

Two Sorts of Self-Creation: On Galen Strawson’s “Basic Argument”

Griffin Klemick

In a handful of works over the past few decades, Galen Strawson has formulated and defended an argument he calls the “Basic Argument” against moral responsibility.¹ This argument leverages the seemingly demonstrable impossibility of control over one’s character against any account of responsibility robust enough to ground our moral practices. Strawson has given the Argument in myriad formulations, but one has become particularly prominent in the literature. I give this version below; in my view, it omits nothing central to Strawson’s argument.²

The Basic Argument

- (1) In any given situation, you do what you do because of the way you are.
- (2) So in order to be ultimately responsible for what you do, you must be ultimately responsible at least for particular mental aspects of the way you are.³
- (3) But you cannot be ultimately responsible for the way you are.

¹ I will always give *Basic Argument* in capital letters. Where *Argument* is given with a capital letter, it refers to the Basic Argument as given below.

² The first four premises of the argument are based on Strawson (2002, version 1); the regress component is based on Strawson (1994, sec. I).

³ The qualifier “at least for particular mental aspects of” is added to forestall the idea that one must be responsible for *all aspects* of oneself—for one’s hair color, ancestry, intellectual aptitude, and so on. The mental aspects in question are characterized in premise 6. I will take this qualifier for granted in the remainder of the argument.

***Two Sorts of Self-Creation:
On Galen Strawson's "Basic Argument"***

- (4) Therefore, you cannot be ultimately responsible for what you do.

Why can't you be ultimately responsible for the way that you are? Because:

- (5) To be ultimately responsible for the way you are, you must have consciously and explicitly chosen to be the way that you now are. Call this choice "C."
- (6) But for you to have *chosen* to be the way you now are, you must have existed prior to your being that way, and at that prior time you must already have held to certain principles by which you might choose C—preferences, values, ideals, and the like. Call these principles collectively "P." (In what follows, I will call these principles one's *volitions* and their collection one's *volitional makeup*.)⁴
- (7) But then to be ultimately responsible for C, you must be ultimately responsible for having held P.
- (8) But to be ultimately responsible for having held P, you must have consciously and explicitly chosen to hold P. Call *this* choice "C*."
- (9) But to have chosen C*, prior to C* you must have existed and held to certain principles by which you

⁴ As I will use the term, one's *volitions* include one's first-order desires as well as one's higher-order desires and deep concerns, values, and ideals. Perhaps this use of the term is non-standard, but another lump term was not forthcoming.

LYCEUM

might choose C*. Call *these* principles collectively “P*.”

- (10) This cycle—ultimate responsibility for mental aspects of yourself requiring choosing those aspects, and ultimate responsibility for those choices requiring an appeal to still prior mental aspects—is a backward infinite regress.
- (11) So to be ultimately responsible for the way that you are, you must have actually made an infinite series of choices.
- (12) But it is impossible for a finite being to have actually made an infinite series of choices.
- (13) Therefore, it is impossible to be ultimately responsible for the way that you are.

For some time, I found myself in the position in which Berkeley left Hume: the Argument produced no conviction in me, but admitted of no decisive response. I now think this stemmed from my misconception of what the first step toward a response would have to be. Strawson’s argument ultimately derives its support from an appeal to intuition: the notions of control over self and moral responsibility to which the Argument makes reference are “central to ordinary thought about moral responsibility and justice” (Strawson, 1994, p. 222). For this reason, it seemed to me that compatibilists should argue that our intuitions do not support the requirement of Strawson’s “ultimate responsibility” for moral responsibility, for all the reasons given by the standard arguments for

Two Sorts of Self-Creation: On Galen Strawson's "Basic Argument"

compatibilism. And libertarians should argue that ultimate responsibility is possible,⁵ for all the reasons given by the standard arguments for libertarianism.

I see matters differently now. To be sure, the fact that libertarians and compatibilists can so easily leverage their standard arguments against Strawson's premises does call into question the Argument's dialectical value. But framing the issue in this way allows a potential deeper point of agreement between libertarians and compatibilists to be drowned out by the familiar "dull thud of conflicting intuitions." In this paper, therefore, I wish to consider a more fundamental answer to the Basic Argument, one arrived at through a less direct route. I propose to examine in some detail two accounts of self-creation—one expressed by the Basic Argument; the other, a more modest proposal, detailed by Harry Frankfurt and Charles Taylor—and to consider which sort of self-creation places us in a better position to attain the values we seek through our agency. Does each form of self-creation allow us to identify with our selves, or does it leave us alienated from our volitions and our actions? In answering this question, I will develop a challenge to Strawson's claim about our ordinary conception of responsibility based on premises both libertarians and compatibilists can accept.

First, I will give an account of what it is to be *causa sui* and a preliminary description of the goods the *causa sui* agent can and cannot achieve. Next, I will sketch an alternate account of our ability to shape ourselves, arguing that our particular actions and our attempts to articulate our fundamental values imbue us with significant control over our wills. Finally, I will argue that the latter form of self-shaping better positions us to attain the goods desired in agency, precisely because it situates self-creation within human contexts and communities.

⁵ Or perhaps they should argue that intuition requires only a less demanding type of incompatibilist responsibility, which type we actually possess.

LYCEUM

Strawsonian Self-Creation

For Strawson (1994), if we are to have a meaningful measure of control over our selves, we must create them—we must be *causa sui* with respect to them (see p. 212). Obviously, this formulation is unhelpful by itself: we need a detailed account of what it is for human agents to be self-creators in the relevant sense.⁶ We may look to the regress component of the Basic Argument for a concrete illustration: what is required is that, for every aspect of an agent's volitional makeup, she has consciously, explicitly and reflectively chosen that aspect, and at least one such choice is not motivated by any further aspect.⁷ It seems appropriate to impose further affective conditions: let us say that the agent must care about which kind of will she comes to have, and she must recognize her responsibility for her choice (rather than, e.g., ducking it through Sartrean bad faith). If these conditions are met, we have a self that carefully and deliberately builds itself up from nothing. We might think that the “finished product” is a self for which the agent can rightly feel ultimately responsible.

Note that these conditions actually go beyond the requirements described in the Basic Argument's regress scenario. This is necessary: while the Argument

⁶ Strawson's parody of overly general accounts is apt: to be *causa sui* is “really quite simple . . . in the case of stones, for example, it's just for the stone to be—truly—the origin or creator of itself” (2004, p. 362). Obviously, this doesn't help us understand why we should identify with the self we create.

⁷ A few points on this rough summary. First, concerning the requirement of conscious, explicit reflection, Strawson tends to restrict the scope of the Basic Argument to “fully intentional and consciously deliberated actions” (2002, p. 442), but for expository rather than theoretical reasons. So it might seem that this requirement goes too far. But Strawson views uncaused actions as random occurrences for which no agent can be responsible (see his 2000), without apparent regard for whether the absence of causality occurs at the psychological level or the physical one. Second, since the Basic Argument is challenging responsibility for actions, it requires that an agent be responsible only for those aspects of the agent's will that issue in or contribute to subsequent actions. But since the issue here is with identification with the self, it seems reasonable to extend the requirement to all aspects of the agent's will. (In any event, how can a volition be a part of an agent's self in any deep or important way if it has no effect at all on her behavior?)

Two Sorts of Self-Creation: On Galen Strawson's "Basic Argument"

takes its cues from the problem of being *causa sui* in time, the property of being *causa sui* is purely *structural*; it has no essential *temporal* components, and so the condition that all aspects of our will be chosen is not sufficient for our being *causa sui*.⁸ For an agent who, in an infinite series of choices, chooses every aspect of herself may yet fail to be *causa sui*. This could be true, for instance, if she made each choice in this infinite series of choices with reference to “still prior” decision-making criteria; on this account, the structural question of whether the agent controls her will or is controlled by it is unclear.⁹ On the other hand, it seems that an agent who exists entirely outside temporal boundaries could be *causa sui*, if the right structural relationship obtained between her volitional makeup and her choice—if, in one timeless act, she eternally chooses all her values and desires without reference to further, un-chosen principles of choice. (Strawson seems to recognize this when he softens his claim that “nothing can be ultimately *causa sui* in any respect at all” to the claim (1994, p. 224) that “even if God can be, we can’t be.”)¹⁰ To be *causa sui* in the Strawsonian sense, then, all an agent’s choices must structurally stem ultimately from only the choice of the agent’s “bare self.”

⁸ This does not, of course, mean that Strawson is wrong to draw out the impossible necessary conditions of a *temporally bound human's being causa sui*. It is only to note that these conditions are particular to humans, and that the property is not itself temporal in nature.

⁹ Of course, it is unclear what we should say about this agent’s control over her will, given the introduction of the impossible actual infinite sequence. But we cannot for this reason disallow the example, since the whole point of the Basic Argument is that key human goods depend on our possession of a kind of control over ourselves that is achievable only by meeting impossible necessary conditions. It is only fair, then, to ask whether the impossible necessary conditions would actually help us achieve any key human goods. And this seems in part to hinge on the unanswerable question whether it is more important, in cases like the one above, that *the agent chooses all her reasons* or that *each of her choices is controlled by a prior reason*. I contend that in such examples, there is no reason to assume that the first fact is more important.

¹⁰ And even if God can’t be *causa sui*, I doubt that this is due to God’s transcendence of temporality.

LYCEUM

What sort of self can take its place at the top of this structure, in which all aspects of its will depend ultimately on reason-less choice? Only a bare subject of experience, which Thomas Nagel (1986) describes as a “pure, featureless mental receptacle” (p. 33), fits the bill.¹¹ This self is to the mental aspect of a human being what prime matter is to a human body. Like prime matter, such a self cannot actually exist, and some compatibilists have argued that this should gainsay the claim that it figures crucially in meaningful agency: for instance, Daniel Dennett (1984a) contends that “if the ideal of freedom we hold out for is simply self-contradictory, we should hardly feel bereft when we learn we cannot have it. There’s no sense wringing our hands because we can’t . . . create ourselves *ex nihilo*” (p. 172).

In one sense, this is quite right: it is *useless* to regret our inability to be *causa sui*,¹² since it’s not our fault we lack this ability and we have no prospects of acquiring it. But perhaps this does not mean that we cannot intelligibly wish that we had it or that we cannot recognize the goods we might attain by it. To borrow an analogy from Strawson (2004), it is only because we can state with precision what it is to be a square circle that we know this is impossible (p. 360); likewise, he contends, we know that being volitionally *causa sui* is impossible because we understand roughly what it is to have this ability and, perhaps, what further goods it would confer on us. In view of this understanding, we can intelligibly wish that we could achieve these impossible goods, and we may feel alienated from ourselves because we lack them.^{13,14} At any rate, I will assume

¹¹ Cf. Nagel’s attempt to identify this subject: “I know what I mean by ‘I.’ I mean *this!*” (1986, p. 33).

¹² Cf. R. E. Hobart (1934, p. 18).

¹³ For a compelling presentation of these points, see Nomy Arpaly (2006, pp. 126-33).

¹⁴ One might object that, while we can understand the goods that our possession of nomologically or otherwise contingently impossible abilities would confer on us, we cannot state with certainty the goods that logically impossible abilities would confer on us. The claim *if I had the ability to be in two places at once, I could go to more parties than I can* seems straightforwardly true, even though its antecedent is impossible. By contrast, it is totally unclear how we should determine the truth or falsity of the claim *if my basketball were shaped like a three-dimensional*

Two Sorts of Self-Creation: On Galen Strawson's "Basic Argument"

that we can determine what goods the ability to be *causa sui* might confer on us and will concern myself solely with the question of what goods these are.

It seems doubtful that the ability to be *causa sui* could confer on us any goods at all. Since the bare subject lacks any volitions by which it could form preferences, it is not clear how it could choose anything at all (or how it could *understand itself* to have this power of bare choice, which might be necessary for it to have a sufficient measure of control over its exercise of this power).¹⁵ Further, even if the bare subject could make choices, there seems no reason to think that it would have the tendency to make *good* choices, for it lacks criteria for distinguishing the good from that which is not good. So the genuinely *causa sui* self cannot dispel our worries about our ability to be moral agents or to stave off alienation.¹⁶

circular square, I could make more impressive bank shots. And the claim if I were causa sui, I could better identify with my will seems more like the latter than the former.

I think this objection is correct, but I would prefer not to make it bear the weight of my whole argument, particularly when the argument in the text potentially yields even more conclusive results by granting Strawson even more and yet confuting his argument.

¹⁵ Hobart (1934) points out that even if an agent could be *causa sui*, we would nevertheless have difficulty *understanding him* to have this ability: "If there were a being who made his 'original character,' and made a fine one, and we proceeded to praise him for it, our language would turn out to be a warm ascription to him of a still earlier character, so that the other would not have been original at all" (p. 18). I am suggesting that it is possible that not only the observers but also the agent would have this difficulty.

¹⁶ One might respond that there remains a chance that the bare subject will choose the good, whereas some human beings are deprived of even a chance by the influence of heredity and environment. But there are two reasons to reject this response. First, unless human beings are fully incapable of apprehension of moral reasons, they have some resources by which they may choose the good, which the bare subject totally lacks; this seems at least as relevant to assessing grounds for alienation as the chance to which the response appeals. Second, it is not clear that any choice for good the bare subject makes could constitute moral agency, since morality makes claims not only on an agent's choices but on her reasons for choosing. Since a bare choice just is a choice not made for reasons, it is not, *a fortiori*, made for moral reasons either. Charles Taylor's critique of Sartrean radical choice seems to apply just as well to Strawsonian self-creation: by demanding that we choose not only our actions but also the values that motivate them, it "dress[es] up as a moral choice

LYCEUM

Lest the discussion grind to a halt so quickly, however, I think a softened, yet still robust form of “*causa sui* agency” might be plausibly thought to provide us with the goods desired in moral agency. While no self that must choose all its reasons will do, there might be some reasons that a self may take as given without thereby diminishing its control. Perhaps some reasons are binding for all rational agents as such; perhaps, along Kantian lines, we might stipulate that these include moral reasons. Even if this is not a tenable metaethical position, it seems to be an assumption underlying our wish to be *causa sui*, at least as we typically express it in everyday life. For what we ordinarily wish is that we could strip away those particular volitions that hold us back from acting in line with normative reasons, including moral ones, and that we could replace these volitions with dispositions to act according to these reasons. We take for granted that moral reasons will continue to be reasons for us even when we have stripped away many of our actual volitions, and this seems plausible only if moral reasons are reasons binding on rational agents as such.

Let us stipulate, then, that the “bare self” has access to reasons binding for rational agents as such, including moral reasons. This self seems capable of creating a will from which it would not be alienated, since it has an ability to choose that is unencumbered by a particular context and has the required knowledge to comport itself with moral norms. Full self-creation may not be desirable, then, but this slightly restricted self-creation still might be. Two questions remain, then: whether a still more modest form of self-creation is possible for human beings, and whether this still more modest contender is equally desirable—or more desirable—than the slightly restricted version. I argue that both questions should be answered in the affirmative; I treat the former question in the next two sections and the latter in the penultimate section.

what is really a *de facto* preference” (1976, p. 121). If this is right, then totally unrestricted choice deprives one of the capacity for moral agency far more than any typical constraints on agency faced by human agents.

***Two Sorts of Self-Creation:
On Galen Strawson's "Basic Argument"***

The Relationship Between Actions and Volitions

In this section and the next one, I will focus on the simple question of how we might have significant causal influence and control over our volitions if we cannot be *causa sui*. One plausible answer is that we can attempt to articulate our deepest volitions, identify with some of them rather than others, and take steps to comport our actions with the chosen volitions. In the next section I will describe this process in more detail.

But first I want to give this simple question a simpler answer, and so I invite the reader to perform a simple thought-experiment. Simply think of a few of your significant volitions: your interests, concerns, loves and values. Now consider how many of these volitions became so important to you without your actions in any way contributing to the process. Speaking for myself, I find the answer is *none of them*. Of course, I did not typically make the choices that deepened the importance of these volitions to me with full knowledge of how my choices would affect me. Nor, in making any such choice, was I free from the influence of other desires or concerns—still less from environmental influences. Nevertheless, in each case I willingly (and more-or-less reflectively) chose to act in ways that caused my will to take its current shape, and I am convinced this is significant.

Perhaps this answer is not only simple but also shallow. After all, one convinced by the Basic Argument might make the familiar retort: Even if these choices affected my volitional makeup, weren't they in turn caused by other volitions, and doesn't this chain infinitely regress? Perhaps, but I argue that it is equally true that the volitions that influenced my choices did not take shape until I acted in ways that furthered or gave rise to them—the chain of dependence of volitions on actions also infinitely regresses. It may be true that simplistic forms of libertarianism may disproportionately emphasize our ability to choose, considered in abstraction from the influence of volition. But it is equally true that simplistic forms of compatibilism or hard determinism may disproportionately

LYCEUM

emphasize the dependence of actions on one's volitions. Of course, each of my actions stems from some subjective motivation. But there is no reason to think that every such subjective motivation will answer to labels like *my deepest volition* or *my strongest desire*. After all, as Carl Ginet (1989) points out in response to Mill's contention that one's "chosen action will [always] be the one that satisfies whichever of the conflicting motives is stronger than all the others":

One can secure confidence in Mill's law only by making it true by definition: "the strongest motive" *means* the motive that prevails. If the term is defined by some logically independent criterion, so that the proposed law will be a non-trivial proposition, then it is an open question whether the facts would give us reason for confidence in it. (pp. 26-7, Ginet's italics)¹⁷

If one operates with a definition of "strongest desire" such that all subjective motivations that lead to action satisfy the term, this simply evinces the triviality of the established claim.

I suggest, then, that a different model of the relationship between actions, volitions and the self is required: a model that keeps to the middle ground between these two disproportionate emphases. We should reject models that identify the self with the subjective power of decision and locate it above volition, such that its decisions can be made apart from—and can always subsequently alter—the influence of its volitions. We can never choose apart from the influence of volition in general; nor should we underestimate the difficulty of acting to diminish the hold of particular strong volitions over our powers of decision. But equally worthy of rejection are those models that constitute the self from a set of particular, static, mutually consistent volitions and traits, identifying free action with the mechanistic function of these central

¹⁷ The wording of the contention termed "Mill's law" is Ginet's own.

Two Sorts of Self-Creation: On Galen Strawson's "Basic Argument"

volitions.¹⁸ This type of model unduly minimizes the importance of an agent's actions in shaping her volitional makeup; it also underestimates the degree to which the volitions central to her identity may be mutually inconsistent. Interestingly, Galen Strawson blends these two models: for Strawson, the self that is *causa sui* can be nothing but the bare subject, but its actions are a function of its volitional makeup. So, the self must create itself "from the top down": to be responsible for its actions, it must choose (without motives) all the intermediary volitions that will determine them.

I suggest instead that an agent's actions are motivated by given oft-mutually inconsistent volitions, each of which in turn increases or decreases in importance to her based on her actions and experiences. Crucially, neither the will's influence on action nor action's influence on the will is more fundamental than the other. The agent's self does not stand behind her volitions; rather, the self is constituted by the interplay between volitions and actions (as these reflect the influences of her experiences and the communities to which she belongs).

In practicing self-creation through our actions and particular choices, we do not directly create our volitions; they are, rather, self-creation's raw material and indirect consequences. But even if volitions exist in us unbidden, our actions make them a part of us in a more significant way. Harry Frankfurt (1987) grasps this point acutely:

When the decision is made without reservation, the commitment it entails is decisive. Then the person no longer holds himself apart from the desire to which he has committed himself. It is no longer unsettled or uncertain whether the object of that desire—that is, what he wants—is what he really wants: the decision determines what the

¹⁸ Hume seems to have adopted such a model of free agency when he argued that free actions are just those that express the agent's deep or settled character traits. Cf. John Martin Fischer (2005, p. 157).

LYCEUM

person really wants by making the desire on which he decides fully his own. To this extent the person, in making a decision by which he identifies with a desire, *constitutes himself*. (p. 170, Frankfurt's italics)

For Frankfurt (ibid.), it is not the *existence* of the person's desires that is important for moral agency; it is whether the person *takes responsibility* for these desires, and thus makes them a part of his self (see pp. 171-2). And he can do this through his particular actions just as much as by choosing his volitions from the top down.

John Martin Fischer (1999) finds the middle ground between the two disproportionate emphases I've described by contending that the value of free agency consists in the agent's expression of "the meaning of the sentence of the book of his life. And this meaning is fixed in part . . . by the overall narrative structure of the life" (p. 116). In morally responsible agency, we further the story of the self; we express further elements of the interplay between our actions and our choices. In some cases, the new chapter will show how our values, already relatively stable and fixed, worked themselves out in action. In other cases, it will show how, in a moment of uncertainty, our actions resolved a battle between our desires and subsequently altered our volitional makeups.¹⁹ In still other cases, including cases of *akrasia*, it will describe how our actions were out of sync with our central volitions. In many situations of each kind, we act responsibly and further the stories of our selves. We should resist the urge to restrict attributions of responsibility or meaningful self-creation to any one type of narrative.

¹⁹ Cf. Frankfurt's (1987) claim that deliberate decisions serve as guides for other decisions and thus tend to integrate us over time as people who care about the volitions that motivated the decisions (p. 175).

Two Sorts of Self-Creation: On Galen Strawson's "Basic Argument"

Articulating Our Values

One way in which we can shape our wills, then, is through our decisions to act in particular ways. But we can also shape our wills more indirectly, and by laying out a model of the acting self in the previous section, I have given the context for a detailed account of this capacity. For I have suggested that our selves are constituted in part by a number of our values, concerns, and projects, not all of which mutually cohere. Further, our volitional makeups are not static; they alter as a consequence of our actions and our experiences over time. Charles Taylor (1976) supplies yet another complexity: when we attempt to articulate our most fundamental values—and so to unify many of the inconsistent volitions that are a part of our selves—it is precisely these fundamental values that are “least clear, least articulated, most easily subject to illusion and distortion” (p. 124). This means that any product of our attempts at articulation will not only be incomplete but may be quite wrong, and so the question of whether we understand ourselves rightly can always arise anew.

This inchoateness of our fundamental values makes possible indirect control over our wills. The formation of our wills is, as Frankfurt (1982) writes, a matter of coming to care about particular things rather than others, and we may indirectly affect this process by *caring about what we care about* (p. 91). This involves reflecting on the worthiness of the things we care about and, since our criteria for worthiness will necessarily be shaped by just these same cares, it also involves remaining open to different criteria than those we currently hold (cf. Taylor, 1976, pp. 124-5). In reflecting on our fundamental values, we neither choose them *ex nihilo* nor simply express the immutable, as Taylor (*ibid.*) notes:

Our attempts to formulate what we hold important must, like descriptions, strive to be faithful to something. But what they strive to be faithful to is not an independent object with a fixed degree and manner of evidence, but

LYCEUM

rather a largely inarticulated sense of what is of decisive importance. An articulation of this “object” tends to make it something different from what it was before. And by the same token a new articulation doesn’t leave its “object” evident or obscure to us in the same manner or degree as before. In the act of shaping it, it makes it accessible and/or inaccessible in new ways. (pp. 123-4)

In articulating our fundamental values, we do shape ourselves in genuinely new ways, but always attempting to be faithful to our own desires and loves, the values of those communities with which we identify, and our best understandings of what is good.

This self-shaping, obviously, does not involve the ability to make “bare choices.” In fact, the decisive importance we place on our fundamental values suggests that articulations of these values often constrain us from acting against them (though, of course, not in every case; *akrasia* often remains possible). This might seem like a limitation of our control over our selves rather than a manifestation of it. But, as Gary Watson (2002) reminds us, we should remember that “the assertion of moral or volitional necessity is not just the assertion of a constraint on one’s deliberative judgments . . . but part of the content of those judgments” (pp. 108-9)—that is, it expresses the vital importance of the judgment in question to us. We often do not consider this constraint on deliberation to rob us of responsibility; rather, as Frankfurt points out, we often take responsibility for our deep values, approving of our inability to go against what we take to be fundamentally important.²⁰

²⁰ Moreover, even if not everyone thinks worries about volitional necessity’s control-undermining tendencies can be so easily dispelled, those who take *causa sui* status with respect to one’s volitions to be necessary for responsibility for our actions should not be among those worried; after all, on their account, one’s actions are a function of one’s character traits, and this allows for ultimate responsibility for one’s actions (provided one is *causa sui*).

Two Sorts of Self-Creation: On Galen Strawson's "Basic Argument"

What about cases in which this volitional necessity *does* have the phenomenology of constraint—when we really wish we could seriously consider acting against our deep volitions? The standard example of volitional necessity in the literature is Martin Luther's (possibly apocryphal) closing words to the Diet of Worms: "Here I stand; I can do no other. God help me."²¹ But, as Nomy Arpaly (2006) wryly remarks, "one can feel chosen by God and yet cry, 'God! Can't you for once choose someone else?'" (p. 126). The constraints these volitions impose on our deliberation sometimes prevent us from performing other actions—perhaps more prudent or noble actions—that we simultaneously desire to perform.²²

But this framing exaggerates the extent of the problem by considering volitional necessity only from the perspective of the agent during deliberation. From a broader, narrational perspective, the import of necessity may be quite different. In some cases, we will ultimately come to identify with the constraint and embrace it, ceasing to reserve acceptance of our desires as they are.²³ In other cases, we may endure the constraint of cares and loves we do not endorse for long periods of time before finally prevailing and throwing them off. Taylor (1989) is right to insist that "the issue of [the human] condition can never be exhausted for us by what we *are*, because we are always also changing and *becoming*" (pp. 46-7, italics original). This is as true for our ability to shape ourselves as for any other aspect of our moral existence. And so, as long as the

²¹ I think the example was first used by Dennett (1984b).

²² This answers a question posed by one referee for this journal, namely, whether the sort of self-creation I argue that humans actually have is sufficient to dispel all moments of frustration at our lack of Strawsonian self-creation. This question should be answered in the negative, as the quotation from Arpaly above suggests. Nevertheless, as will be argued below, while it is at times frustrating to lack the capacity for something like Strawsonian self-creation, on balance, it would even be more frustrating to *possess* this capacity.

²³ See Frankfurt (1977, pp. 63-4). It is important to note that this occurs not only in cases where we finally accede to our desires while, in a significant sense, continuing to affirm our previous judgment that they were not part of our ideal selves, but also in cases where we accede to those desires precisely as we realize that our previous evaluations of them were mistaken.

LYCEUM

states of our wills confront us as problems, we can never rule out the possibility that our moments of impotence will find their places in narratives of developmental or redemptive self-shaping.

But some people never prevail against their truly bad volitions. So this modest sort of self-creation does not preclude our failing to achieve the goods we desire as moral agents. The question, then, is whether it achieves more of them than the ability to be *causa sui* (in, recall, our somewhat restricted sense). In the next section, I will argue that it does, and thus that it paradigmatically enables identification with self.

Causa Sui or Contextual Self-Creation?

Let's take a step back for a moment. What aspects of being *causa sui* seem desirable to us? The answer may seem obvious: it is the ability to *cause* our selves—that is, to have complete and direct control over them. But I think the correct answer goes deeper, for the desirability of causal influence in moral agency is derivative: causal influence correlates with intrinsically significant desiderata. What are these desiderata? They include our sense of responsibility for our wills and our reflective endorsement of our volitions. Some measure of causal influence over our wills is necessary for each of these desiderata; their absence can give rise to feelings of alienation from oneself.

We may call these desiderata *personal*, since they involve only facts about the agent's relationship to her will. But there are other important conditions for identifying with our wills.²⁴ Some of these are *interpersonal*: we want to have

²⁴ Below I discuss *interpersonal* and *suprapersonal* desiderata for moral agency. But Robert Bishop suggests that some desiderata for moral agency are also *impersonal*; these describe states of the mind-independent world that are necessary conditions for the possibility of moral agency. For instance, if universal determinism and causal closure obtain, then perhaps our qualitative experience is epiphenomenal and responsible agency is impossible. Even if this is correct, however, I think it is fitting that this category of desiderata receives less coverage than the other three, because the claim that there really are desiderata for moral agency assumes that the impersonal desiderata are met; if

Two Sorts of Self-Creation: On Galen Strawson's "Basic Argument"

values which some community or communities around us will share. And some might be called *suprapersonal*: many agents desire to value those things they *ought* to value, quite apart from considerations of what they or any communities around them actually endorse. Of course, we cannot step outside our present commitments or all the communities to which we belong, seeing bare ethical facts from “the view from nowhere.” But this doesn’t preclude conformity with objective normative truth from being a sensible regulative ideal (and thus a sensible desideratum).²⁵ Now, consider: does the bare self have the resources to enable us to achieve these suprapersonal goods? I think not, for an important and perhaps surprising reason: the suprapersonal goods depend strongly on the interpersonal goods—i.e. on acting intelligibly within the norms and values of communities to which the agent belongs—in three ways.

First, while there are truths about the manners of existence suitable for human beings that are not grounded in the beliefs of human beings (either individually or collectively),²⁶ this does not entail that either questions of the good life or their concrete answers can be formulated intelligibly without

they are not, we do not in fact have any desires at all. And if the impersonal desiderata are met, we do not need to act to meet them further; these desiderata do not obtain in degrees.

²⁵ As my framing of this point may suggest, I favor some version of ethical realism, but I don’t think the point depends on it: even if, say, moral expressivism were true, there seems no reason why there should not be norms governing what the person expresses in particular circumstances, quite apart from her beliefs or the beliefs of the surrounding communities about them. (That is, even if moral expressions have no truth conditions independent of the commitments of the speaker, their expression may be subject to prudential and epistemic norms.) Moreover, there is no reason in principle why these suprapersonal desiderata could not be constructed from intersubjective materials along quasi-realist lines.

²⁶ The independence in question concerns the *grounds* of these truths, not their *reference* or *formulation*. I am only arguing that even if all human beings informed one another about their contextual visions of the good and came to an agreement about “the good lives for human beings,” this formulation might still be incomplete and even mistaken. As will become clear, I am quite ready to grant that the meaningfulness of these truths depends on their translations into limited, contingent contexts. Further, these truths are, no doubt, multivalent; I am not suggesting that any claim as simple as, e.g., “the good human life is the life of study” is true or plausible.

LYCEUM

reference to a socio-historical context or the norms it espouses. As Charles Taylor (1989) remarks, questions like *who am I?* and *who should I be?* find their meaning for us in intersubjective dialogue (see p. 35). Further, to be a *good human agent* is, in part, to achieve excellence in various elements of one's social life, including comporting oneself with norms particular to one's communal situation. But this condition, of course, has no actual content when abstracted from all particular social contexts, and it is just these contexts from which the bare self is removed.

Now Galen Strawson might retort that this is precisely the problem: that these facts only highlight the all-encompassing influence of environment that gives rise to the problem of the Basic Argument in the first place. But Taylor (1989) disagrees, contending that the social embeddedness of questions of identity "cannot be just a contingent matter, [since] there is no way we could be inducted into personhood except by being initiated into a language" (p. 35). And language, in turn, cannot be divorced from its social context. So even if moral facts are not socially contextual, both moral language and moral life are—and so the *causa sui* self exists outside both.

Second, even granting that some moral facts are *eo ipso* binding reasons for all rational agents, it seems doubtful that all goods that strongly influence our moral agency can be translated into this "dialect of universal reason." Suppose a person is deeply concerned about which vocation will be hers. In virtue of the particular vocation she pursues and about which she is concerned, she will be drawn to behave in particular morally significant ways rather than others and to achieve particular moral goods rather than others. So the question of which vocation an agent will pursue is a question that is pivotal for her moral agency. But no plausible universal moral law will dictate that she ought to pursue any particular vocation. This is why even Kant, who argued for as strong an association between the dictates of morality and those of universal reason as anyone, recognized (roughly) that moral agents require contingent desires to fill in the gaps left by the moral law, if they are to get around to any concrete moral

Two Sorts of Self-Creation: On Galen Strawson's "Basic Argument"

action.²⁷ But as the bare self is removed from the grip of all contingent desire, it loses the ability to determine its will sensibly toward particular moral actions.

Third, not only does the performance of moral actions depend on contingent particularities, but the *moral character* of moral actions is similarly context-dependent. For if one lacks all such particular interests and cares, one cannot pursue any good with the affective dispositions necessary for genuine moral action.

Here a less abstract formulation may help. Marriage is a form of community that normatively carries with it particular moral goods; people who are married ought to be loyal, self-sacrificial and encouraging toward one another in distinctive ways. Presumably, *ex hypothesi*, the bare self recognizes these moral features of marriage; let us assume (somewhat fatuously) that, in virtue of these features, it decides to marry. How would it decide whom to marry? Stripped of contingent desires and contexts, it could do so only by considering the magnitude of the moral goods to be achieved by marrying any particular person (i.e., who would benefit the most from its loyalty, who would benefit *it* the most by their self-sacrifice, etc.). But I submit that this sort of calculation simply is not a morally acceptable basis for making this decision—not to mention that, in practice, it seems questionable if such considerations could actually sufficiently motivate even a *causa sui* human being to behave morally in these ways over a long period of time. The moral basis for devoting oneself to marriage, I argue, is affection for one's partner *considered as a particular*, and while this affection must be sensitive to the (morally salient) universals involved, it cannot be exhaustively expressed by reference to them. And, I submit, similar remarks apply to the selection of our vocation, our

²⁷ See *Religion Within the Bounds of Mere Reason*, Bk. I. More precisely, Kant recognizes that human beings are necessarily motivated by their desires and requires only that moral agents subordinate their self-serving desires to their desire to follow the moral law for its own sake. There is never any question of the moral agent's quelling all "non-dutiful" desires within her, since this is neither possible nor productive for human beings. Indeed, the very concept of an imperfect duty makes clear the great role that contingent experiences and desires have to play in moral agency.

LYCEUM

interests, the causes important to us, and so on. Moral agency, then, is motivated by loves with irreducibly particular dimensions. But our loves for bare particulars must be fixed by the contingent contexts in which we are embedded.

In summary, the question of the good *for us*—us contingent, contextually limited human agents—cannot be simply a question of the Abstract, Timeless Good. Taylor (1989) perceptively asserts that “our orientation in relation to the good requires not only some framework(s) which defines the shape of the qualitatively higher but also a sense of where we stand in relation to this” (p. 42). The arguments above show that even if we could consider outside our limitations the question of where we should stand, we would not be able to find an answer. For any space in which we could place ourselves would fall within the limitations we would have transcended, and selection between such spaces is possible only by entering into them. The question of where we stand with respect to that which is valuable is answerable only if we are given somewhere to stand, helped (and *made*) to begin moving, and finally prompted to select—increasingly, though never fully, on our own—between the paths that confront us. By preventing us from beginning in interpersonal space, total self-creation precludes our attaining not only the interpersonal desiderata of agency but also the suprapersonal ones.

So, even if total self-creation allows us more responsibility for ourselves and thus achieves the personal desiderata of agency better than modest self-creation, I argue that it prevents us from achieving goods at least as vital (and probably more so). But I doubt that even this much can be said for total self-creation. For whatever sorts of control over oneself being *causa sui* may involve, it structurally rules out being influenced by one’s context in ways that aren’t already latent in one’s prior self. And, insofar as it stipulatively rules out one’s acting in ways that are not determined by one’s deep character, it precludes not only one’s reformulation of oneself but also one’s coming to identify with oneself through protracted struggle against *akrasia*. But if the self one has is not rendered intelligible by any particular context or location, and if one does not come to identify with it through genuine development or through struggling to

Two Sorts of Self-Creation: On Galen Strawson's "Basic Argument"

express one's fundamental values in action—in short, if one does not have a history, a narrative by which one becomes intelligible to oneself—how could one identify with or take responsibility for oneself, the strong sort of control one has over oneself notwithstanding? How could one fail to experience alienation? Taylor (1989) is right to claim that “to ask what a person is, in abstraction from his or her self-interpretations, is to ask a fundamentally misguided question” (p. 34). But it is one's *story*, not one's control over oneself, that allows one to interpret oneself.

Conclusion

If the argument of the preceding section is correct, then we have strong evidence against Galen Strawson's claim that the understanding of responsibility for self found in the Basic Argument is central to ordinary thought about morality and justice: namely, that only self-creation of a very different sort can provide us with the goods we seek in moral agency. We should not be distressed, then, if Strawson's ultimate responsibility is unattainable, for the self-creation that mature humans practice better allows for meaningful agency and identification with self.

The argument of this paper, then, provides libertarians and compatibilists alike with the necessary resources to found a defense against a prominent skeptical argument. But I think it does this by stepping back and considering the points that are indispensable to any account, libertarian or compatibilist, that would accurately describe the agency we value—namely, that meaningful agency develops out of the influences of heredity, environment and community; through the interplay of volitions and actions; and by our gradually coming to take responsibility for our selves and to reflectively articulate our senses of what is fundamentally valuable. If a theory of free will, from whichever side of the

LYCEUM

divide, takes these points as pivotal, it will not fail to capture much that is deeply important concerning human life and activity.²⁸

References

Arpaly, Nomy. (2006). *Merit, meaning, and human bondage: An essay on free will*. Princeton, NJ: Princeton University Press.

Dennett, Daniel C. (1984a). *Elbow room: The varieties of free will worth wanting*. Cambridge, MA: The MIT Press.

———. (1984b). “I could not have done otherwise—so what?” Edited version reprinted in Kane (2002a), pp. 83-94.

Fischer, John Martin. (1999). “Responsibility and self-expression.” Reprinted in Fischer (2006), pp. 106-23.

———. (2005). “Free will, death, and immortality: The role of narrative.” Reprinted in Fischer (2009), pp. 145-64.

———. (2006). *My way: Essays on moral responsibility*. New York: Oxford University Press.

———. (2009). *Our stories: Essays on life, death, and free will*. New York: Oxford University Press.

Frankfurt, Harry G. (1977). “Identification and externality.” Reprinted in Frankfurt (1998), pp. 58-68.

²⁸ My thanks are due to Mark Talbot, Robert Bishop, Jay Wood, two anonymous referees for this journal, and especially Stephen Scheidell for helpful comments and suggestions.

***Two Sorts of Self-Creation:
On Galen Strawson's "Basic Argument"***

- . (1982). "The importance of what we care about." Reprinted in Frankfurt (1998), pp. 80-94.
- . (1987). "Identification and wholeheartedness." Reprinted in Frankfurt (1998), pp. 159-76.
- . (1998). *The importance of what we care about*. New York: Cambridge University Press.
- Ginet, Carl. (1989). "Reasons explanation of action: An incompatibilist account." *Philosophical Perspectives* 3, pp. 17-46.
- Hobart, R. E. (1934). "Free will as involving determination and inconceivable without it." *Mind* 43(169), pp. 1-27.
- Kane, Robert. ed. (2002a). *Free will*. Malden, MA: Blackwell Publishers.
- . (2002b). *The Oxford handbook of free will*, 1st ed. New York: Oxford University Press.
- Nagel, Thomas. (1986). *The view from nowhere*. New York: Oxford University Press.
- Strawson, Galen. (1994). "The impossibility of moral responsibility." Reprinted in Watson (2003), pp. 212-28.
- . (2000). "The unhelpfulness of indeterminism." *Philosophy and Phenomenological Research* 60(1), pp. 149-55.
- . (2002). "The bounds of freedom." In Kane (2002b), pp. 441-60.

LYCEUM

- . (2004). “Free agents.” Reprinted in G. Strawson (2008), pp. 359-86.
- . (2008). *Real materialism and other essays*. New York: Oxford University Press.
- Taylor, Charles. (1976). “Responsibility for self.” Reprinted in Watson (1982), pp. 111-26.
- . (1989). *Sources of the self*. Cambridge, MA: Harvard University Press.
- Watson, Gary. (2002). “Volitional necessities.” Reprinted in Watson (2004), pp. 88-122.
- . Ed. (2003). *Free will*. 2nd ed. New York: Oxford University Press.
- . (2004). *Agency and answerability: Selected essays*. New York: Oxford University Press.