we cannot directly experience the past in the way we do the present”.

The alternative, then, is to have the momentary self subsume the diachronic person. How would this work, absent some kind of Leibnizian monad? One suggestion builds on Schechtman’s long held insight that a view of oneself as extended in time can color one’s present experience.

She argues that it is artificial to suggest that somehow memories can exist apart from the person who has them as if neutral. Instead, my experience of, say, a dinner with friends, is colored by my knowledge of them and history with them, and is phenomenologically different from a memory that would be created by somebody in exactly my physical location, but without my history. This lays the groundwork for the idea that an extended self can in some sense be incorporated into current consciousness, by taking into account one’s sense of one’s diachronicity. As illustration, she quotes George Eliot in *Middlemarch* describing Lydgate’s ability both to be in the grips of an infatuation and yet view that same infatuation from some future point at which he knows it will have passed:

Strange that some of us, with quick alternate vision, see beyond our infatuations, and even while we rave on the heights behold the wide plain where our persistent self pauses and awaits us.²³

In a recent talk, Schechtman has gestured at developing this idea that the current momentary self can have more than one viewpoint (a feature that she cites Peter Goldie arguing is distinctive of narrative) and suggests that in switching between the viewpoints, and attempting to reconcile them, create within its consciousness the diachronic awareness she has argued Locke requires. (I think of this as analogous to the way our brain creates 3-d images by processing the differences between the views of each of our eyes.)

This view still seems to contain a division between the viewpoints: there is the “me-now” view, and there is the long view, as if from above. However Schechtman clearly believes that not only can we reconcile these two viewpoints, it is the very act of reconciling them that brings us into being as moral entities. She has yet to develop this view formally in print, but it is worth considering its strengths and weaknesses in outline.

First, this view wants to incorporate both the idea that we “contain multitudes” such that we can experience disconnect between past or future selves and our current one, but that we can also experience an overarching unity. Perhaps, by analogy, one could think of how one could work to reconcile two apparently conflicting commitments one has (say to pacifism but also opposing authoritarianism) by a mental process of starting from one and working towards the other and then doing it in reverse. Both of them, however, are core beliefs of one’s own. Similarly, one can take the viewpoint of a being very much in the present, but also step back from it and see the “bigger picture”. This is meant to explain the appeal of a view like Strawson’s but also overcome the flaw with psychological continuity theories that there is too great a disconnect between current and future selves to bear the moral weight of personhood.

To respond, I would like to return to the issue of self-constitution. It must be said that the experiencing self (in Strawson’s sense) does not create itself. It *has* experiences, it is not itself an experience. Whereas what *is* produced by the self is a narrative that, on this view, only metaphorically exists in four dimensions. Rather it exists in the consciousness (or possibly subconsciousness) of the experiencing self. But if this mental narrative could somehow be

copied and made available to another experiencing self, I wouldn't think that *I* had been duplicated. So clearly the narrative is *not* me, it is a product of me. As a corollary, I identify with a future self because I think that it is the same experiencing self. What that is is still a mystery. According to Locke it is neither body nor soul, but on a Strawsonian view it is likely that it is a particular organization of matter that has the right sophistication to be conscious. But then the narrative and the viewpoints seem irrelevant. If you convinced future-me that he was *not* the same experiencing self as me-now, but instead a self that had had somebody else's narrative implanted (as happened to the replicant Rachel in *Blade Runner*), then future-me would be perfectly justified in believing he did not have to honor any commitments made by me-now.

Furthermore, even assuming some omnipotent being assures us that me-now is the same experiencing self as future-me, it looks like there could be *rival* narratives held by each of those consciousnesses. For example, suppose I'm a hippy in the 60s but become a yuppie in the 80s. Looking back from the 80s my view of my youth is as a misguided dropout who hadn't got serious yet. But looking forward from the 60s, my view of my 80s self is as a sell-out who's been corrupted. Which is the correct narrative? There is no one single narrative that unifies us. Instead, whatever unity we have has to be explained by something other than a narrative, thereby undermining Schechtman's project.

What, then, are we to make of the Narrative Self-Constitution View? I think it contains important insights. It is certainly true that one's view of oneself as persisting through time, as continuing relationships with people one has encountered before, and as being at a certain stage on a life path that stretches into the future colors one's current experience, in both positive and negative ways, both of which can be seen in the sensation of nostalgia. It is also true, I think, both that one needs to answer the characterization question, and so doing will have forensic relevance. *Who I am* at my true core is important for issues of responsibility. Am I the really the person who assassinated the famous philosopher at the beginning of this piece? According to Schechtman, only if that experience is woven by me into my narrative. Perhaps the true insight is that it is not up to me if it is part of the events of the story, but what is up to me is whether a reader would see that as out of character. At any rate, it seems right to say that whether or not I have reformed is a matter of whether or not I would still recognize that behavior as in character or if I now renounce it, and perhaps the ability to do those things is a mark of moral competence. But we are all familiar with unreliable narrators, and indeed may suspect ourselves of being them to greater or lesser degrees. Self-discovery is at least as familiar as self-constitution, and a good deal less paradoxical.

Let us end with an overview. Both Kane and Schechtman see some measure of self-creation as essential for moral agency. For Kane, a self-formed will is one that provides an answer to the consequence argument and yet allows that Martin Luther be free while at the same time have no alternative to act as he does. For Kane, the only way that "the buck stops here" with the agent, and thus she can be rightly held responsible, is if she had a role in making herself. I have little time left to make my case, but I think this is misleading. The very capacity to look at oneself in the light of experience and judge whether one needs to chance is itself a moral power, and a sign that one is *already* a moral agent. But this power is bestowed by nature. And in response to the Deranged Disemboweler worry, I would say that it would indeed rob us of agency if our characters were shaped *by another agent*. For we would then be their puppets. However, nature itself has no agency, and if our characters are the result of genetics and unintended environment, then there is no such worry that our agency has been coopted.

For Schechtman, there cannot be a moral agent who does not view herself as a being who extends across time, and without this active incorporation of past and future into the present, there is no entity of which forensic judgments can justifiably be made. However, I believe that her claim can either be interpreted weakly, in which case it is acceptable, but does not deserve the name "self-constitution," or it must be interpreted as the Lewisian view but with the added false claim that the four-dimensional person is aware atemporally. I have argued that the second cannot stand. The first, weak claim, then, is simply that the experiencing self colors the present with the sense that it is a stage in a lifespan of finite duration. I share Schechtman's skepticism that one can be a true "episodic," as Strawson claims, and see *no* connection with the experiences of so-called former selves. But I side with Strawson in thinking that even someone with a very happy-go-lucky, live-in-the-moment attitude can be a moral agent rightly subject to praise or blame. I think it is an inescapable fact of human nature that we see ourselves in the past and future, on a par with love for our children or aversion to pain. It is bequeathed us by nature and not something we can self-create, nor should we claim the credit for so doing.

References

- Dennett, Daniel. 1984. *Elbow Room*. Cambridge, MA: MIT Press.
- Kane, Robert. 1996. *The Significance of Free Will*. Oxford: Oxford University Press.

- Kane, Robert. 1999. "Responsibility, Luck, and Chance: Reflections on free will and indeterminism." *Journal of Philosophy* 96, (5): 217-240.
- Kane, Robert. 2005. *A Contemporary Introduction to Free Will*. Oxford: Oxford University Press.
- Korsgaard, Christine. 2009. *Self-Constitution: Agency, Identity, and Integrity*. New York: Oxford University Press.
- Lewis, David. 1976. "Survival and Identity." In *The Identities of Persons*, edited by Amelie Oksenberg Rorty, 17-40. Berkeley: University of California Press.
- Olson, Eric T. 1998. "There is No Problem of the Self." *Journal of Consciousness Studies*, 5 (5-6): 645-57
- Schechtman, Marya. 1996. *The Constitution of Selves*. Ithaca: Cornell University Press.
- Schechtman, Marya. 2007. "Stories, Lives, and Basic Survival: A Refinement and Defense of the Narrative View." *Royal Institute of Philosophy Supplement* 82 (60):155-178.
- Schechtman, Marya. 2014. *Staying Alive: Personal Identity, Practical Concerns, and the Unity of a Life*. Oxford: Oxford University Press.
- Schechtman, Marya. 2015. "Interview by Simon Cushing." *Journal of Cognition and Neuroethics (Philosophical Profiles)*: 1-31.
- Strawson, Galen. 2005a. "Introduction." In *The Self?* edited by Galen Strawson, vi-xii. Oxford: Blackwell.
- Strawson, Galen. 2005b. "Against Narrativity." In *The Self?* edited by Galen Strawson, 63-86. Oxford: Blackwell.
- Wolf, Susan. 2002. "Sanity and the Metaphysics of Responsibility." In Robert Kane, ed., *Free Will*. Oxford: Blackwell.

Simon Cushing is Associate Professor of Philosophy at the University of Michigan-Flint. Yes, *that* Flint: Roger & Me and lead in the water. Since November 2016 he has lived in a state of creeping dread, occasionally interspersed with periods of denial.

¹ Olson 1998, p. 645.

² Schechtman 2014, 102.

³ Wolf 2002, 154.

⁴ Kane 2005, 130.

⁵ Dennett 1984, 133, Kane 1996, 38 and 2005, 81.

⁶ Kane 1999, 224.

⁷ Kane 1999, 224-5.

⁸ Locke, *Essay*, 2nd Ed., Chapter XXVII, §11.

⁹ *Ibid.*, §28.

¹⁰ Eric Olson has been at the forefront of a resurgence of the view that the *human* is the more basic entity. However, Olson also concedes that it is perfectly reasonable to care more about the persistence of a Lockean person than a human being that has become a non-person.

¹¹ Schechtman 2014, 99-100.

¹² Schechtman 2007, 166.

¹³ Schechtman 1996, 94

¹⁴ "[T]he narrative self-constitution view provides a unity of consciousness over time that is just as deep, and has implications just as important, as unity of consciousness at any given time." Schechtman 1996, 144.

¹⁵ Schechtman 1996, 149.

¹⁶ Schechtman 2015, 16. "No one is tempted to say: 'how can the giraffe make itself into itself unless it is already there?' The picture here is not of a craftsman who is, mysteriously, his own product. The picture here is of the self-constitutive process that is the essence of life. The paradox of self-constitution, in this context, is no paradox at all." Korsgaard 2009, 42.

¹⁷ And lest it be said that the *kind* "giraffe" is self-constituting, giraffes as a species emerged from non-giraffes by evolution.

¹⁸ This seems to be the way that Galen Strawson uses the term. "It is very natural for us to think that there is such a thing as the 'self' – an inner subject of experience, a mental presence or locus of consciousness that is not the same thing as the human being considered as a whole." Strawson 2005a, vi. Strawson regards each self of being of finite length, say 2-3 seconds, and thinks that the human being Galen Strawson has been controlled/inhabited by a string of them, like a string of pearls. This is the ninth version of the self that Olson (1998) considers.

¹⁹ Strawson argues that episodics can still recognize duties to honor commitments made by earlier selves because of a recognition that others regard him as the human being rather than the self. This claim has met with some skepticism.

²⁰ See Strawson 2005b.

²¹ See Lewis 1976.

²² Schechtman 1996, 156-7.

²³ George Eliot, *Middlemarch* (New York: Penguin), 182. Quoted in Schechtman 1996, 156.