“Spinozistic Selves”

Abstract

Spinoza’s *Ethics* promises a path for sweeping personal transformations, but his accounts face two sets of overarching problems. The first concerns his peculiar metaphysics of action and agents; the second his apparent neglect of the very category of persons. Although these are somewhat distinct concerns, they have a common, unified solution in Spinoza’s system that is philosophically rich and interesting, both in its own right and in relation to contemporary work in moral philosophy. After presenting the core of the problem facing Spinoza's action theory, I turn to his overlooked account of selves, one that can be illuminated by contemporary work on so-called deep-self theories. I then show how Spinoza’s distinctive account of selves prevents his action theory from collapsing into metaphysical incoherence, and conclude with an implication for Spinoza’s broader account of transformation.

Agents and Selves: Two Problems, One Solution

Spinoza’s *Ethics* promises a path for sweeping personal transformations: from bondage to freedom, foolishness to wisdom, weakness to empowerment, fear to love, passivity to activity, even corruptibility to eternality. Spinoza’s accounts of such personal transformations face two sets of overarching problems. The first concerns his peculiar metaphysics of action and agents; the second his apparent neglect of the very category of persons. Although these are somewhat distinct worries, they have a unified solution in Spinoza’s system that is philosophically rich in its own right and in relation to contemporary work in moral philosophy.

Spinoza's metaphysics of actions and agents provides the foundation for his accounts of moral improvement in the *Ethics*, but that action theory faces significant internal worries. Although some of the problems have been recognized by others, the general structure of the problem facing Spinoza's action theory has been overlooked. This has led interpreters to miss a promising parallel that Spinoza draws between his action theory and his equally striking causal theory, a parallel that points us toward his main, also overlooked, solution.

After laying out these problems and parallels in the first section, I turn to another instance of apparent neglect, this time within Spinoza's own ontology: the category of persons or selves. For all of Spinoza's calls to personal transformation, it is noteworthy that he never offers a formal account of selves, and he has been accused of outright *excluding* selves and associated first-person features (such as subjectivity, perspective, qualia, or even finitude *tout court*) from what he takes to be the true, *sub specie aeternitatis* account of the world.

However, I argue in section two that Spinoza actually has an interesting account of selves, one that can be illuminated by contemporary deep-self theories. I also contend that Spinoza's distinctive account of selves prevents his action theory from collapsing into metaphysical incoherence. I conclude with an implication of all this for Spinoza's broader account of personal transformation.

1. Actions and Agents

Spinoza outlines the core of his action theory in *Ethics* part III, which serves as the bridge between the metaphysics of parts I and II and the applied ethics of parts IV and V. His account is built on a contrast between actions and agents, on the one hand, and passions and patients, on the

other. Although he defines his main categories in causal and conceptual terms, I focus on the causal versions:

> I say that we *act* [*agere*] when something happens, in us or outside us, of which we are the adequate cause, that is (by 3d1), when something in us or outside us follows from our nature, [an effect] which can be clearly and distinctly understood through our nature alone. On the other hand, I say that we *are acted on* [*pati*] when something happens in us or follows from our nature, of which we are only a partial cause [IIId2].

> By affect I understand affections of the body by which the body's power of acting is increased or diminished, aided or restrained, and the same time, the ideas of these affections. *Therefore, if we can be* [*possimus*] *the adequate cause of any of these affections, I understand by the affect an action; otherwise it is a passion* [IIId3, emphasis original].

Before digging into the details, there is a small translation issue worth flagging. Modern English lacks a simple word for *pati*, the Latin counterpart of *to act* (*agere*). English translators have been forced to awkwardly state Spinoza's clean *agere*/*pati* distinction as a difference between 'acting' versus 'being acted on' (Curley 1985) or "being active' versus 'being passive' (Shirley 2002). But the main difference between these categories for Spinoza is *not* the difference between being a cause versus being an effect, as if an action is just something I *do* and a passion is just something *done to* me. Spinoza's distinction is entirely within the category of *being a cause*, and the main difference is between something I do *all by myself* versus something I do *with the contributions of others*, which can include past external causal influences on me. This is reinforced in other passages like IIIp3s and IVApp2, and Spinoza even uses Scholastic language

of *concursus*, or cooperative causing, to characterize the passion/pati category (IIp13s, G II/97). This clarification will become important later, so I will be explicit: the domain of an individual's actions and passions includes only those changes for which the individual is causally responsible to at least some degree.

These definitions suggest that an individual body or mind is an *agent* of a change just in case it is entirely causally responsible for its occurrence, in which case the change is an *action* performed by that agent. By contrast, we could say that an individual is a *patient* of a change if it is only partially responsible for its occurrence, in which case the change is a *passion* of that patient. But Spinoza denies that finite individuals are ever fully causally responsible for their effects (Ip28). It would follow that no finite thing is ever an agent, which suggests that being *fully* causally responsibility for an effect is an unachievably high standard for us.

Spinoza usually prefers scalar accounts whenever he can have them, part of what Don Garrett (2008: 19–20) calls Spinoza's 'incremental naturalism'. In the case of actions, Spinoza is often interested in the comparative degree to which an individual is active or passive with respect to a change (IIIp3; IIIp3s; IVPref2; Vp40), which suggests that his definitions of actions and agents could be construed in scalar terms as well. We could say that an individual mind or body is an *agent* of a change to the non-zero extent to which the individual is causally responsible for it; to that extent, the change is an *action* of that agent. By contrast, an individual is a *patient* of a change that it contributes to bringing about to the extent to which it is not responsible for its occurrence. (For ease of expression, I use mostly the non-scalar versions in this section, though the same problems can be expressed using scalar versions.)

Spinoza's focus in IIId3 concerns a special class of actions, namely *affective* actions. These are changes in an individual's degree of power for which the individual is causally

responsible. Although Spinoza is usually interested in such increases in power when he discusses moral transformations, an individual's actions need not be limited to only affective actions, since individuals can be causally responsible for more than just changes in their own power. The same is true for passions, even though we typically think of passions as purely internal states akin to emotions. But nothing in principle limits an individual's passions to affective passions, and Spinoza sometimes characterizes an individual's non-affective changes as actions and passions (IIIp1; IIIp3).

Spinoza's framework implies that actions and passions are subject-relative in various ways. For example, on the scalar version, an individual can be more or less of an agent of a change, depending on how much causal responsibility she bears for it. An individual can also be both an agent and a patient relative to different effects, and one individual might produce more effects by itself than another, in virtue of which it could be classified as more active than the other.

Those kinds of relativity seem innocuous. However, Spinoza appears to think that actions and passions exhibit another kind of variability that is much harder to understand. He suggests that a particular change that is a passion for an individual can also be an action *for the same individual*. Admittedly, Spinoza sometimes uses a subjunctive mood to make this point (IIId3), in which case maybe he means only that what is actually a passion for an individual could have been an action instead (Marshall 2012: 148–49). This would be less odd, even if counterfactuals are generally tricky to evaluate for Spinoza.

But Spinoza also uses the indicative mood, such as in IVp59: 'to every action to which we are determined from an affect which is a passion, we can [*possumus*] be determined by reason, without that affect' (see also Vp4s, G II/283). It is tempting to render even indicative

claims like this differently, perhaps as a claim about *types* rather than tokens of passions, or as a hidden subjunctive about what *could have* determined us (though, as a matter of fact, did not), or as an epistemic modal about what we can *know* about what determines us. For an unvarnished reading implies that a token change that is passion can also be an action for the same individual. But given the way Spinoza constructed his categories in IIId2–3, this would involve a strange variability in causal structure. An individual that is merely a partial cause of an event can also be, in some sense, the complete cause of that same event.

But how could that be? Facts about causal responsibility seem like invariant and binary facts about the structure of the world. Either an individual is the sole cause of a particular event or it is not; surely it cannot be both. This might seem like a sufficient reason to abandon the unvarnished reading, but I later show how to make good sense of Spinoza's surprising indicative. For now, I will let it stand as a potential worry, which I label the CAUSAL PROBLEM OF AGENCY:

> CAUSAL PROBLEM OF AGENCY: an individual can be both an agent and a patient with respect to the same change.

There were neighboring disputes among Scholastics about the individuation of actions and passions (Löwe 2018), but it was taken as obvious by those disputants that at least in cases of non-immanent changes, the agent and patient were different individuals. That non-identity is what generated the Scholastic debates in the first place, as it was not immediately clear how many events a given action/passion involved and in which of the subjects the relational accident(s) inhered. In fact, when Descartes endorses the consensus view that a token action and passion are one and the same event (Schickel 2011), he points out that this identity holds *despite* the obvious nonidentity of the agent and patient (Descartes 1985: 328). So Spinoza would be in unchartered territory here.

It gets weirder. Spinoza lays out a strategy for ethical improvement in part V, one that promises greater contentment, happiness, blessedness, and even a kind of salvation (Vp36s). Part of this strategy centers on converting one's passions into actions, a change that is supposed to be somewhat within our power: 'We can devise no other remedy for the affects which depends on our power and is more excellent than this' (Vp4s, II/283–84). Spinoza promises that by becoming more active, we thereby become more perfect (Vp40), more virtuous (Vp20s), more satisfied (Vp27), wiser (Vp42s), more blessed (Vp42), and more God-like (Vp17). In short, Spinoza thinks not only that a particular passion can become an action for an individual, but also that individuals have moral motivations to convert their passions into actions as much and as often as they can.

How do we accomplish this? Spinoza claims that 'an affect which is a passion ceases to be a passion as soon as we form a clear and distinct idea of it' (Vp3). That is, by forming a clear and distinct representation of a change for which an individual had been only partially responsible, the individual converts that change from a passion into an action. There are at least two glaring problems with this strategy.

For one, it seems like the relevant change—the event of which an individual had been only a partial cause—has already happened when she goes about converting it into an action. The representational conversion Spinoza proposes is backward-looking. He is not saying, 'Henceforth, be more causally responsible for new changes!' He appears to be claiming that individuals can now become more causally responsible for, more active with respect to, a change that has already occurred. But that seems to involve changing the causal structure of the past.

As with the CAUSAL PROBLEM, one might take this to motivate a less straightforward reading of Spinoza's claim in Vp3, but for now, I present it as another potential problem, the TEMPORAL PROBLEM OF AGENCY:

> TEMPORAL PROBLEM OF AGENCY: an individual can now change the causal structure of the past.

Ever vigilant, Jonathan Bennett (1984: 336) notices the apparent temporal upshot of Spinoza's account and calls *bullshit*:

> But really it is nonsense, for no one could possibly acquire an adequate idea of an event after it has occurred. If x now exists and is a passion in me, then its cause y was [partially] outside my body; so [the idea of y] was [partially] outside my mind, and thus [the idea of x] is inadequate in my mind. And that's that! I can no more make [the idea of x] adequate by bringing it about that the idea of (y) was inside my mind than I can become royal by altering who my parents were. The absurdity remains if we set aside adequate ideas and attend merely to the notion of turning a passion into an action. That means making a change in what the cause of the event was, and once the event has occurred it is too late for that.

It was unusually generous of Bennett to set aside the role of representations in this conversion process. For how can passions become actions through mere *representational* shifts? This points to a second glaring problem with Spinoza's proposal. His call to become more active and more powerful is not primarily a call to *do* (in our sense) more, to bring about more and different effects than had previously been brought about. He is not offering the advice of a liberal arts guidance counselor: seek out new opportunities! Take up kitesurfing, biology, Russian literature.

Rather, the process of becoming more active involves *representing* oneself and one's effects differently. In Vp3 (quoted above), Spinoza claims that by more clearly and distinctly representing an affective change, that change thereby becomes more of an action, and therefore more virtuous. As he explains in the corollary, 'The more an affect is recognizable [*notior*] to us, then, the more it is in our power and the less the mind is acted on by it'. Spinoza's advice is more like that of a therapist: think more clearly about yourself, and by changing how you represent yourself, you can thereby become more active and more of agent with respect to the very same changes.

But this is yet another strange proposal when combined with Spinoza's causal account of actions and agents. It implies that facts about causal responsibility for changes are sometimes sensitive to how the change is thought about, and it is unclear why the world's causal structure should be sensitive to such representational shifts. I call this the REPRESENTATIONAL PROBLEM OF AGENCY:

> REPRESENTATIONAL PROBLEM OF AGENCY: an individual can convert a passion into an action by representing it differently.

Previous interpreters have noticed some of these concerns and have offered partial solutions, though none have pinpointed what I think is the underlying structural problem facing Spinoza's account. For but one recent example, Colin Marshall (2012) addresses what I call the REPRESENTATIONAL PROBLEM by arguing that, according to Spinoza, an individual converts passions into actions by first representing them *as* passions and then her shifting attention to more general truths about that passion. (Marshall also provides a critical summary of previous attempts to answer Bennett's particular objections.) But while this provides a psychologically astute reading of Vp3, Marshall's account does not readily generalize to cover the more

metaphysical problems I have indicated, and it overlooks the key step of internalization that, as I argue in the next section, is at the heart of Spinoza's conversion strategy.

The underlying worry uniting these different agency problems is that in his metaphysics, Spinoza ties facts about actions and agents to the world's causal structure, but when discussing moral improvement, he implies that this causal structure can vary, depending partly on how it is represented. That definitely sounds odd, perhaps odd enough to motivate an alternative reading of those previously cited passages. But when the underlying problem is expressed this way, Spinoza aficionados might instead perk up. After all, the thesis that the world's causal structure can vary, depending on how it is considered, is a hauntingly familiar refrain from another part of Spinoza's metaphysics.

Here is a quick refresher for non-aficionados. Spinoza thinks there exists only one substance—call it 'God' or call it 'Nature'. This sole substance causes both physical and mental effects. Furthermore, each individual state of substance—what Spinoza calls a 'mode'—also causes both physical and mental effects. And yet, Spinoza thinks that only mental things can affect mental things and that no mental thing can have any physical effects (and mutatis mutandis for physical things). Put differently, Spinoza thinks that minds and bodies are entirely causally isolated from each other (IIIp2), despite the fact that each mind is identical to some body and vice versa (IIp7s; IIp21s). By simple substitution, it seems to follow that a mind both is and is not the cause of some physical change, which looks incoherent.

Spinoza's solution is to argue that whether a thing causes mental or physical effects depends partly on how that thing is conceived or considered. 'The modes of each attribute have God for their cause only insofar as [*quatenus*] God is considered under the attribute of which they are modes, and not insofar as he is considered under any other attribute' (IIp6). Spinoza

relativizes differences in kinds of causal structures—mental versus physical—to the attribute under which an individual is conceived. Conceived under the attribute of extension, an individual causes only physical effects; conceived as thinking, that same individual does not cause any extended effects. More abstractly, Spinoza preserves the identity of individuals across seemingly incompatible causal variations by (1) tying causal variability to attribute variability and (2) tying attribute variability to changes in the ways individuals are conceived. As Spinoza expresses the resulting picture, 'the thinking substance and the extended substance are one and the same substance, which is now comprehended under this attribute, now under that. So also a mode of extension and the idea of that mode are one and the same thing, but expressed in two ways' (IIp7s).

Although this compressed account of the mind-body case in Spinoza is controversial (see Newlands 2018: 42–55 for elaboration and defense), it points to an intriguing parallel with the agency case. For Spinoza also ties the agency status of an individual to how that individual is considered, sometimes using the very same language that he employs in the mind-body case. For example, Spinoza concludes that 'the passions are not related to the mind except insofar as [*quatenus*] it has something which involves a negation, or insofar as it is considered as [*sive quatenus consideratur*] a part of Nature which cannot be perceived clearly and distinctly through itself without others' (IIIp3s). That is, whether or not a particular change is a passion or an action for a mind depends partly on how that mind is considered (here, whether or not it is considered in relation to external causes). He repeats this variability in Part V: 'In this way, all the appetites, or desires, are passions only insofar as [*quatenus*] they arise from inadequate ideas, and are counted as [*accensentur*] virtues when they are aroused or generated by adequate ideas' (Vp4s, G II/283).

Expressions like '*quatenus consideratur*' repeat Spinoza's wording in IIp6 concerning mental and physical causal variability. Spinoza's *quatenus* operator, especially when paired with some form of *conceptus*, is often part of his general, realism-preserving strategy for relativizing variations without undercutting identity (Newlands 2018: 43–44). So perhaps in his action theory, Spinoza is similarly claiming that the very same change could be an action *and* a passion for the same individual, depending on how the individual and the change is considered or conceived.

This parallel is strengthened in IVp59d:

> Any action is called [*dicitur*] evil insofar as it arises from the fact that we have been affected with hate or with some evil affect (see IVp45c1). But no action, considered in itself, is good or evil…instead, *one and the same action is now* good, *now* evil. Therefore, to *the same action which is now* evil, or *which arises from* some evil affect, *we can be* (by IVp19) led by reason (G II/255, emphasis mine).

Spinoza's 'now . . . now' wording echoes his claims from IIp7s about the identity of substance and the identities of its modes across causal and attribute differences. In that vein, Spinoza seems to be claiming in IVp59d that the source of one and the same change can be considered in different ways, a conceptual shift that corresponds to genuine variation: the same change can be a passion or an action of the same individual, depending on how the causal history of the change is considered (here, whether in relation to 'some evil affect' or 'reason').

Spinoza then illustrates this conceptual shift with a vivid example that reinforces the general parallel:

> The act of beating, insofar as it is considered physically, and insofar as we attend only to the fact that the man raises his arm, closes his fist and moves his whole arm forcefully up

> and down, is a virtue, which is conceived from the structure of the human body. Therefore, if a man moved by anger or hate is determined to close his fist or move his arm, that (as we have shown in Part 2) happens because one and the same action can be joined to any images of things whatsoever. And so we can be determined to one and the same action both from those images of things which we conceive confusedly and [from those] we conceive clearly and distinctly. (IVp59s)

Setting aside some of the details, notice how tightly Spinoza ties the status of a change (whether it is an action or a passion) and its cause (whether an agent or a patient) with how the event and cause are conceived or considered in relation to external causes. Conceived in a narrower way, a man's striking someone is an *action* brought about by an *agent*; conceived more broadly in relation to the activities of external things, the same event is merely a *passion* of the same man, who is only a patient or partial cause of it. The extent of an individual's agency for a change tracks the extent to which the source of the change is internal to him, and the extent to which that source is internal to him somehow depends on how narrowly or broadly the individual is conceived. Spinoza even ties the moral valence of the effect—whether or not it is virtuous—with how the effect and its cause is considered or conceived.

Although the parallel between the agency and the mind-body cases is suggestive, it breaks down in at least one important respect. In the mind-body case, Spinoza ties differences in causal structures to differences in attribute contexts. How a thing is conceived with respect to an attribute partly determines the nature of its effects. But in the case of actions and agents, the worrisome variation can arise *within* an attribute. A body can be both the entire and also merely the partial cause of a physical change, and attributes are too coarse-grained to make sense of that.

However, Spinoza also makes a more fine-grained identification *within* an attribute that is supposed to parallel the identity of a mind and body *across* attributes:

> We have shown that the idea of the body and the body, that is (by IIp13), the mind and the body, are one and the same individual, which is conceived now under the attribute of thought, now under the attribute of extension. So the idea of the mind and the mind itself are one and the same thing, which is conceived under one and the same attribute, namely, thought. (IIp21s)

Spinoza's identification of an idea ('the mind itself') with an idea of that idea implies that one and the same representation can have at least somewhat different representational contents. (For other examples in Spinoza of representation identity despite representational content and character variation, see IIp11c, IIp19–20; Della Rocca 1996: 44–67.)

More importantly for what is to come, it seems that distinguishing an individual mind/body from higher-order representations of that individual need not involve positing additional mental modes to the first-order causal sequence of individual minds/bodies. Similarly, making at least some changes in the representational character or content of higher-order ideas—making them more adequate, say—need not involve changing the causal order among first-order modes. Just as there can be causal variation *across* attributes without loss of identity, Spinoza thinks there can be variation among higher-order representations without multiplying or altering the first-order objects of those ideas.

This is not the place to defend Spinoza's striking higher-order representation identity theory (see Melamed 2013: 174–79), but it reminds us that he is committed to surprising identities both across and within attributes. It also prompts us to look for a structurally similar move in the agency case. Abstracting a bit from the mind-body case, we can see that Spinoza

justifies seemingly problematic causal variation by appealing to a feature of an individual substance or mode that (1) can be many-to-one (i.e., is more fine-grained than identity); (2) co-varies with the relevant causal differences; and (3) is intrinsically sensitive to conceptual shifts. The relevant feature in that case is *attribute* contexts. One and the same individual has multiple attributes; differences among attributes corresponds to differences in causal structure (e.g., mental versus physical); and attribute differences supervene on certain differences in how the individual is considered or conceived.

As I have presented it, Spinoza's action theory needs something structurally similar: an even more fine-grained feature of individuals that can be many-to-one, co-varies with the differences in agency status, and is intrinsically sensitive to representational shifts. For ease, I later refer to these as the FINE-GRAINED, CO-VARIATION, and REPRESENTATIONALLY SENSITIVE desiderata, respectively.

Happily, Spinoza has such a feature: *selves*. Seeing this will, in turn, provide elegant and systematic solutions to the three agency problems. For it turns out that Spinoza does not think that individuals can change the past or alter the world's fundamental causal structure just by thinking about it differently. Rather, he thinks that an individual can represent their *selves* differently, and these changes in self-representation correspond to the changes in agency status. The past and the fundamental causal structure of the world do not change; an individual's *self*-identification changes.

Nothing comes cheaply in metaphysics, however. In the mind-body case, Spinoza needed a surprising and controversial multiplicity: each individual has multiple attributes or fundamental natures, and that is a strange view that was widely rejected by his contemporaries. In the agency case, Spinoza's strategy requires another proliferation: individuals have multiple selves or, less

jarringly, can adopt distinct practical identities. Seeing how this is supposed to work will then pave the way for a new understanding of Spinoza's account of personal transformation.

2. Self-Identity and Self-Acceptance

It might seem like a complete non-starter to try to solve any problem in Spinoza's metaphysics by appealing to a category of selves. Spinoza has been accused of neglecting or excluding various elements that seem central to any plausible account of human selves, such as finitude, subjectivity, qualia, and first-person perspectives. Spinoza's *Ethics* has sometimes been read as encouraging us to *forget* ourselves, or at least to forget those messier, subjective aspects of ourselves that do not readily conform to a neatly naturalized, strictly *sub specie aeternitatis* account of the world. On this reading, Spinoza is the paradigm of a hyper-objective, 'view from nowhere' approach to philosophy that later philosophers like Søren Kierkegaard (2009) and Thomas Nagel (1979) criticize.

I myself do not accept this as a general reading of Spinoza. I have argued elsewhere that Spinoza's monism has an important place for finitude and partiality (Newlands 2018: 18–41), and others, such as Ursula Renz (2010), Karolina Hübner (2015), and Julia Borcherding (2016), have shown how some of the perspectival, experience-laden features of our lives figure importantly in Spinoza's arguments. Still, one potentially missing category in Spinoza's ontology that has not received much attention is the category of *selves*. (A possible exception is Koistinen 2009, which offers a curious account of selves—or at least the individuation of finite minds—in terms of a quantity of divine force.) Although Spinoza uses various reflexive pronouns throughout the *Ethics*, he never formally defines a category of selves, and many of

Spinoza's claims about 'our' psychology could be rewritten entirely in third-person terms of a finite thing's mental and brain states.

Nevertheless, I think human selves is one of the central categories operating in the back half of the *Ethics*. True, Spinoza does not explicitly treat the self as a basic ontological category like modes, minds, and bodies. But we should not infer neglect or nonexistence from non-fundamentality. As we will see, Spinoza thinks that complex entities like human individuals can construct and relate to selves, and he attaches great ethical significance to how well this task is done.

To get some traction, it will be helpful to briefly step away from Spinoza. I believe that Spinoza's interest in selves can be fruitfully understood as a distant relative of so-called deep-self views that have been discussed in contemporary philosophy and moral psychology by Harry Frankfurt, Gary Watson, Susan Wolf, Angela Smith, and many others. (For a pair of recent discussions, see Sripada 2016; Shoemaker 2015; for a more experiment-driven discussion in moral psychology, see Strohminger, Knobe, and Newman 2017.)

Advocates of the deep-self theory begin by considering the attribution of physical /or mental states to selves or persons. (From here on, I use *selves* and *persons* interchangeably.) What makes a given mental or physical state *mine*? In effect, they are asking about boundaries of a responsible self: which traits, dispositions, attitudes, activities, or, more generally, which *states* are my own in such a way that I can be held responsible for them? Crucially, they argue that being a state of an individual human body or mind is insufficient for grounding responsibility, and that other nearby conditions, like voluntary control, are not necessary.

Frankfurt famously claimed that there are plenty of everyday cases in which we do not ascribe physical or mental states and actions to any person, even though they clearly are states of

a human body or mind. We can ignore complex cases of unwanted addiction or self-alienation and just think about a muscle spasm or a fleeting desire:

> We acknowledge that in this strict sense, there is *no person* to whom [a particular involuntary bodily movement] can be attributed—no person of whom it is 'just as much part of him' as his actions and activities are. Now why may a desire not, in a similar way, be an event in the history of a person's mind without being that person's desire? Why may not certain mental movements, like certain movements of human bodies, in this sense belong to no one [i.e., no person]? (Frankfurt 1988a: 61)

In this vein, deep-self views distinguish what we might call *human states*—states of a human body/mind—from *personal states*, which are human states that also belong to a person. A person's 'deep self' is constituted by her personal states, which are a proper subset of her human states. Importantly, appeals to selves and personal states are not supposed to incur ontological commitments to additional fundamental entities or basic classes of properties beyond whatever fundamental individuals and mental or physical properties to which one is otherwise committed. But even though selves are derivative constructs, they are still real and morally significant on this view.

What converts a merely human state into a personal state? In general, personal states are human states that have been appropriated in the right sorts of ways. Advocates of deep-self theory disagree about the details, intramural debates that often replicate familiar metaethical debates (e.g., reason-responsive versus conative accounts; historical versus structural accounts of agency; individual versus social constructivism). To keep things manageable, I will stick with Frankfurt as a representative spokesperson for this family of views.

For Frankfurt, an agent accepts and identities with some of her human states and rejects and disavows others. This is how she appropriates some of her human states into personal states for which she can be held responsible and which, collectively, constitute her as a self. As Frankfurt (1988b: 170) describes this process, 'It is these acts of ordering and of rejection—integration and separation—that create a self out of the raw materials of inner life,' adding, 'to this extent, the person, in making a decision by which he identifies with a desire, *constitutes himself*.' When this appropriation process yields a stable and coherent set of personal states, persons attain a kind of integrity that Frankfurt (1988b: 165; 175–76) calls *wholeheartedness*.

Although the beginning of this appropriation process sounds like it involves some kind of spooky self-creation, Frankfurt sometimes characterizes the process in more mundane terms that move us back toward Spinoza. He describes appropriation as an internalization of what had been treated as external, a process of taking ownership of merely human states and thereby converting what had been the passions of a mere patient into the actions of an agent:

> A person is active with respect to his own desires when he identifies himself with them, and he is active with respect to what he does when what he does is the outcome of his identification of himself with the desire that moves him in doing it. Without such identification the person is a passive bystander to his desires and to what he does. (Frankfurt 1988c: 54)

For Frankfurt, becoming a more active person is not fundamentally about doing more or different things. Rather, by identifying more with the source of a change, an individual can integrate that source more into who she is as a person, and thereby become more active and responsible for the change.

Frankfurt has simple cases in mind, such as when an individual brings about an effect on the basis of a desire. By becoming more active through appropriating the desire, a person does not change the desire itself nor its effect. What changes is that the desire, which had been merely a human mental state, becomes a personal state as well. This converts the desire's effect into an effect of the appropriating person and makes it something for which she is responsible as an active agent. What had been a passive desire becomes a source of the agent's own action, and the desire's effect becomes something she herself does instead of something done merely through or within her mind.

Here is Spinoza's version of a broadly similar idea about the connection between the internality and agency status of a desire:

> The desires which follow from our nature in such a way that they can be understood through it alone are those which are related to the mind insofar as the mind is conceived to consist of adequate ideas. The remaining desires are not related to the mind except insofar as it conceives things inadequately, and their force and growth must be defined not by human power, but by the power of things which are outside us. The former, therefore, are rightly called actions, while the latter are rightly called passions. (IVApp2)

Put in more Frankfurtian language: desires which I do not appropriate (desires which are 'related to the mind insofar as it is conceived to consist' of mental states that do not include them) are *external* influences, and their effects 'must be defined . . . by the power of things which are outside us'. Those desires might well cause me to bring about various effects, but I will be only an inadequate or partial cause of those effects, a mere patient with respect to them, a Frankfurtian bystander. By contrast, insofar as a desire has been appropriated ('related to the mind') and internalized ('follows from [my] nature'), the changes that the desire brings about will be *my*

effects, the result of my *actions* for which I am the causally responsible *agent*. For Spinoza, the difference between the two groups of desires is a difference of internalization, which, as I argue below, corresponds to whether a desire has been properly appropriated as part of one's self.

Frankfurt's account of appropriation is not exactly something Spinoza would have endorsed. One obvious contrast is that Frankfurt's account (1988b:170–71) includes a volitional component that Spinoza would not accept, since Spinoza rejects any account of agency that appeals to 'a consent, or a deliberation of the mind, or a free decision (for we have demonstrated that this is a fiction in IIp48)' (IIIAppVI, G II/192). Instead, Spinoza treats identification and appropriation as a matter of *representational* ownership. Individuals appropriate human states by representing them as their own, a process of higher-order internalization that fixes the boundaries of one's self and, by extension, delineates the effects for which one is causally responsible as an agent. In Spinoza's strict ontological categories, this happens when an individual changes their reflexive, higher-order representation (IIIp53d), a change in what Spinoza classifies as an 'idea of the mind' (IIp21s).

Spinoza's account of this representational appropriation process centers on what he calls 'acquiescentia in se ipso'. This is a very rich notion in Spinoza's thought. It would not be an exaggeration to say that encouraging a particular form of *acquiescentia in se ipso* is one of the main goals of the entire *Ethics*. But I will pass over many of the details (see Rutherford 1999; Carlisle 2017), as I am most interested in the *role* it plays in his account of selves, agents, and actions.

Acquiescentia in se ipso is also a difficult expression to translate, as most English candidates, such as *self-esteem*, *self-satisfaction*, or *self-contentment*, have misleading pop-psychology baggage attached to them. Most fundamentally, it names an attitude that an

individual bears toward himself—hence the explicitly self-reflexive *in se ipso*—as he represents his causal powers. Spinoza highlights the self-regarding nature of *acquiescentia in se ipso* when he defines it as 'the joy that arises from a person's thinking about himself and his own power of acting' (IIIDefAff25; Silverthorne and Kisner translation). That is, *acquiescentia in se ipso* is a stance that an individual takes by representing and thereby affirming (IIp48) various capacities, abilities, and changes *as his own*, a kind of identification and acceptance of power *as his*. It is an active process of self-identification or, as I translate it from here on, a form of self-*acceptance*.

Spinoza first introduces self-acceptance as 'joy accompanied by the idea of an internal cause' (IIIp30s), which emphasizes both its affective and representational dimensions. As a form of joy, it is a felt quality that accompanies an individual's increase in power (IIIDefAff2) when she represents that increase as a change for which she is causally responsible as an agent. It is a kind of self-celebration: just look at what I myself have done!

But self-acceptance is more than just a self-*regarding* attitude for Spinoza, as if the boundaries of the self and its activities were already fixed, and self-acceptance just names a way of relating to those invariant facts. Self-acceptance can also involve a kind of appropriation. Partly by representing oneself as the cause responsible for this increase in power, an agent can thereby take ownership of the source of that increase and incorporate it into her self-identity (IIIGenDefAff, G II/204). Spinoza says that self-acceptance concerns 'what [a man] affirms of himself [*quod de se affirmat*]' (IIIp55s), which I take to be a veiled reference to a process of internal appropriation and integration. This how a desire that I had regarded as an external, determining influence can become an internalized part of my motivational structure, how a passion can become an action, how a desire comes to be 'counted as [a] virtue' (Vp4s, G II/283).

I identify and accept an increase in power as *mine* partly by representing its cause as a part of myself ('the idea of an internal cause'), an intellectualist version of deep-self appropriation.

An individual cannot legitimately appropriate just any ole' state of the world as part of her self-representation. To borrow a term from Marya Schechtman (1996), there is a *reality constraint* on this process for Spinoza. As we saw in the previous section, the domain of passions and actions, patients and agents, includes only an individual's sole or cooperative effects. An individual can internalize and appropriate only those changes for which that individual is at least a partial cause. This bedrock, fundamental causal structure of minds and bodies blocks some wilder attempts at self-constitution. For example, during Wrestlemania IV, 'Macho Man' Randy Savage landed an epic elbow drop on 'Million Dollar Man' Ted DiBiase. By Spinoza's lights, I could not straightforwardly represent Savage's movements as internal parts of me in such a way that I could appropriate and celebrate the defeat of DiBiase as *mine*.

Still, Spinoza's reality constraint is not as restrictive as one might expect, at least insofar as individual minds and bodies are parts of long and entangled causal chains in the way Spinoza thinks they are (Ip28). Upon watching Savage's feat on television as a child, I immediately tried to land a similar elbow drop on my cousin (still sorry, Eric!). Suppose that Savage's bodily movements were part of the complicated causal chain that led to my imitation effort. While I certainly could represent Savage's movements as merely external influences on me ('he made me do it, Mom!'), Spinoza's account leaves open the possibility that I might incorporate those televised movements and burgeoning desire to imitate Savage into my own self-identity in such a way that elbow-dropping my cousin was an event for which I was comparatively more active and more causally responsible.

In later parts of the *Ethics*, Spinoza introduces other, more normative-sounding criteria for self-acceptance, while still affirming the core definition of self-acceptance from part III (IVp52d). For example, he contrasts an imagination-based version of self-acceptance with a reason-based version. The joy of the former 'is more and more encouraged the more a man imagines himself to be praised by others' (IIIp52c), a dependence that Spinoza later criticizes as non-salutary. By contrast, the kind of self-acceptance that 'arise[s] from reason is the greatest [kind] there can be' (IVp52).

There are two main differences between these better and worse forms of self-appropriation, neither of which bottoms out in being issued from distinct mental faculties. The first axis of evaluation concerns the character of the representation itself. Spinoza explains that what makes a person's rational self-acceptance so powerful—and hence virtuous and preferential—is that it includes only 'clear and distinct' representations of his power (IVp52d). This does not entail that unclear or more inadequate representations are guaranteed to be false, only that they are not guaranteed to be true, in which case the possibility of self-deception remains ever present when one's self-representation is less than fully clear and distinct (IVApp3).

For example, by relying on non-clear and distinct representations in self-acceptance, an individual might inaccurately represent the world's causal structure and try to appropriate sources that are not actually co-causes with her of anything. This would be a kind of self-overreaching. Alternatively, an individual might fail to notice various causal influences on his behavior, and therefore not even try to internalize and appropriate them. This would be a kind of under-selling of oneself. In both cases, Spinoza warns that there will be unhealthy consequences, such as a destructive pursuits of false goods (IIIp51s) or a life characterized by self-imposed

limitations and unjustified feelings of helplessness and impotence (IIIDefAff28). Striving for clear and distinct self-representations is a path to a more reliable form of self-constitution, which is why Spinoza commends it as superior.

Another evaluative dimension concerns the *source* of the representational and affective components of self-acceptance. In the inferior version in IIIp52c, an individual relies on the opinions of others to fix the boundaries of her practical identity. For example, what Spinoza describes in IVp58s as a vainglorious individual represents the boundaries of her power based on the opinion of others. Such a person might represent herself as intellectually gifted because her parents repeatedly told her she was smart. *I am who they say I am*, a form of external dependence that makes the resulting practical identities risky and unstable, 'for the multitude is fickle and inconstant' (IVp58s). More generally, Spinoza warns that individuals whose self-identity depends on the presence and activities of transitory things are themselves unstable and ultimately doomed to frustration and destruction (Vp42s). Better forms of self-acceptance will insulate a person's self-identity against such fragility, which is why Spinoza commends them.

One might expect Spinoza's overall recommendation to take the form of *withdrawal*, as if the best way for an individual to go about self-acceptance is to isolate herself from external things. But that is not what Spinoza actually recommends. As we have seen, Spinoza's primary recommendation is for an individual to *integrate* the external more into herself. She thereby converts what had resulted from an external influence—a passion—into something that follows from herself—an action. Less metaphorically, Spinoza recommends an individual to take greater ownership of physical/mental changes to which she contributes by shifting how she represents those changes—that is, by appropriating and integrating more of their causal influences into her

own self-representation. In Frankfurt's normative language, this is how Spinozistic agents *wholeheartedly constitute a self*.

There remain significant differences between Spinoza's account and contemporary deep-self theories. Most obviously, deep-self theorists claim that attributability is part of the basis for moral responsibility, whereas Spinoza arguably rejects the notions of praise and blame altogether (IApp, G II/78; IVp37s2) in favor of an evaluative framework based solely on power, desire-satisfaction and causal responsibility. Still, if I am right here, Spinoza and deep-self advocates would agree that whether a particular desire 'is counted as' a virtue (Vp4) depends partly on how well the desire has been incorporated into a person's self-identity, even though Spinoza thinks the resulting kind of agential responsibility is purely causal.

We can see more illuminating points of contrast by comparing Spinoza's account to Frankfurt's version. Frankfurt (1988b:171; 2006:16) thinks that the process of appropriation is ahistorical, in the sense that the causal origins of one's human states are irrelevant for issues of appropriation. (This is part of the foundation for his famed compatibilism.) For Spinoza, becoming more active through self-acceptance involves internalizing what had been represented as an external 'other', including past causal influences. That is, a successful Spinozistic agent will examine and embrace past influences as essential parts of who she now is as a person and what she does as an agent. Unlike Frankfurt, Spinoza makes coming to terms with one's origins, development, and influences an essential component of successful self-acceptance.

Like many deep-self theorists, Frankfurt (1988b:170–75) is drawn to cases in which individuals have conflicting mental states, such as mutually exclusive desires. On his account, persons decisively settle such inner conflicts by identifying with some of those human states and rejecting others. For Frankfurt, a successful agent will actively prevent some of her human states

from becoming personal states, abdicating responsibility for some of what is in her mind and body. *That's not part of who I really am!*

By contrast, successful Spinozistic self-acceptance is more about absorption than exclusion. Externality and separation is actually the fundamental source of conflict and danger, since Spinoza thinks only external sources can destroy a thing (IIIp4). Insofar as I exclude something from myself and represent it as alien to who I am, I become vulnerable to it in ways that ultimately cut against my self-interests (IVp2–4). Hence, the most stable form of Spinozistic self-acceptance involves saying to more and more causal influences: *Yes, that's a part of who I am too!* It is a process of expanding what an individual had previously represented as the more limited boundaries of herself, thereby converting what she had considered an externally imposed limitation into something more like a self-owned source of strength.

This leads to the wildest point of speculation—and I emphasize that this is highly speculative. Like other deep-self theorists, Frankfurt implicitly limits which human states a person can appropriate. The boundaries of the self can only narrow from an individual's human states. That is a very natural restriction, especially if the bearers of human states are something like finite substances, as it is hard to see how one substance could directly appropriate another substance's states.

But suppose, with Spinoza, that there are no finite substances. Suppose individual human minds and bodies are nothing but organized collections of mental and physical states of the one sole substance. In that case, we might wonder whether Frankfurt's restriction holds. Might an individual incorporate into her self-identity an influential desire that she had previously represented as being a part of someone else's mind? Might she successfully represent herself as expansive enough to include among her personal states a physical state that she had previously

represented as belonging to an external body? Nothing in Spinoza's account rules out the possibility of such expansive forms of appropriation, so long as his reality constraint is satisfied (i.e., there is actual mental or physical causal cooperation).

In a different context, Renz (2011:112–13) raises something close to this speculative possibility for Spinoza, and she rightly distinguishes it from an even wilder alternative that I am not suggesting either:

> It becomes fundamentally conceivable that several variations are possible, not in answer to the question 'with which body do we identify ourselves?' (no one in his right mind would identify himself with the body of his postman), but as regards the perception of the individual boundaries between our body and others. The concept of one's *own* body is thus something that can vary. (emphasis mine)

Regardless of how expansive these possible variations turn out to be for Spinoza, we have now discovered the feature of his account of selves that can address the potential problems facing his account of agency and actions. As we will see in the final section, the differences among the states and changes that an individual can represent as her *own*—that is, variations in an individual's higher-order self-identity—is precisely the many-to-one, agency-tracking, and representationally sensitive feature of individuals that we were looking for at the end of the first section.

3. Personal Transformation Reconceived

Let us now weave these two threads together—Spinozistic agency and Spinozistic selves. In section one, I presented Spinoza's puzzling suggestion that the same individual can be an agent *and* a patient with respect to the same change, making the change an action *and* a passion for the

same individual, depending on how the individual and change is represented. We now have a better sense of what this does and does not involve.

Spinoza's positive thesis is that an individual can adopt different higher-order, self-constituting representations of first-order mental and physical states, and these different representations fix the boundaries of an individual's self and domain of agency differently. This possibility is implicit in his urging readers to pursue better forms of self-acceptance (e.g., IVp52 and IVp58s, discussed above, and more generally in IIIp11s). It also satisfies FINE-GRAINED and CO-VARIATION. Furthermore, converting a passion into an action by shifting how a change is represented involves incorporating what had been represented as an external influence into one's self-identity, an identity that, according to Spinoza's intellectualist account of self-acceptance, is intrinsically tied to these higher-order representational facts. That satisfies REPRESENTATIONALLY SENSITIVE.

Seeing what Spinoza is *not* saying allows us to defang the three agency problems. In making such a representational shift, neither the past nor the fundamental causal structure among first-order physical and mental modes changes. No matter how anyone now represents anything, a physical desire still caused this hand to bring this bottle to these lips an hour ago. There is nothing to be done now about that causal fact. What can still change, if Spinoza is right, is the way an individual incorporates that past desire into her present self-identity. What can change is the extent to which what had been represented as an externally influential desire becomes representationally internalized. What can change is the boundary between what is external and internal to one's *self*. What had been an event happening merely *through* a person can become something she herself did, a passion can become an action.

But given what I presented in section one as Spinoza's identification of higher-order representations with their lower-order counterparts (IIp21s), these higher-order representational changes need not induce changes in the causal sequence of first-order mental modes or minds in which they occur. In a similar vein, I noted above in section one that there is a sense in which becoming more powerful and more active for Spinoza is not really about doing (in Spinoza's sense) anything new or different. The moral thrust of the *Ethics* is not for readers to try to become causes of more or new effects. It is for an individual to try to become a more adequate cause of *the very same effects* of which she had been an inadequate or merely partial cause:

> Finally, insofar as joy is good . . . it is not a passion except insofar as a man's power of acting is not increased to the point where he conceives himself and his actions adequately. So if a man affected with joy were led to such a great perfection that he conceived himself and his actions adequately, he would be capable—indeed, more capable—of the same actions to which he is now determined by affects which are passions. (IVp59d)

According to Spinoza's account of self-acceptance, this transformation gradually occurs as an individual's effects flow more from what she increasingly identifies and represents as her own personal desires and power. In this way, moral improvement stems not from an individual performing a better or even a different act, but from the same event coming to follow to more fully from an agent's own motivational structure, where 'own' is fixed by her higher-order representational identification and appropriation.

That sounds rather abstract, so I conclude with an example from Spinoza's closing comparison of the wise and the ignorant:

> For not only is the ignorant man troubled in many ways by external causes and unable to ever possess true peace of mind [*vera animi acquienscentia*], he also lives as if he neither knew himself, nor God, nor things; and as soon as he ceases to be acted on, he ceases to be. On the other hand, the wise man, insofar as he is considered as such [*quatenus ut talis consideratur*], is hardly troubled in spirit, but being, by a certain eternal necessity, conscious of himself and of God and of things, he never ceases to be but possesses true peace of mind [*vera animi acquienscentia*]. (Vp42, G II/308)

Although Spinoza contrasts the wise and the ignorant here, they are not necessarily distinct *individuals*. Notice how Spinoza appeals to a wise man *insofar as he is considered as such*, a rather momentous phrase in light of his representationalist accounts of selves and agency. The wise and the ignorant could just be different *selves*, different ways that one and the same individual can represent themselves in relation to minds, bodies, God, and the fixed, fundamental causal order.

The ignorant self is anxious about what he represents as external, threatening causes. He never achieves the highest form of self-acceptance. Spinoza describes such a person as being 'at the mercy of his affects [and] not under the control of himself, but of fortune' (IVPref, G II/205). By contrast, the wiser self internalizes more of what had been represented as external influences, and thereby becomes more active, more powerful and less vulnerable to destruction. The wise person represents more of the sources of his activities as essential parts of himself and integrates more of his motivational structure into his self-identity, thereby transforming what had been passions into actions. In this way, an individual increasingly becomes a better person, a new self. So understood, Spinoza's closing question to his individual readers is this: which person will you become, which self—the wise or the ignorant—will you appropriate?

Samuel Newlands

University of Notre Dame

snewlands@nd.edu

References

Bennett, Jonathan. (1984) *A Study of Spinoza's 'Ethics'*. Indianapolis: Hackett.

Borcherding, Julia. (2016) 'A View from Nowhere? The Place of Subjectivity in Spinoza's Rationalism'. In Jari Kaukua and Tomas Ekenberg (eds.), *Subjectivity and Selfhood in Medieval and Early Modern Philosophy* (Dordrecht: Springer), 215–62.

Carlisle, Claire. (2017) 'Spinoza's *Acquiscentia*', *Journal of the History of Philosophy*, 55, 209–36.

Della Rocca, Michael. (1996) *Representation and the Mind-Body Problem in Spinoza*. New York: Oxford University Press.

Descartes, René. (1985) *The Philosophical Writings of Descartes*. Vol. 1. Translated by John Cottingham, Robert Stoothoff, and Dugald Murdoch. Cambridge: Cambridge University Press.

Frankfurt, Harry G. (1988a) 'Identification and Externality'. In *The Importance of What We Care About: Philosophical Essays* (Cambridge: Cambridge University Press), 58–68.

Frankfurt, Harry G. (1988b) 'Identification and Wholeheartedness'. In *The Importance of What We Care About: Philosophical Essays* (Cambridge: Cambridge University Press), 159–76.

Frankfurt, Harry G. (1988c) 'Three Concepts'. In *The Importance of What We Care About: Philosophical Essays* (Cambridge: Cambridge University Press), 47–57.

Frankfurt, Harry G. (2006) ‘Taking Ourselves Seriously’. In Debra Satz (ed.), *Taking Ourselves Seriously and Getting It Right*. (Stanford: Stanford University Press), 1–26.

Garrett, Don. (2008) ‘Representation and Consciousness in Spinoza’s Naturalistic Theory of the Imagination’. In Charles Huenemann (ed.), *Interpreting Spinoza: Critical Essays* (Cambridge: Cambridge University Press), 4–25.

Hübner, Karolina. (2015) ‘Spinoza on Negation, Mind-Dependence, and the Reality of the Finite’. In Yitzhak Y. Melamed (ed.), *The Young Spinoza: A Metaphysician in the Making* (Oxford: Oxford University Press), 221–37.

Kierkegaard, Søren. (2009) *Concluding Unscientific Postscript to the Philosophical Crumbs*. Translated by Alastair Hannay. Cambridge: Cambridge University Press.

Koistinen, Olli. (2009) ‘Spinoza’s Eternal Self’. In Jon Miller (ed.), *Topics in Early Modern Philosophy of Mind* (Dordrecht: Springer), 151–69.

Löwe, Can Laurens. (2018) ‘John Duns Scotus versus Thomas Aquinas on Action-Passion Identity’. *British Journal for the History of Philosophy*, 26, 1027–44.

Marshall, Colin. (2012) ‘Spinoza on Destroying Passions with Reason’. *Philosophy and Phenomenological Research*, 85, 139–60.

Melamed, Yitzhak Y. (2013) *Spinoza’s Metaphysics: Substance and Thought*. Oxford: Oxford University Press.

Nagel, Thomas. (1979) ‘Subjective and Objective’. In *Mortal Questions* (Cambridge: Cambridge University Press), 196–213.

Newlands, Samuel. (2018) *Reconceiving Spinoza*. Oxford: Oxford University Press.

Renz, Ursula. (2011) ‘The Definition of the Human Mind and the Numerical Difference between Subjects (2p11–2p13s)’. In Michael Hampe, Ursula Renz, and Robert Schnepf (eds.), *Spinoza’s Ethics: A Collective Commentary* (Leiden: Brill), 99–118.

Renz, Ursula. (2010) *Die Erklärbarkeit von Erfahrung: Realismus und Subjektivität in Spinozas Theorie des menschlichen Geistes*. Frankfurt: Klostermann.

Rutherford, Donald. (1999) ‘Salvation as a State of Mind: The Place of *Acquiescentia* in Spinoza’s Ethics’. *British Journal for the History of Philosophy*, 7, 447–473.

Schechtman, Marya. (1996) *The Constitution of Selves*. Ithaca: Cornell University Press.

Schickel, Joel A. (2011) ‘Descartes on the Identity of Passion and Action’. *British Journal for the History of Philosophy*, 19, 1067–84.

Shoemaker, David. (2015) ‘Ecumenical Attributability’. In Randolph Clarke, Michael McKenna, and Angela M. Smith, (eds.), *The Nature of Moral Responsibility: New Essays* (Oxford: Oxford University Press), 115–40.

Sripada, Chandra. (2016) ‘Self-Expression: A Deep Self Theory of Moral Responsibility’. *Philosophical Studies*, 173, 1203–32.

Spinoza, Baruch. (1985) *The Collected Works of Spinoza*. Vol. 1. Edited and translated by Edwin Curley. Princeton: Princeton University Press.

Spinoza, Baruch. (2002) *The Complete Works*. Translated by Samuel Shirley. Edited by Michael L. Morgan. New York: Hackett.

Spinoza, Baruch. (2018) *Ethics: Proved in Geometrical Order*. Edited by Matthew J. Kisner. Translated by Michael Silverthorne and Matthew J. Kisner. Cambridge: Cambridge University Press.

Spinoza, Baruch. (1925) *Opera*. 4 vols. Edited by Carl Gebhardt. Heidelberg: Carl Winter.

Strohminger, Nina, Joshua Knobe, and George Newman. (2017) 'The True Self: A Psychological Concept Distinct from the Self'. *Perspectives on Psychological Science*, 12, 551–60.