

FREE CHOICE AND MORAL CHARACTER: A DIFFICULTY FOR LIBERTARIANS

Many libertarians now concede to the compatibilists, as I believe they should, that an action can be free even when determined by an appropriately formed character, and their intuition seems to be that an agent's character is appropriately formed only when the agent is at least partly responsible for it. James F. Sennett thus writes as if we sometimes choose our own character: "A character that is libertarian freely chosen is the only kind of character that can determine compatibilist free choices."¹ Laura Ekstrom likewise suggests that our judgment that an action is praiseworthy "may presuppose the idea that the agent's good character is ultimately of his own making . . ."² And Robert Kane explores the idea of a "self-forming action" in great detail and with considerable insight.³ But in this paper, I shall raise some questions concerning this strange idea that, in the words of Kane, free agents "make themselves into the kinds of persons they are" and pose something of a challenge to it as well. I shall also challenge the common libertarian assumption that a bad character is typically the product of bad choices even as a good character is typically the product of good choices. For more often than not, the effects that our choices have upon our own character are neither predictable nor under our own control.

¹James F. Sennett, 'Is There Freedom in Heaven?' *Faith and Philosophy*, **16** (1999), p. 74.

² Laura Ekstrom, *Free Will: A Philosophical Study* (Boulder, Colorado: Westview Press, 2000), p. 165.

³ See Robert Kane *The Significance of Free Will* (London: Oxford University Press, 1996), ch. 5.

Just what might it *mean*, in the first place, to say that someone has made, or formed, or produced his or her own character? Robert Kane speaks of certain “voluntary ‘self-creating’ or ‘self-forming’ actions (including refrainings) in the life histories of agents for which the agents are personally responsible.”⁴ These self-forming actions (or SFAs), says Kane, are “both undetermined...and such that the agents willingly performed them and ‘could have voluntarily (or willingly) done otherwise’”⁵ But though undetermined—and, as some might say, self-generated or self-originated—they are also *self-forming* in the sense that they help to determine or shape the agent’s present motives, purposes, and character traits: “Agents with free will...must be such that they could have done otherwise on some occasions of their life histories with respect to some character- or motive-forming acts by which they make themselves into the kinds of persons they are.”⁶

It is not, however, the supposedly indeterministic origin of these self-forming acts (or SFAs) that here concerns me. Where “UA” is shorthand for “an undetermined action such that the agent who performed it categorically could have done otherwise,” my present concern is with the relationship between the UAs in one’s life history and one’s present character. Just what must that relationship be in order for the latter to be considered a product of one’s own making? Kane suggests that the relevant UAs must make “*a difference in what you are* (or in the character and motives you now have).”⁷ And in a similar vein, Laura Ekstrom appeals to the idea of a *partial* causal explanation, or a contributing cause, when she offers the following necessary condition of moral responsibility:

⁴ *Ibid.*, p. 75.

⁵ *Ibid.*

⁶ *Ibid.*, p. 72.

⁷ *Ibid.*

A person S is morally responsible for doing X at *t* only if *either* S could have done otherwise than X at *t* or S's inability to do otherwise than X at *t* is causally explicable at least in part by S's own act(s) at some time(s) other than *t*, such that S could have done otherwise at that (those) other time(s).⁸

Although Ekstrom herself finally rejects this principle on account of Frankfurt-type counterexamples,⁹ I would here emphasize just how easily a determined action can satisfy it. Suppose that a woman has only one UA in her life history, namely her decision as a youngster to spend her allowance on swimming lessons rather than on violin lessons. If that UA partly explains why she later became an expert swimmer, indeed an Olympic champion, rather than a concert violinist, and if her swimming expertise partly explains why she found it unthinkable and therefore psychologically impossible to stand by as a child drowns—why she leapt into a dangerous river in an effort to save the child—then she evidently meets the Ekstrom necessary condition of being morally responsible for a determined action. But I doubt that many libertarians would regard this one UA buried in the woman's past as having any relevance at all to the praiseworthy character of her present action.

Of course Ekstrom never intended for anyone to treat her necessary condition as a sufficient condition. But elsewhere she speaks of Honest Abe and suggests that he “deserves praise only if we assume that he is now of such a constitution that he is unlikely to lie because of the

⁸ Ekstrom, *op. cit.*, p. 210. Ekstrom uses the expression “at some time(s) other than *t*” rather than “at some time(s) prior to *t*” so as not to exclude by assumption certain time travel cases involving backwards causation.

⁹ According to Ekstrom, “It is reasonable to expect the absence of Frankfurt-type counterfactual interveners.” So to handle the Frankfurt-type counterexamples, she revises her principle in the following way: In a case where S is unable to do otherwise than X at *t*, S is morally responsible for doing X at *t* only if at some time or another S could do something “that would reasonably be expected to have the result that S would do otherwise than X at *t*” (p. 211). But since this revision adds nothing further, so far as I can tell, to the idea that “the agent's good character is ultimately of his own making”—indeed, it serves only to weaken this idea further—I shall not deal with it here, however important it may be in its own context.

past free acts—in an incompatibilist sense—that have shaped his character.”¹⁰ And that is just my question: What does, and does not, count as a UA in the past actually shaping one’s present character? Suppose that a young and somewhat irresponsible married man succumbs to temptation and falls into a rather frivolous affair; suppose also that his wife subsequently finds out about the affair and seriously considers divorcing her husband on account of this and other irresponsible actions on his part; and suppose, finally, that the young man then comes to appreciate what he is about to lose and, terrified by the prospect of losing the wife he genuinely loves, feels utterly compelled to re-establish a relationship of trust. Once his wife finds out about the affair, in other words, it is fully determined that he will change his wayward ways; never again, let us suppose, does he even consider an affair, lest it undermine the very relationship that he now values so highly. If the man’s decision to have an affair qualifies as a UA, then this UA may have made a huge difference to the kind of person he eventually becomes. It also seems to qualify as a contributing cause. For had he not made his foolish choice at this precise time and in circumstances where he would eventually be caught, perhaps he would have gone through his entire adult life sneaking around and taking his wife for granted.

So do we have here a case where the free decision to have an affair helped to shape a more trustworthy and faithful character? And do we also have a case where a man acquires his faithfulness in an appropriate way? I think we do. The man’s frivolous affair was contrary to his own interests even as he perceived them; and once he discovered this or perhaps came to see it more clearly, he responded in a perfectly rational way. He also acted freely, or at least so I would argue. For not even a libertarian would deny that a determined action can sometimes be *voluntary*; and if an action is both voluntary and determined by one’s own fully rational judgment concerning the best thing to do, then it surely remains a paradigm of free action. Its being

¹⁰ *Ibid.*, p. 166.

voluntary rules out what Kane calls “constraining control,” such as being held at gunpoint, and its being determined by one’s own fully rational judgment concerning the best thing to do rules out what he calls “nonconstraining control,” such as might be “exemplified by . . . cases of behavioral conditioning and behind the scenes manipulation . . .”¹¹

Now libertarians almost always seem to presuppose a picture similar to what Kane sketches in the following passage:

The Probabilities for strong- or weak- willed behavior are often the results of agents own past choices and actions, as Aristotle and other thinkers have insisted. Agents can be responsible for building their moral characters over time by their (moral or prudential) choices or actions, and the character building will be reflected by changes in the probabilities for strong- or weak-willed behavior in future situations. Each time the [alcoholic] engineer resists taking a drink in difficult circumstances, he may strengthen his will to resist in the future; and conversely, when he succumbs, his will to resist may lessen (or crumble altogether, as sometimes happens with alcoholics).¹²

But though such a picture may reflect accurately *some* of our experience in *some* contexts—very limited ones, I believe—the way in which UAs in a life history, assuming there are such, affect one’s character and motives may be just the opposite of what Kane has imagined; worse yet, the effect is apt to depend upon intervening factors utterly outside the agent’s control.

Kane is right, of course, about the alcoholic engineer, at least partly. One biochemical effect of alcohol on an alcoholic seems to be that it undermines the will to resist another drink. So in a case such as this, it stands to reason that the longer an alcoholic continues to resist taking a

¹¹ See Kane, *op. cit.*, p. 64.

¹² Kane, *op. cit.*, p. 180.

drink, the stronger his or her will to resist might become; and the more often an alcoholic succumbs to the temptation to have another drink, the harder it will be to resist in the future. But that is merely a predictable effect of the alcohol,¹³ and it is not the whole story either. For as an alcoholic friend of mine once pointed out, the longer she stayed off the alcohol, the easier it became during times of stress to deceive herself into believing that this time a couple of drinks would do no harm; so curiously, the longer she resisted the temptation, the stronger her temptation became. Indeed, it was not until she had succumbed to temptation and had binged terribly on a good many occasions that she finally learned to recognize such deception for what it was. So in that sense, her experience was just the opposite of what Kane describes: the more often she successfully resisted temptation, the harder it became to resist such temptation in the future; and the more often she succumbed to it and experienced the destructive consequences of doing so, the easier it became to resist such temptation in the future.

So here is an obvious case where some bad choices helped to undermine a bad character. Experience also provides examples where freely resisting temptation, particularly in difficult situations, seems to weaken the will over time rather than to strengthen it. I daresay that many men—and this would include some Christian ministers I know—have sincerely (even fervently) resisted sexual temptation for many years, only to succumb to it, finally, in middle age. For it may happen that the harder a man tries, for the most earnest of reasons, to suppress his childish yearnings and unrealistic fantasies, the more intense his temptations become and the more likely he is to succumb to them in an explosion of destructive behavior. Perhaps it would be misleading, however, to describe this as a case where good choices help to undermine a good character.

¹³ Nor should one make the mistake of supposing that the effect of alcohol on the alcoholic lasts only for the duration of the high. There are many long-term effects as well, including an ever-increasing craving for the drug and the gradual destruction of the very anti-anxiety centers of the brain upon which the alcohol works to provide temporary relief.

For if we suppose that the described behavior really is destructive and really is the product of childish yearnings and unrealistic fantasies, then it is also, perhaps, the product of deeper character flaws of which the agent is unaware—character flaws that first need to be exposed before they can be dealt with effectively.

Is my point, then, merely that the ultimate springs of human action are mysterious and incredibly complex, so that only a God could assess moral responsibility with any degree of accuracy? Not at all. My point is that we still have no coherent account of what it might even *mean* to say that free agents “make themselves into the kinds of persons they are”; at the very least, we need something more than a requirement for a life history to include some UAs. The relevant UAs must also qualify, in Kane’s own words, as “self-forming actions” (SFAs), and this in turn requires that an agent be personally responsible not only for the relevant UAs themselves, but also for the effect that the UAs have, in conjunction with a complex variety of other circumstances, on the agent’s character. But the problem is that actions too often have unexpected consequences in our lives. One person lies and cheats in pursuit of wealth and fame, only to discover that the result is emptiness and misery; and the circumstances surrounding this discovery may causally determine (even compel) a life transformation. Another may sincerely cultivate moral integrity and inadvertently produce some of the worst character traits: moral rigidity, self-righteousness, and a lack of compassion.

As Bernard Williams once observed, “One’s history as an agent is a web in which anything that is the product of the will is surrounded and held up and partly formed by things that are not” products of the will.¹⁴ Indeed, the assumption that even God could consider how people exercise their libertarian freedom and, on that basis, divide them into the good and the bad, or into those who deserve a reward and those who deserve punishment, now seems to me radically confused.

¹⁴ Bernard Williams, *Moral Luck* (Cambridge: Cambridge University Press, 1981), p.29.

As surprising as it may now seem, I remain an unrepentant libertarian myself, at least in this sense: I doubt that genuine freedom could exist in a fully deterministic universe. I also find most insightful Kane's view concerning the locus of indeterminism in what he calls a "self-forming willing" (SFW). It is as if an SFW, as Kane understands it, involves an undetermined leap of the imagination not altogether unlike the leap of imagination that might occur when a scientist formulates a new theory for testing. There are relevant reasons for the undetermined leap and also relevant reasons for a similar leap in some other direction, but there is no sufficient causal explanation of why the leap goes in the one direction rather than in the other. Kane himself puts it this way:

Every free choice (which is an SFW) is the initiation of a "value experiment" whose justification lies in the future and is not fully explained by the past. It says, in effect, "Let's try this. It is not required by my past, but it is consistent with my past and is one branching pathway my life could now meaningfully take. I'm willing to take responsibility for it one way or the other. . . . [In doing this, I am] guided by my past, but not determined by it."¹⁵

This seems to me most insightful, as I have said. We all emerge as self aware beings and begin making choices in a context of ambiguity, ignorance, and misperception; and furthermore, we all begin making choices in response to a host of complex (and in some cases wildly different) genetically determined inclinations and environmental (including social and cultural) influences. So Kane's idea that our decisions concerning the best course of action often amount to something like a "value experiment" makes a lot of sense to me. Elsewhere Kane goes on to elaborate: "To initiate and take responsibility for such value experiments whose justification lies

¹⁵ Kane, *op. cit.*, p. 145.

in the future, is to ‘take chances’ without prior guarantees of success. Genuine self-formation requires this sort of risk-taking and indeterminism is part of it.”¹⁶

But here one might want to ask: In what sense is one morally responsible for the outcome of a value experiment conducted in a context of ambiguity, ignorance, or misperception, a context in which there is, according to Kane, no certainty concerning the best thing to do¹⁷ and no “prior guarantee of success”? Suppose that I am trapped in a burning building with two apparent escape routes, and suppose further that, although only one of these apparent escape routes will indeed enable me to escape, I have reasons with respect to each of them for thinking that it might be the best route out. Here, then, is a situation that involves risk without any certainty concerning the best escape route or any guarantee of success. And though the presence of such risk may indeed add an element of drama to life, perhaps even something of great value, it may also seem incompatible with any personal responsibility for the outcome. So if risk taking, indeterminism, and a host of unexpected consequences are part of the process whereby we become the kind of person we are, as I agree they are, then we must also confront the question: In what sense are we morally responsible for the kind of person we finally come to be?

Now the remarkable thing is that the most virtuous among us rarely try to credit themselves for their own virtuous character. A loving mother, for example, will not credit herself for the love that controls her, however thankful she may be for the opportunity to care for (or even to sacrifice on behalf of) her children. But the virtuous also take responsibility for their moral failures. So here, very briefly, is how I view the matter. The concept of moral responsibility applies only to those whose character is compatible with moral failure. If the Christian God should ex-

¹⁶ Robert Kane, “Responsibility, Luck and Chance,” in Laura Waddell Ekstrom, *Agency and Responsibility* (Boulder, Colorado: Westview Press, 2001), p. 176. Originally published in *The Journal of Philosophy* XCVI, 5 (1999), pp. 217-240.

¹⁷ Kane writes: “If there are persons who need to be certain in advance just exactly what is the best or right thing to do in every circumstance . . . , then free will is not for them.” *Ibid.*

ist, then the concept of moral responsibility would have no application to him because moral failure would not be possible for him. And similarly for the perfected saints in heaven, as Christians have traditionally conceived of them: the concept of moral responsibility no longer applies to them either. They are not beneath moral responsibility in the way that an utterly irrational individual might be; instead, they are beyond moral responsibility in the sense that moral obligation no longer has any relevance to them. Paradoxically, however, the very conditions that make moral failure possible—the ambiguity, the ignorance, and even the indeterminism—may also seem to threaten responsibility, as our example of the man in the burning building illustrates. But a value experiment of the kind that Kane describes is also different from that example in one important respect: A relevant “value experiment” requires a context in which hypotheses concerning the best course of action, or the best way to live, can be put to the test, so to speak—that is, a context in which one can learn from mistakes and correct moral failures. In the case of a bad character in particular, being morally responsible for it depends not on its genesis, but on an agent’s present ability to learn moral lessons and thus to do something about it. On that point, I believe, the standard compatibilist understanding of moral responsibility has always been closer to the truth than the standard libertarian understanding.

I believe, in any event, that no adequate theory of moral responsibility is possible within the context of the traditional libertarian understanding of intrinsic desert: the unintelligible idea, as I see it, that certain punishments (or certain rewards, as the case may be) are *intrinsically fitting* responses to certain actions. But that is a much longer story and requires a much longer paper.