

## Can We Ever Be Really, Truly, Ultimately, Free?

MARK BERNSTEIN

I begin with the tale of Dora who is charged with stealing clothing from a department store. After a short trial, she is found guilty and sentenced to a year in jail. Soon thereafter, her defense attorney discovers that Dora had been injected with a serum that fixed her will. The serum affected her neurophysiology in such a way as to (deterministically) cause Dora to desperately want to steal and, as a result, when the opportunity presented itself, she acted accordingly. Her defense attorney brought this new information to the judge and urged that since her client did not steal "on her own free will," she ought to be exonerated. The judge concurred and rescinded the jail sentence. A week later, however, the prosecuting attorney discovered that although Dora was indeed injected with this powerful serum, she had knowingly authorized its use. Apparently Dora had wanted to have the experience of being a thief but believed that her generally timid character would prevent her from gathering the nerve that it takes to walk into a department store and steal its wares. To ensure that she would have this experience, she contracted with a disreputable scientist to inject her with the theft-causing serum. When the prosecuting attorney brought this new information to the judge and argued that Dora was, after all, responsible for the theft, the magistrate agreed, reversed course and reinstated his one-year sentence. Fortunately for Dora, her energetic defense attorney dug still deeper into the history of the case and found out that Dora's desire to experience a theft was itself a product of another serum that had been injected into Dora without her knowledge. This serum interacted with Dora's internal chemistry to produce this irresistible desire. The defense attorney urged that, after all, Dora's theft was not a product of her own free will and so she should not be held responsible for the theft. Upon hearing this new evidence, the magistrate wearily (and warily) reversed course yet again and retracted his sentence.

The story mercifully ends here although it takes little imagination to see how Dora's ups and downs could continue indefinitely. Stripped of its legal trappings, we have what amounts to a framework for understanding what philosophers mean when they speak of "really," "truly," or "ultimately" being responsible for one's will and the actions that emanate from it. To have free will, in this uncompromised, full-blooded way, is to have the power to be the ultimate or final creator of her own intentional mental life. Each time we believe that Dora ultimately originates her will, we are inclined to ascribe free will to her; each time we think that the serum serves as final source, we are inclined to retract our attribution.

Reflections on Dora's tribulations belie the caveat issued by some that we must exercise caution in distinguishing between being "really," "truly," and "ultimately" responsible for our will and being responsible for our will *sans phrase*. These separatists urge that the modified notion is a mere philosophical contrivance whose relevance is limited to the verbal games held so dear to philosophers. The unadorned idea, on the other hand, is purveyed as the one used by *hoi polloi* and, as such, the only one with practical significance. But this none-too-subtle attack on philosophy's relevance, although sometimes with merit, here misses the mark. The story of Dora is indeed a bizarre one, but in this setting the judge's changing decisions are not only intelligible but also warranted, constituting a commentary about a single notion of responsibility. Even some who are sympathetic to the significance of "true" responsibility have, I believe, mischaracterized the relationship between being "really responsible" and simply "responsible," by conceiving of these kinds of responsibility as distinct, although, perhaps mutually entailing. My suggestion is that there is only one kind of responsibility that concerns us here. At the end of our narrative, the judge does not believe that although Dora was not "truly" responsible for her will and her subsequent actions, she was, at certain times during her tribulations, responsible for her will and concomitant actions. When we ask whether someone is "truly" responsible for her or his volitional acts, we are not conceding her or his (simple) responsibility and wondering whether it is proper to attribute yet another special type of responsibility to her or him. Rather, the question is asked when we are unsure of the will's etiology; we have doubts whether the will is, for example, drug-induced. To my mind, then, we can delete not only the scare quotes around the adjectives that describe the responsibility that concerns us but also omit the adjectives altogether. "True" ("ultimate," "real") responsibility is nothing more than true (ultimate, real) responsibility, which in turn, is nothing other than responsibility. There is no error in attaching the modifying clauses or in retaining the scare quotes as long as we are not misled into believing that we are referring to some arcane segment of our ontology when we do so.

Dora's tale engenders two sources of angst in philosophers. The first is the thought that, if we occupy a deterministic world, we are all Doras. While there are no evil inoculators that cause concern, there are states of affairs analogous to the injections Dora incurs, that, in conjunction with the laws of nature, ensure how our wills are constituted. It is not merely that we cannot have a will other than the one we have, but also that the *cause* of both the content and inevitability of our will is something beyond our say. This depressing state parallels the culmination of Dora's story where the drug brings about both what she desires and the

ineluctability of it. The second, and equally necessary cause of dread, comes from our belief that eradication of free will erases something of enormous value from our lives. Absent free will, we lose creativity and agency, desert and responsibility, dignity, and autonomy. Theists inherit an additional insult. Without these prized attributes, we release all claims to being the stars of the divine drama with our heavenly tickets to eternal salvation in hand.

1

What to do? Historically, the favored option is compatibilism. Dora's deterministic world is invested with free will but to see this we need to have a clearer understanding about the constitution of this power. Normally — although not inevitably — we are provided a hypothetical account of freedom in terms of how we would act if we were in a particular volitional state. Whatever merits this tactic has as an account of free (overt) action, it attracts enormous problems as an account of free *will*. This is not the venue for a sedulous discussion of compatibilism; I simply register my life-long dissatisfaction with any of its formulations. Resurrecting the ideas of creativity, desert, and dignity within a compatibilist framework has always struck me as requiring scare-quotes about the pertinent terms. Even the grandest efforts never seem to produce the robust type of free will that I find most worth having.

We segue to a second option, one gaining in popularity over the last two decades. Libertarianism concedes that determinism precludes responsibility over our wills but that, nevertheless, we do have free will and in the full-blooded sense that is necessary to wholeheartedly ground our most prized values. Thus, Dora's world cannot be our world. We live instead in something akin to Linda's world. Linda's case remarkably parallels Dora's up to and including the point where her defense attorney finds that she, too, was unknowingly injected with a drug. However, unlike Dora's inoculation that determined her desire to steal, Linda's inoculation inserted indeterminism into the volitional proceedings. The serum from this injection interacts with Linda's neurophysiology resulting in some chance of her forming the desire to steal and some chance of her not forming this desire. As it turns out, Linda does form the desire and acts upon it, stealing just as Dora had after her injection. While Dora's desire is determined (it *has* to happen given the antecedent state of affairs and the governing laws), Linda's desire is undetermined (it need not happen given the antecedent state of affairs and the governing laws).

Place yourself as the magistrate in Linda's case and ask whether, at this juncture, rescinding the original one-year sentence as the judge did with Dora is the proper course of action. It would seem that without more information any decision is arbitrary. The only fact that Linda's inoculation ensures is that there is some possibility that she will form the desire to steal and some possibility that she will not form this desire. This is how non-philosophers think of virtually all the actions we do perform. It is the rare case when we think that either some volitional event or some ensuing overt act is one that *had* to occur given the antecedent state of the world prior to these events. Common sense, unlike much philosophical reflection, suggests that free will occurs only in indeterministic contexts. Moreover, that

there is some probability of Linda not forming the desire to steal does not, in itself, militate against the notion that her desire—were she to form it—was not “up to her” or under her control. This is true even if the probability of forming the desire is relatively small.

Why have philosophers traditionally thought that an indeterministically produced—and therefore undetermined—act cannot be under the control of an agent, and so, that all libertarian theories of free will that essentially depend upon the viability of such a notion are doomed from the start? The reasoning is initially clear and compelling. To say that Linda’s decision to steal is undetermined is elliptical for saying that her decision to steal is undetermined by any and all conditions that preceded her decision. This means that it is possible that the world is identical up to the very moment that Linda decides to steal and yet, contrary to what actually is the case, Linda decides not to steal. Thus, either of two incompatible actions could have occurred given identical antecedent circumstances. Linda’s decision, therefore, appears arbitrary or capricious. If either of two (or more) events could have occurred from identical precedents— from all of Linda’s thoughts, preferences, motivations, deliberations, and external circumstances— then it appears as though Linda’s input into the decision is exhausted. It seems no longer “up to her” whether she decides to steal or decides against it. “Undetermined free mental event,” in the sense entailing voluntary control of a piece of personal behavior, is an oxymoron.

Consider a gambler throwing a pellet into a legitimately randomizing roulette wheel. Her toss is compatible with (say) two possible alternatives. That is, an exact duplicate of the world up to the instant the pellet landed in slot 32 is consistent with the pellet resting in slot 33. The anti-libertarian invites us to realize how silly it would be to give the gambler credit for selecting the correct slot 32; for everything she did, it could have landed in losing slot 33. We are to see Linda’s decision on a par with the gambler’s winning selection. Just as indeterminism ensures that the gambler’s winning toss was not “up to her,” but rather a product of unadulterated luck, so too is Linda’s decision to steal nothing more than a fluke. In neither case, do we have the control of an action necessary for the robust free will that we desperately seek.

## 2

Until very recently, this criticism of libertarianism seemed insuperable, at least insofar as scientific respectability constrains our answers. Allowing ourselves magic—in the form of Cartesian egos, noumenal selves, transemprirical power centers, and the trendy agent (i.e., non-event) causation—we can “solve” most any problem. Enter Robert Kane. His theory, developed over some 30 years, is perhaps best epitomized in the tale of Betty the businesswoman.<sup>1</sup> Betty is on her way to

1. Kane’s tale of the businesswoman—unnamed in his works—appears repeatedly. See, for example, “Free Will: The Elusive Ideal” in *Philosophical Studies*, vol. 75, nos. 1–2, 1994, pp. 44 ff; *The Significance of Free Will* (New York: Oxford University Press, 1996), pp. 126 ff; “Some Neglected Pathways in the Free Will Labyrinth” in *The Oxford Handbook of Free Will*, ed. R. Kane (New York: Oxford University Press, 2002), pp. 417 ff.

an important business meeting when she happens upon an assailant attacking a young woman. In virtue of her complex character, Betty's will is motivationally split. On the one hand, she is ambitious and knows that the meeting is crucial to her career. On the other hand, she is a moral person and so is inclined to try to ward off the assailant. An internal struggle ensues, the outcome of which is metaphysically uncertain. Eventually she decides to help the victim at the expense of her career. This decision is undetermined; she could have decided to continue to her meeting. This "could have" is to be understood in the strong, non-hypothetical libertarian sense; at that very moment with the history of the world congruent to the way it actually is, it is physically possible for her to decide either to go to her meeting or help the victim against her assailant. Despite the fact that an indeterministic process (the conflicting efforts of will) produced her decision, she has *plural voluntary control* over her set of options; that is, she can bring about both options (helping the victim, continuing to her meeting) voluntarily (i.e., willingly), intentionally (i.e., knowingly), and rationally (i.e., having reasons to do so). She has control over what she does, since she can bring about either of her options voluntarily or at will as a result of her efforts. A bit more laboriously, whatever Betty decides, the decision is "up to her" because she can do what she most wants to do, when she wants to do it, for the reasons that she wants to do it, where her decision is neither coerced nor compelled. There is a type of control—*antecedent determining control* that necessarily the businesswoman lacks regarding her self-willed choice since there is nothing prior to her decision that suffices for the occurrence of one act rather than the other. Nevertheless, the lack of this sort of control does not eradicate the possibility of acting from one's own free will, an action that paradigmatically occurs in the making of *self-forming actions* as exemplified in Betty's choice to help the victim of the assailant. She passes the tests for being personally responsible for her moral act, an evaluation that is confirmed by the fact that she "takes responsibility" for her act. She endorses (accepts the decision as a product of her will) the choice and, from a phenomenological point of view, the act feels as if it came from her and was not forced upon her from some alien source. If she had instead decided to bypass the victim and act self-interestedly (a possibility that is ensured by the fact that her decision to aid the assailant is undetermined), she again would have done so voluntarily, intentionally, and knowingly, although, of course, the reasons for this act would have been in terms of her ambitious motivation. Had this option occurred, she would have been responsible for it in the same way that she bears accountability for the moral act that she actually did perform. In self-forming actions, the nature of a person does not suffice (i.e., determine) what the agent will do but, as the case of Betty makes clear, undetermined actions need not be random or lucky in any sense that precludes personal responsibility.

After her self-forming decision, Betty's newly formed character undoubtedly determines numerous behaviors. Nevertheless, these actions can truly be ones for which she bears full responsibility since, although they are determined by her character (and so, given her character at the time, she could not—again in the strong libertarian sense—do otherwise than perform the act she does), her character was one that *she*, at least in part, created. We can see Luther's break with the church

in this way. When he announced that he could do no other, he spoke literally. Given his character at the time of pinning the 95 theses to the church door, his act was determined. However, he is responsible for his act since the character that he had at the time descended from prior self-forming actions that he brought about voluntarily, intentionally, and rationally. Luther is responsible for his break with the church even if his act is determined at that time because he freely brought about undetermined acts that came to constitute a relevant segment of his present nature.

The indeterminism situated in the efforts that terminate in Betty's self-forming action hinders both her attempt to help the victim and her attempt to get to her meeting; In order to decide to continue on her walk to her meeting, Betty would need to successfully resist the motivation to help the victim; in order to decide to help the victim (the choice that actually occurred), Betty would need to successfully resist the inclination to continue walking. But it is important to note that this indeterminism is a product of her will and not of some external source. Subtract one of the two sets of conflicting reasons and we are left with no obstacle for her to overcome, and, as a result, there is a sense in which her control is total. But the price for this is that she can no longer rationally or voluntarily choose the other course of action, for there are now no distinctive motivations on which rational and voluntary action can be founded. Kane establishes a high threshold for the type of control required for these paradigmatic free acts that create one's self. The agent needs not merely "one-way" control of the sort Betty would have were she to lack competing motivations and desires, but requires plural voluntary control where she is able to purposively bring about any of the undetermined options when she wishes, without being coerced or compelled, and for the reasons that she has at the time of her choice. As it turns out, indeterminism, although a hindrance to one-way control, is absolutely necessary for plural control and, consequently, is necessary for individuals to be ultimately responsible for any of their acts.

Speaking to the innovation and importance of Kane's theory, a barrage of attacks has unsurprisingly followed. I do not spare myself from this chorus, but I now believe that it largely misses a significant mark.<sup>2</sup> The heart of the free will problem remains even if we accept Kane's attempt to show that we can legitimately said to be in control of undetermined acts. At the end of the day, the mode of production—deterministic or indeterministic—is not the crucial fulcrum about which the free will issue pivots. The problem, instead, is one of ultimate origins; to be ultimately responsible for one's actions requires that one be ultimately responsible for the person who performs those acts. There is a convincing argument to the effect that only by accomplishing the impossible—by being a *causa sui*—could we ever be free in the fundamental sense that we require. We now turn toward these considerations, show why Kane's libertarianism does not help answer this problem, and speculate how, if at all, a solution may be framed.

2. See "Lucky Libertarianism" co authored with M. Almeida in *Philosophical Studies*, vol. 113, no. 2, 2003. I am now less than thoroughly happy with this article for other reasons as well.

## 3

Galen Strawson has forwarded this pessimistic conclusion in varied guises over a long period of time.<sup>3</sup> The idea, as Strawson himself admits, is neither greatly original nor esoteric. We begin by noting the apparently innocuous point that persons act the way they do because of the way they are (at least in certain significant mental respects such as one's preferences, beliefs, intentions, and the like). Ultimate responsibility for a person's acts (in particular, her will) requires that the individual be ultimately responsible for how she is. (This is why when we believed that Dora's inoculation caused her to be a certain way, we were inclined to relieve her of ultimate responsibility for her decision.) However, we cannot be ultimately responsible for the way we are. Why not? To be ultimately responsible for one's nature, one must intentionally bring it about that you are that way (that is, your mental life needs to be *your* knowing and willing product, and not say, a product of drugs, hypnosis, another agent, or even from inanimate forces). But to intentionally bring about this nature, Z, say, a prior nature Y, with certain significant mental states such as preferences, desires, beliefs, and intentions, must exist by virtue of which Z was intentionally brought about. But, to be responsible for Z, you must also be responsible for Y, and this requires a still prior nature X that, in virtue of its important mental states, Y becomes intentionally produced. This reasoning—which appears unexceptionable—sets off an infinite regress. The only solution is the impossible one of ultimate self-origination; we would need to be able to be responsible for the way we are without already possessing the mental life in virtue of which responsibility for our nature emanates. But this is to say, at least for mere mortals, that there is no solution and so ultimate responsibility is an unexemplifiable ideal.

Kane's theory, ingenious as it is, provides little help. Recall our intrepid Betty who—according to Kane—is ultimately responsible for her undetermined decision to aid the assailant's victim. The key to this ultimate responsibility is that both this decision and the optional choice she could have made but did not (i.e., continue to her business meeting) meet the conditions of plural voluntary control. Let us grant that either alternative act is one that Betty could voluntarily, intentionally, and rationally perform. But, obviously, for Betty to perform either of these possible acts in these ways is a function of her mental life at the time, that is, of the way she is (in certain crucial mental respects) when she decides to help the victim of the attack. We are, then, back in the throes of the regress, for if the voluntary, intentional, and rational performance of an act is to count in favor of an agent being responsible for that act, the agent would need to be responsible for the nature that allows her to perform acts in these ways. And this, as we have seen, requires yet another prior character from which the present character is responsibly generated.

3. Galen Strawson has argued for the soundness of what he calls the "Basic Argument" in many venues. See *Freedom and Belief* (Oxford: Oxford University Press, 1984); "The Impossibility of Moral Responsibility" in *Philosophical Studies*, vol. 75, 1994; "The Unhelpfulness of Indeterminism" in *Philosophy and Phenomenological Research*, vol. 490, no. 1, 2000; "The Bounds of Freedom" in *The Oxford Handbook of Free Will*.

And so it goes. Without the possibility of self-origination, there is no possibility of ultimate responsibility, and the terms “voluntary,” “intentional,” and “rational” that describe self-forming acts that comply with the conditions of plural voluntary control *should* be understood as inheriting scare-quotes.

By looking at Kane’s somewhat technical account of *ultimate responsibility* (one that tries to track the commonsense notion that drives our assessments of Dora and Linda), we can see the lacuna from a different perspective.

(UR) An agent is *ultimately responsible* for some (event or state) E’s occurring only if (R) the agent is personally responsible for E’s occurring in a sense which entails that something the agent voluntarily (or willingly) did or omitted, and for which the agent could have voluntarily done otherwise either was, or causally contributed to, E’s occurrence and made a difference to whether or not E occurred; and (U) for every X and Y (where X and Y represent occurrences of events and/or states) if the agent is personally responsible for X, and if Y is an *arche* (or sufficient ground or cause or explanation) for X, then the agent must also be personally responsible for Y.<sup>4</sup>

Focusing on (U)—the “ultimacy” clause—we see Kane’s sensitivity to the commonsense notion of ultimate responsibility that is required for the type of free will that grounds the goods and values of creativity, autonomy, agency, and desert that we rightly laud; agents must be responsible for any state of affairs that suffices for any mental state. Kane, therefore, views determinism as precluding ultimate responsibility since there are determining events (“sufficing” events) for our mental lives (as well as for every other event or state of affairs) for which we are not responsible. All we need do to supply uncontentious examples is consider events that occurred before the agent was born. Kane must revert to an indeterministic world where, somehow and somehow, our mental lives do not have sufficient causes (thus breaking this regress) and yet—and here is the sticking point—are responsibly ours. Kane’s work can be viewed as an attempt to show how this can be accomplished absent verbal chicanery and occult entities. Unfortunately, the ultimate, Strawsonian problem is left untouched. While we can accept, *arguendo*, that our character or nature *locally* allows us to be responsible for some undetermined states of mind that eventually will lead to actions, and even accept that there are undetermined segments of this effective character of which we can be *regionally* responsible (deriving, as they do, from prior natures), without having our original natures being responsibly created—an apparent impossibility—we can never attain *global* (“ultimate” in the sense Kane himself recognizes) responsibility for any of our mental life or its consequent overt acts.

Our first nature or character cannot be derived from a preceding character; so much is guaranteed by being the original. On a positive note, we can confidently claim that we develop a nature—a constellation of mental states including beliefs, preferences, intentions, and the like—at an early stage in our lives by virtue of inherited qualities and early environmental factors, neither of which can be under

4. UR has undergone various minor changes throughout Kane’s writings. I use the one recorded in *The Significance of Free Will*, p. 35.

our intentional control (for this we would need an antecedent character). From these humble argumentative beginnings, free will appears precluded at any time during an agent's life. Whether the agent's subsequent mental furniture is deterministically or indeterministically produced, absent prestidigitation, there is no means by which we can have (global) say over its contents. We need, at the end of the day, to have intentionally, knowingly, and willingly brought about our initial nature, and this, even if somehow a possibility for deities, is beyond our mortal grasp. The essential impossibility of moral responsibility rests on the idea that we cannot be responsible for any act that ultimately results from a nature that is not a product of our responsible choice.

4

Can we summon help from anywhere? Some may think that Harry Frankfurt's work may prove useful. Let us accept the verdict that Frankfurt examples demonstrate that the ability to perform an act other than the one that is actually performed is not required for moral responsibility. Let us further concede that this shows that legitimate attributions such as praise, blame, dignity, and autonomy can attach to a person in virtue of her performing an act for which there is no alternative. Let us even go so far as to admit that this shows that we can be held responsible for our original nature despite the fact that we could not have chosen a different one. Unfortunately, none of this helps with our conundrum. If we are inclined to agree with Frankfurt's assessment of his cases and believe it justified to assess the agents as morally responsible, we are assuming that, in fact, the agent freely decided to act as she did. We hold Judy morally responsible for killing Bill, despite the fact that an inactive mad scientist would have intervened to ensure her murderous decision had she wavered, because we believe that Judy's will was "her own." If we discovered that some madman who injected her with a kill-Bill serum caused her lethal desire, we presumably would redact our evaluation, just as we did with Dora in her bizarre circumstance. Frankfurt yields no help in countering the Strawsonian insight that in the formation of our original natures by events foreign to our control we are all Doras, and so are forever distanced from buck-stopping responsibility. Free will deniers impute the lack of ultimate responsibility not to the fact that an agent has no alternative original character from which to choose—the situation that Luther found himself in on All Saints' Eve in 1517—but rather to the fact that no agent can be in control of what constitutes his original nature.

We need to entertain a bold step that questions one of the apparently unavoidable assumptions. Perhaps we can make some headway by rejecting the notion that control and responsibility of our character must derive from a preceding nature.<sup>5</sup> If we can make sense of the idea of the simultaneous occurrences

5. An avenue that some may find worth exploring involves the notorious concept of retro-causation. If the intelligibility (and practical possibility) of backward causation can be made out, then we may start a defense of being in control of our initial nature subsequent to the time it begins to exist. The suggestion here is that our initial nature is caused by a later character of which we have control and responsibility in virtue of it emanating from our initial nature and our intentional and deliberate choices. Although I am less skeptical than many about making sense of back-

of our original nature and our control of it, the regress would be terminated. We need a consistent and somewhat plausible story about how we can intentionally, purposively, and knowingly bring about what we originally are like, where the bringing about of our nature occurs at the same time that we begin to have one.

*Pace* Hume, who believed that causes always temporally precede events, we now recognize this temporal sequencing as nothing more than pure stipulation. There are many examples of causes and effects occurring coincidentally; indeed there are plausible candidates for simultaneous mutual causation. Consider two moving bodies whose motion is affected solely by gravitational forces of the other body. At some instant of time, the infinitesimal motion of one body affects the gravitational force of the other and conversely. Thus, at a given instant, the state of a body is both the cause and effect of a state of another body.<sup>6</sup> A more pedestrian example is provided by two playing cards, each at 45-degree angles to the ground supporting each other. Each card's position is brought about and sustained by the other card's position. While being careful not to invoke intentionality, we can say that the position of each card is controlled by the position of the other, or that each card's position is responsible for the position of the other, or finally, that each card's position is "up to" the position of the other card.

In addition to providing examples of mutual simultaneous causation, these cases, conjoined with the assumption that causation is transitive, give us a way to understand how self-creation is possible. If the position of card B that, at the very same moment, causes the position of card A, transitivity of causation yields that the position of A is the cause of the position of A (A causes B, B causes A, so by transitivity, A causes A); that is, we have a case where a state of affairs causes itself to exist.

Suggestive as these cases may be, it may be insisted that more is required. We need to insinuate the notion of intentionality; the idea that our self-creation is up to us. It is one thing to grant the intelligibility of things bringing themselves into existence, even bringing ourselves into existence; it is another to claim that this self-creation can be purposively and deliberately performed. Some may be tempted to reject even this supposition and suggest that all we need to block the regress and get a robust sense of free will off the ground is to account for the self-creation of our original nature. Responsibility for our original character comes not

ward causation, I will not pursue this line of defense, for even if a persuasive case can be made for intelligibility of the concept in general, it strikes me as implausible, in the extreme, that it can be applied to the case at hand. The source of the control we have over our initial character is some later character, that itself is controlled by us in virtue of its relation to the earlier character. Thus, the present problem would involve a particularly difficult instance of backward causation since unlike "ordinary" examples of backward causation, we here have a case where each of the two states is both cause and effect of the other. The transfer of responsibility runs both backward and forward between the antecedent and subsequent natures. This seems far more problematic than the simpler case where a temporally later event brings about an earlier event that does not, in turn, bring about the subsequent event. And, even if this complication of "causal loops" can somehow be incorporated in an intelligible notion of causation, I find it beyond miraculous that we are the type of creatures capable of effecting such a sequence of events.

6. See Quentin Smith's "The Reason the Universe Exists Is That It Caused Itself to Exist," *Philosophy*, vol. 74, 1999, for other ways in which self causation may be possible.

from it being rationally, voluntarily, and intentionally selected, but from our subsequent *endorsement* of the character that is forced on us by genetic and environmental factors. Suppose that Frank is unwittingly injected with a serum that causes him to become a huge Yankee fan. Although thoughts of rooting for the Yankees never crossed his mind, after a short period of time he accepts this preference and even identifies himself with it; in conversations, Frank proudly announces himself a staunch supporter. Just as Frank did not freely choose a mental state that he comes to call his own, the free will apologist believes that the subsequent endorsement of a nature suffices for responsibility of it despite its initial non-voluntary and non-intentional acquisition.

This will not do. The quickest way of dealing with this tactic is to question the efficacy of endorsing a prior nature. Endorsing a constellation of mental states obviously requires having certain other mental states, but if these latter (i.e., the endorsing mental states) are not under the control of the agent, the process—at least as far as culminating in making particular mental facets one's own—is a sham. The Strawsonian concerns highlighting the difficulty of having a free will at any point in an individual's life apply with full force to these endorsing beliefs and desires that are, after all, just a proper subset of an agent's mental life. If Frank's endorsement of Yankee fanaticism is a result of beliefs and desires that themselves are products of an unwitting injection, then, it would count for naught as far as manifesting some truth about what Frank is truly like. We have made no progress in answering the charge that since our original nature results from forces beyond our control, no subsequent mental events can ultimately be labeled as our own.

We remain with the quandary of inserting voluntariness, rationality, and intentionality into the self-creation of our original character. The truly troubling problem is not the self-creation *per se* (at least here there are some meaningful analogies on which to draw), but rather how an agent can be responsible for the initial creation of his nature.

But why—granting the viability of self-creation—should this be found especially mysterious? While a special explanation would be required to explain how a plant, say, can control its behavior even conceding self-creation—after all the plant lacks any mental characteristics—what we are and what we are creating just is our initial nature that comprises beliefs, desires, intentions, motives, and other psychological traits. Presumably, there is no question about the nature of control; there is, for the sake of this discussion, at least, no debate about control being a function of these psychological states functioning in certain ways. The deep question is how we can be in ultimate control of this entire process; how we can have our original natures be up to us. Perhaps the chasm between self-creation and intentional self-creation is not unbridgeable. Here's a rudimentary story.

We enter the world without a nature. In the relevant sense, then, we are not a "we" yet. By virtue of our genetic endowment and environmental factors, we may own, even at this time, dispositions to form a particular nature, but there is sufficient plasticity so that our character can instantiate a very large array of options. There is no first moment at which a person has a character, much like there is no first moment when someone begins to read. Just as mimicking and parrotting eventually evolve into reading, nature-like phenomena ("proto-choices,"

“proto-intentions,” “proto-beliefs”) eventually develop to full-blown choices, intentions, and beliefs. Like their mature counterparts, these proto-phenomena lack any sharp point of origin. Proto-control evolves during this indeterminate period of time. (Perhaps this proto-control is something like the “control” a very simple animal has over its movements or even the “control” that a plant has when it bends toward the sun.) Over time, complexity increases resulting in a nature that is caused and shaped, at least to some extent, by us, that is by the motivational structure that constitutes our nature. Thus, within the limits set by genetic and environmental factors over which we have no control, the options we pick that make us what we are, are intentionally and purposively selected by us. What is being intentionally molded is that which is doing the intentional molding.

In fact, the very idea of a fixed or determinate nature is likely a fiction. Talk of an initial nature being followed by subsequent natures suggests that our mental lives are a series of whole and complete characters. A more realistic picture may be that our natures constantly evolve and so are better conceived as processes rather than substances. Grammar misleads here; rather than having a nature, “naturing” may be something we do over the course of a lifetime. If this is right-headed, then the “self-creation of our initial nature” is still not a sufficiently radical departure from the suppositions of the free will denier. It may be best to think of ourselves as continually creating ourselves; constantly, and for the most part, responsibly, making ourselves what we are.<sup>7</sup>

7. My gratitude to Mike Almeida for comments upon an earlier draft.