

Libertarian Free Will and the Erosion Argument *

Gerald Harrison
University of Aberdeen

Abstract. Libertarians make indeterminism a requirement of free will. But many argue that indeterminism is destructive of free will because it reduces an agent's control. This paper argues that such concerns are misguided. Indeterminism, at least as it is located by plausible Libertarian views, poses no threat to an agent's control, nor does it pose any other kind of threat.

I

Libertarians hold that indeterminism is a requirement of free will. A common criticism is that indeterminism would erode an agent's control and so damage or annul free will. The argument typically goes as follows. If an agent's decision is nondeterministically caused by his prior deliberative process if we roll back the clock to the moment just prior to decision-making and run the sequence through again an innumerable number of times, in some replays they will decide one way, and in some another. Because the agent's prior desires, beliefs, and all other motivational states are being kept constant it seems to be a matter of chance or luck that the agent decided one way rather than another. Indeterminism damages free will, it simply renders decisions chancy rather than determined. This criticism has been made in one form or another by amongst others: Bernstein (1995), Double (1991, pp. 198–99), Fischer (1999b), Haji (1999a; 2002), G. Strawson (2003), Waller (1988). Recently it has tended to be referred to as the 'Mind' argument (Van Inwagen, 2002). However, it is ambiguous whether the Mind argument claims that indeterminism reduces an agent's control compared to the level of control available in deterministic settings, or whether the Mind argument claims only that indeterminism fails to enhance control over the level available in deterministic settings. What I am interested in here, is whether indeterminism damages or erodes control. Whether indeterminism is capable of enhancing an agent's control

* I have discussed earlier versions of this paper with Julia Tanner and Geoffrey Scarre. Their comments, as well as those of an anonymous reviewer at this journal, have greatly improved it.

is a matter that is beyond the scope of this paper. For that reason I will refer to this variation of the Mind argument as the ‘erosion’ argument, as what unites all who make it is the belief that indeterminism erodes control.

In what follows I will argue that agent-internal indeterminism does not damage control, or at least, it does not damage control in any responsibility subverting way. I will also briefly consider three further dangers associated with indeterminism and will argue that they too are misguided.

II

The erosion argument is the claim that we have less control if our decision-making processes are indeterministic than if they are deterministic. As compatibilist control is by definition the only kind of control available in deterministic settings, it must be compatibilist control that indeterminism erodes. And if this supposed erosion in control is to matter, it must be significant enough to make a difference where moral responsibility is concerned. For this reason, when I talk of responsibility-relevant control I assume that I am addressing a compatibilist. However, the point of this paper is to show why indeterminism does not make matters worse than under determinism. I leave open the question of whether either libertarian or compatibilist conceptions of free will are correct. That is why I will assume throughout this paper that I am addressing a compatibilist—not because I believe compatibilism to be true, but because I want to show that in terms of control, libertarians are in no worse position. Whether the libertarians are in any *better* position than the compatibilists, is a matter that I leave open here. I actually believe that they are, but it is not my object here to make that case.¹

III

There are different kinds of libertarian who make different kinds of claims for free will. Some libertarians think that free will requires only indeterminism suitably located within an agent’s deliberative process whilst other libertarians think that something more is required—an ‘extra factor’ (such as a Kantian noumenal self or a special kind of agent-causation)—to deliver the kind of control required by free will and moral responsibility.

¹ I make this case in various ways in my 2005a, 2005b and 2006.

The type of libertarian who tries to do without recourse to exotic extra factors tends to be referred to as a ‘modest’ libertarian or sometimes an ‘event causal’ libertarian. In what follows I am going to focus exclusively on modest brands of libertarianism because if modest libertarian positions can be shown to introduce indeterminism in a way that does not erode control, then the more exotic brands of libertarianism can too.²

Modest incompatibilists insist that free will requires alternative possibilities and thus the falsity of determinism. However, modest incompatibilists maintain that indeterministic event causation, so long as it provides sufficiently robust alternative possibilities, is enough for free will. Robert Kane is the most prominent of these modest incompatibilists, though others who have outlined such views include Wiggins (2003), Sorabji (1980), Nozick (1981, pp. 294–316), Balaguer (1999; 2004), and Ekstrom (2000). For convenience I will focus on Kane’s view in the forthcoming discussion.

On Kane’s account our free will stems from occasions of motivational conflict, where an agent is torn between different courses of action—torn decisions in other words. Kane calls the decisions that resolve such internal conflicts Self Forming Acts or SFAs, and he suggests that there are six kinds (though he does not claim these are exhaustive):

[Self Forming Acts] include acts of the following kinds: (1) Moral choices or decisions, (2) prudential choices or decisions, (3) efforts of will sustaining purposes, (4) attentional efforts directed at self control and self modification, (5) practical judgements and choices, and (6) chances of intention in action. (1996, p. 125)

(1) and (2) involve “conflicts between what an agent believes ought to be done and what the agent wants or desires to do” (Kane, 1996, p. 126). In the case of prudential conflict this involves being torn between doing what is in one’s long-term interests, and what is in one’s immediate interests. In the moral case the conflict is between what one takes to be morally required, and what one wants to do. One has a practical conflict where one is torn between options neither of which has moral or prudential reasons in conflict (so Buridan’s ass cases would be cases of practical conflict as would be conflicts between values).

² It is possible of course that an extra-factor libertarian may well agree that indeterminism, left unsupplemented, erodes control. That, after all, may be the rationale behind bringing in the extra factor. To my knowledge no extra-factor libertarian actually gives this rationale, for typically the extra-factor libertarian believes compatibilist control inadequate for free will and that a wholly different kind of control is needed. It is supplying this different kind of control that the extra factor is brought in for, rather than in some kind of restorative capacity. Nevertheless, this kind of ‘restorative’ view is conceptually possible, and I defend this kind of rationale in my 2005b.

On the occasion of a torn decision the agent recognises reasons for two options, and is torn about which way to go, neither set of reasons seem to outweigh the others. The agent's prior character and motives provide both the reasons why the agent is trying to do, say, their duty, *and* the reasons why the agent is trying to do otherwise. As Kane puts it, the agent's "complex of past motives and character [...] explain the *conflict* within the agent's will *from both sides*" (1996, p. 127). Ultimately the agent ultimately just chooses. It is this kind of case that Kane asks us to imagine could be indeterministic so that he "might choose either way, all past circumstances remaining the same up to the moment of choice." (1996, p. 127)³

In recent years the erosion argument has been directed for the most part at Kane's brand of modest libertarianism. But, as I will show, indeterminism located between the agent's reasons and decision does not damage an agent's control in any responsibility-subverting way, so long as attention is paid to comparing relevant cases.

IV

The compatibilist Ishtiyaque Haji, is prominent amongst those who have made the erosion argument against Kane (1999; 2002, sec. 2). Haji holds that where indeterminism is just prior to decision making, what the agent lacks is antecedent proximal control, which is control "*to see to it that, in that rerun, he smokes rather than that he does not and vice versa*" (2002, p. 110).

With fixed pasts, the difference in outcome in Jones's and Jones*'s cases [these are the Jones and Jones* of the reruns in Haji's version of the luck objection] appears to be merely a function of the indeterminacy in the actional pathways leading to choice. But it would seem that no agent could exert proximal (or any other sort of) control over such indeterminacy to ensure a particular outcome. (Haji, 2002, pp. 110–11)

According to Haji, the relevant, destructive difference between an agent's torn decision being determined to be resolved in a certain way,

³ There is no need to conceive of torn decisions as only concerning important issues such as whether to act morally or self-interestedly. As Balaguer says, "we make [torn decisions] all the time, every day of our lives."

To appreciate this, consider a second case, involving a decision more mundane than Ralph's. Jane is in a restaurant, deliberating about whether to order tiramisu or a fruit plate for desert. She thinks that the former will taste better but that the latter will be better for her health. She has no clue which reason is stronger and feels genuinely torn. Suddenly, it's her turn to order; the waiter is looking at her; she has to pick; Oh, God, "I'll have the tiramisu," she says. i.e. she just chose. Period. (Balaguer 2004, p. 363)

and it being indeterministic which way it will be resolved, is that in the latter case the agent cannot antecedently ensure the outcome. This is an ability that Haji thinks the agent would have in a deterministic setting.

In response, adopt for the sake of argument, a straightforward compatibilist position of the type Haji and many other contemporary compatibilists endorse. Genuine alternative possibilities are not needed, and it is enough that in the actual sequence the agent exercised compatibilist control over the decision that they made, where this is a matter of the right connections being made (the agent decided for reasons and their response was an output of a sufficiently reason-responsive mechanism operating against a suitable background evaluative scheme). Now introduce agent external indeterminism into the picture in the following way. Due to certain properties of the agent's brain it is indeterministic whether or not the agent will suffer a massive stroke at t_1 . There is now a genuine possibility that, rather than make the decision in question at t_1 , the agent will die of a stroke at t_1 .

In one respect this clearly diminishes the agent's antecedent determining control. The agent cannot, prior to t_1 , ensure that they make a particular decision at t_1 , for they cannot ensure they do not suffer the stroke. If we replay the universe innumerable times, sometimes the agent will make a decision at t_1 , sometimes they will be dead at t_1 . Yet it seems quite obvious that if, in the actual sequence, the agent makes the decision in question, then they are as morally responsible as if there had been no possibility of anything else happening. After all, in the actual sequence in which the agent makes the decision, all the right connections were made and so the compatibilist control requirements are met. Whilst the external indeterminism could be said to serve as an obstacle to the agent successfully making a decision at t_1 , it does not serve as a responsibility-subverting obstacle. No-one could plausibly see the genuine possibility of the agent having a stroke at t_1 as working to get the agent off the hook for the decision that they actually made.

Now we can note that if the indeterminism is internal to the agent's decision making processes rather than external, that does not make any difference in terms of the reliability with which the agent can ensure a particular decision is made. In other words, if it is indeterministic whether Jones will make decision X or be dead at t_1 , or indeterministic whether Jones will make decision X, or decision Y, Jones's ability to reliably ensure that he makes decision X is equally damaged in both cases. Jones can no more antecedently ensure that he makes decision X at t_1 when the source of the control-diminishing indeterminism is *external* to his will, than he can when it is coming from within his own will. Thus, if it is the degree of reliability with which an agent

can ensure they make a particular decision Haji takes to be eroded in a way that is damaging to free will, his case fails. For in the cases of external indeterminism it is quite clear that the reduction in reliability does not damage the agent's free will in any responsibility subverting way. So too then in cases of internal indeterminism.

It might be objected that there is a relevant difference when the indeterminism is located between the agent's reasons and decision. For if the indeterminism is external to the agent's will, then although the agent cannot strictly speaking ensure they make (say) decision X at t_1 (they might be dead at t_1), they can nevertheless ensure that *if* they make a decision at t_1 , it will be decision X.

However, the erosion argument gets a lot of its superficial credibility from not comparing relevantly similar cases. What needs to be borne in mind is that on the Kanean view it is only our torn decisions that are indeterministically resolved, for it is only on the occasion of a torn decision that indeterminism is stirred-up in the will.⁴ Assuming that there is absolutely no reason to think that we do not face torn decisions if determinism is true, the cases we need to compare are torn decisions where the only difference is whether they are resolved indeterministically or deterministically.

Take a very crude model of compatibilist control (my argument would apply equally to a more sophisticated version). On this model an agent's decision is appropriately controlled if he would have decided otherwise had he wanted to. It does not matter whether the agent actually could have wanted to decide otherwise in the actual circumstances in which they make the decision, for all we are doing is assessing the responsiveness of the mechanism leading to the agent's decision. Note that in the case of a torn decision we encounter problems. In the case of a torn decision, the agent wants to make two incompatible decisions, but can only actually make one. This model of control cannot say that the agent would have decided otherwise if she had wanted to, because by hypothesis she did want to decide otherwise *in the actual sequence*. Left un-amended this account of control would deliver the curious result that an agent does not control their torn decisions even in deterministic settings.

Thus, unless the compatibilist wants to be in the embarrassing situation of having to say that we are not responsible for our torn decisions, (irrespective of whether they are resolved deterministically or indeterministically) they are going to have to endorse a notion of

⁴ Kane can allow that indeterminism is found elsewhere, however when indeterminism is found in other locations then it is an obstacle to free will (a point that Kane makes in his 1999). It is only when it is internal to the will, as on the occasion of torn decisions, that it positively contributes freedom.

‘moderate’ responsiveness generous enough to include torn decisions. In other words, the compatibilist will have to say that an agent’s decision issues from a suitably responsive mechanism provided there are a sufficiently broad range of circumstances in which the agent would have made a different decision if there was reason to, without committing themselves to the far stronger claim that the agent would always decide otherwise if there was reason to. But note that if the compatibilist makes their reason-responsive conditions moderate enough to cover torn decisions, then indeterministically resolved torn decisions will satisfy such control conditions too.⁵ There is no way for the compatibilist to avoid this without just stipulating that a process or mechanism leading to decision needs to be deterministic. In other words, there does not seem to be any non question-begging way in which the compatibilists can show why indeterministically resolved torn-decisions fail to satisfy moderate reason-responsiveness requirements.

If this is right it seems that there is no reason to see indeterministic efforts of will as diminishing of control in any way that matters. One might want to say that torn decisions are less controlled than non-torn decisions, precisely because when we are torn we do not know which way we are going to decide until we actually *do* decide. But this would be a point about torn decisions, irrespective of whether they are resolved deterministically or indeterministically.

V

I have argued that we have no reason to think that we will have less responsibility-relevant control over our torn decisions if they are resolved indeterministically than if they are resolved deterministically. I now want to turn to one respect in which we might feel that our control is diminished.

I might like the way that I currently am—I identify with my values and so forth—and I might want to remain just as I am. The prospect of conflicts in my will arising—Kane’s ‘self forming actions’ or SFAs—where it will be indeterministic in which direction my future

⁵ The compatibilist might object that in cases of torn-decision making there is not one mechanism or process operating, but two. This interpretation is lent support by Kane’s own conceptualisation of torn decisions as involving parallel processing. But this will not help the compatibilists, for if we just focus on the mechanism that issues in the decision in the actual sequence, then this mechanism will report the same degree of responsiveness irrespective of whether, in the actual sequence, the *dual* process was one that contained indeterminacy. For the *other* process is now to be thought of as an *external* source of interference. And external indeterminism does not affect reason responsiveness.

development takes, will therefore be of concern to me. For these future SFAs on the horizon threaten to change me. These future SFAs seem to pose a threat to what I term ‘retentive’ control. I want to know that I will always act in certain ways in certain situations, I want to know that I can retain my current character, but SFAs deprive me of any guarantees and so deprive me of retentive control.

I think that this is a legitimate worry, but one that is not unique to modest libertarianism. There are some, perhaps many aspects of our characters that we want to ensure do not change, and if indeterministic torn decisions are a necessary part of what it is to be free, then free will entails that our characters are *unfixed* to some extent, and liable to alteration. But we should note two things. Firstly, whilst the prospect of having an unfixed nature can be a cause for distress, the idea that our nature is fixed is one we are likely to find just as unpalatable. An unfixed nature is, we might say, something we can’t live with, and can’t live without. Secondly, whilst determinism gives one the prospect of having a fixed nature, it certainly does not guarantee it (it is this mistake which people make when they wrongly identify determinism with having a fixed nature—as Dennett points out, determinism is the thesis that one has a fixed future, not a fixed nature). There is no reason to think that if determinism is true, that one’s virtuous nature will remain fixed. For determinism does not guarantee that we will not face torn decisions where we are conflicted about the kinds of people we want to be. Thus, the kind of retentive control that modest libertarianism limits, is not something that can be guaranteed in either deterministic or indeterministic settings. Therefore the concerns about retentive control (or lack of it) apply to both compatibilist *and* libertarian conceptions of free will.

We should also note that whilst retentive control may be something we (sometimes) want, it is of questionable relevance to the question of moral responsibility. Whilst it is plausible that our characters need to hold together over time to some degree if we are to be coherent targets of the attitudes associated with moral responsibility, the idea that we need to be able to ensure that our character remains fixed if we are to be responsible is implausible. For as of right now, we cannot, plausibly, ensure that our characters remain fixed. Yet we still consider ourselves morally responsible for our actions despite this fact.

VI

So far I have argued that there is no reason to think that indeterminism, at least as it is located by a Kanean modest incompatibilist, threatens

to subvert responsibility relevant control. I now want to consider three other concerns about indeterminism. One concerns a supposed threat to the availability of relevant types of explanation. The second is a threat to the rationality of the agent's decision. And finally, one concerns the issue of attributability.

Haji thinks that one reason indeterminism rules out moral responsibility is "intimately tied to lack of an explanation in terms of prior reasons of the difference in choices" that an agent makes in "re-runs" of torn decisions. He is not alone. Many others have couched the erosion concern in terms of the lack of a certain type of explanation of why the agent decided as they did, rather than otherwise (Ayer, 1954, pp. 3–20; Double, 1996). The explanation in question is a 'contrastive explanation,' which is to say an explanation of why one thing, rather than another, occurred.

However, contrastive explanations are not the only kind of explanation that there is. There are also 'plain' explanations. A plain explanation is an explanation in terms of the event's prior causes. If an event is indeterministically caused, it is still caused, and as such a plain explanation is still available. The plain explanation will simply cite the prior causes. Consider a television remote control. This particular remote control will respond in the normal way to any particular one of its buttons being pressed. However, if two buttons are pressed simultaneously indeterminism is stirred up in its internal workings and it becomes indeterministic to which channel it will change. In the actual sequence, two buttons are pressed at once—the channel 1 button and the channel 2 button. The channel is changed to channel 1. Now, in this case we cannot explain why the remote changed channel to channel 1 rather than channel 2, for prior conditions were consistent with either of these events occurring. A contrastive explanation is, therefore, ruled out (note, not all contrastive explanations are ruled out—we can still contrastively explain why the remote changed channels to channel 1 rather than channel 3 for instance).⁶ But importantly, we can explain why the channel was changed to channel 1. We can explain this by citing the fact that the channel 1 button had been pressed. It is only when we ask why the channel was changed to channel 1 *rather* than channel 2 that we lack an explanation. This will also be the case with indeterministically resolved torn decisions. We will be able to plainly explain why the agent decided as they did—for we will be able to cite the prior reasons that spoke in favour of that decision. What we will not be able to do is explain why the agent made one decision rather than

⁶ Clarke (1996) has argued that we can still provide adequate contrastive explanations in indeterministic settings.

another. So, Haji is correct—we cannot explain why Jones decides one way in some re-runs, and a different way in others.⁷ But the important point is that the lack of contrastive explanation does not rule out the availability of a *plain* explanation.

What we now need to ask is why the lack of a relevant contrastive explanation should be a problem in itself? It is insufficient to just insist on it. For if contrastive explanations are not available in any setting other than one in which an agent's decision is determined by their prior reasons, then this just begs the question against the libertarian.⁸ It must rather be that the lack of contrastive explanation is taken to indicate either a lack of control—in which case the arguments I have made in the preceding sections take over. Or it might be taken to indicate a lack of rationality or a problem with respect to 'attribution.'

So, let us now turn to the concern about rationality. There is a concern that with the indeterminism located between the agent's reasons and their decision, the decision made is insufficiently connected to the agent's reasons for action (see Double, 1996, pp. 69–76; see also 1991)

However, we've already noted above that for the agent's decision to have been indeterministic does not mean that the decision that was made will not have been antecedently caused. If an agent's decision making process is indeterministic up to the moment of decision, that does not mean that we cannot give a plain explanation of the agent's decision in terms of the reasons that caused it in the actual sequence. So there is no reason to think that we will not be able to link the agents' decision with their prior reasons.

Nevertheless, the critic might insist that unless the agent's decision is entailed by their reasons, then their decision is not sufficiently rational. Following Balaguer we can call such decisions *strongly* rational (2004, p. 395). This is as opposed to a *weakly* rational decision:

[A] decision is *strongly rational* (given the agent's reasons for choosing) if and only if the agent's (conscious) reasons for choosing entail that the option chosen is the best available option. Second, a decision is *weakly rational* (given the agent's reasons for choosing) if and only if it is consistent with the agent's (conscious) reasons for choosing. (notice

⁷ This is something Kane acknowledges, as does Sorabji—another event-causal incompatibilist (Kane, 1996, p. 145; Sorabji, 1980, p. 31; see also Fischer, 1999a, p. 103, n. 18)

⁸ Another equally question begging criticism, which I've chosen not to make part of the main text of this chapter, concerns a claim about choices themselves. For it might be argued by some that 'choices' cannot be identified with undetermined events (see for instance, Bernstein, 1995, p. 154). Undetermined events are just things that happen. As Kane points out though, this implies that "if anything is a choice or action, it is determined" (2002, p. 423). This just rules out incompatibilism by fiat.

that weak rationality is certainly a kind of rationality; in particular, a decision that's weakly rational is not *irrational* in the sense of going against the agent's (conscious) reasons for choosing). (Balaguer, 2004, p. 395)

In the case of a torn decision the decision will clearly fail to be strongly rational, because the decision will not have been uniquely picked out by the agent's reasons. However, the decision will be weakly rational, for it is going to be consistent with the agent's reason set. The question then becomes whether weak or strong rationality is required for responsibility grounding free will. For if only weak rationality is required for responsibility, then the indeterminism central to the robust modest libertarian view does not threaten responsibility. And as Balaguer has pointed out, strong rationality is clearly not a requirement of responsibility:

Suppose someone said the following... Jane did not have compelling reasons that uniquely picked out her choice of tiramisu, so she did not choose the tiramisu of her own free will, because her choice was not sufficiently rational... it's pretty clear that this remark would seem downright bizarre to just about everyone. In our ordinary discourse and thought, we simply don't think that because decisions like Jane's aren't strongly rational, they are not free. (2004, p. 396)

Thus, whilst agent-internal indeterminism does preclude strong rationality for torn decisions, strong rationality is not plausibly a requirement of responsibility. If weak rationality is required (and it is not clear that even *this* much rationality is needed), then the relevant torn decisions satisfy the weak rationality requirement.⁹

What about attributability? According to this type of concern, what agent-internal indeterminism does is prevent us from connecting what the agent decides with any fixed character. When an agent's will is indeterminate, there does not seem to be any stable character to attribute their actions to. Consider Haji's comments:

Reconsider, now, the Jones/Jones* [roll-back] case. Entertain once again, the thought experiment that God has a thousand times caused the world to revert to precisely its state at the moment just before Jones decides to smoke, and that on about half these occasions, Jones decides to smoke

⁹ In fact, even if I am wrong above, it does not seem to me that rationality matters that much. We do not typically excuse someone just because they have acted irrationally. Irrationality, in other words, is not normally taken to indicate an absence of responsibility grounding control. If someone is systematically irrational then we might be inclined to say that they fail to satisfy sanity conditions and for that reason are not morally responsible. But even here, this would be because the agent was incapable of acting rationally, so the problem seems to be one to do with determinism and the absence of relevant alternative possibilities.

and acts accordingly. Assume, again, that in each of the reruns, Jones was trying to do two competing tasks, and that whatever he ended up doing, he would have done voluntarily, intentionally, and rationally. Suppose, as I have proposed, responsibility gauges the moral worth of an agent with respect to some episode in her life—a person discloses what she stands for when she is morally responsible for some deed. Then, given *type identical pasts*, when Jones does one thing in half or so of the reruns but something else in the others, there is no saying what Jones stands for. (2002, pp. 118–19)

The conception of responsibility this criticism presupposes is one according to which unless the decision that the agent makes in the actual sequence is strongly connected (for which we can read deterministically caused) by the agent’s character, then the decision cannot be “deeply reflective of who we are” (Double, 1996, p. 76). The agent is only morally responsible if their decision “discloses in conduct one’s moral stance or commitment vis-à-vis a particular episode in one’s life” (Haji, 2002, p. 118). Haji claims that “[o]ne can’t... disclose what one morally stands for with respect to a particular action without its being the case that there is a contrastive explanation, in terms of prior reasons, of why the agent performed that action rather than some other” (2002, p. 143).

Haji’s conception of responsibility draws heavily on what Gary Watson famously referred to as responsibilities ‘attributability’ face. Watson uses a quote from Dewey to capture the notion he is after:

when any result has been foreseen and adopted as a foreseen, such result is the outcome not of any external circumstances, not of mere desires and impulses, but of the agent’s conception of his own end. Now because the result thus flows from the agent’s own conception of an end, he feels himself responsible for it. . . . The result is simply an expression of himself; a manifestation of what he would have himself to be. Responsibility is thus one aspect of the identity of character and conduct. *We are responsible for our conduct because that conduct is ourselves objectified in actions.* (Dewey quoted in Watson, 1996, p. 227)

Haji takes this to imply that unless the agent’s acts are strongly connected to their prior reasons, meaning *deterministically caused* by prior reasons, then the agent’s act fails to disclose what the agent stands for. As such the agent is not responsible.¹⁰ What is assumed is that we are ‘set,’ and our decisions indicate the kind of person—our character

¹⁰ Haji’s view here can be compared in relevant respects with that of Hume, for whom an agent can only be responsible if their acts are attributable or revealing of their character.

Actions are, by their very nature, temporary and perishing; and where they proceed not from some cause in the character and disposition of the person who performed them, they can neither redound to his honor, if good; nor infamy, if evil. [A] person

values etc.—that we hold. Here attributability is achieved because our decisions can be sourced to some concrete prior ‘us’ that is revealed or expressed in our actions.

I am not at all sure that ‘attributability’ in this strong sense is a requirement of free will and moral responsibility, but I will put my scepticism here to one side. Even if such attributability is needed we can ask the compatibilist exactly how determinism can plausibly help achieve the kind of tight connection between an agent’s action and their character in the kinds of cases we are talking about. We need to remember that the cases we are talking about—the cases where the agent has genuine alternative possibilities—are ones where the agent is torn between different courses of action. We need to compare the relevant cases: torn decisions that are indeterministically resolved with torn decisions that are deterministically resolved. Once we do this it strikes me as implausible that a deterministically resolved torn decision can be said to be more expressive of what the agent stands for than an indeterministically resolved one. In both cases the decision will reveal something about the agent, some thing which is equally enduring.

It is true that the indeterminism present in directly free acts means that the agent’s past character does not entail that they will make one decision or another. But the agent’s decision sets their will (at least until they make another torn decision) in which case we can say that the agent’s decision does indicate how they are—it reveals what they’ve just made themselves, rather than revealing what they were that led to them making the decision. As far as I’m concerned this achieves everything in terms of attributability available in deterministic settings. The agent’s indeterministically resolved torn decision reveals what the person stands for *now*. And as for durability—there is no more or less reason to think that the person’s character as it now stands is any less durable under indeterminism than under determinism. For as already pointed out above, it is torn decisions which pose the threat to durability, and we face torn decisions irrespective of whether determinism or indeterminism is true.

Whatever sense of attributability the compatibilist critic has in mind indeterminism either does not threaten it at all (once one focuses on the relevant cases—so torn decisions) or the kind of attributability wanted is equally jeopardised in a deterministic setting as it is in an indeterministic one.

is not answerable for [immoral acts] if they proceeded from nothing in him, that is durable and constant (Hume, 1955, p. 98).

VII

In this paper I have argued that agent-internal indeterminism does not erode control in a way that could plausibly make a difference to an agent's responsibility. Focussing on Kane's modest libertarian position I argued that when one compares deterministically resolved torn decisions with indeterministically resolved ones, there was no reason to see the indeterminism as eroding control in any relevant way.

I considered other ways agent-internal indeterminism has been seen as threatening, including jeopardising the availability of certain types of explanation; the rationality of the agent's choices; and the attributability of the agent's choices. The concerns were, in all these cases, misguided or question begging.

References

- Ayer, A. J. (1954). *Philosophical Essays*. New York: St Martin's Press.
- Balaguer M. (1999). Libertarianism as a Scientifically Reputable View. *Philosophical Studies*, 93, 189–211.
- Balaguer M. (2004). A Coherent, Naturalistic, and Plausible Formulation of Libertarian Free Will. *Nous*, 38, 379–406.
- Bernstein, M. (1989). Review of R. Kane, *Free Will and Values*. *Nous*, 23, 557–59.
- Bernstein, M. (1995). Kanean Libertarianism. *Southwest Philosophy Review*, 11, 151–57.
- Clarke, R. (2002). Libertarian Views: Critical Survey of Noncausal and Event-Causal Accounts of Free Agency. In Kane (Ed.), 2002.
- Dennett, D. (1978). *Brainstorms*. Cambridge, Mass: MIT Press.
- Double, R. (1996). *Metaphilosophy and Free Will*. Oxford: Oxford University Press.
- Double, R. (1991). *The Non-Reality of Free Will*. New York: Oxford University Press.
- Ekstrom, L. (2000). *Free Will: A Philosophical Study*. Boulder, Co.: Westview Press.
- Fischer, J. (1995). Libertarianism and Avoidability: A Reply to Widerker. *Faith and Philosophy*, 12, 119–25.
- Fischer, J. (1999). Alfred R. Mele, Autonomous Agents: From Self-Control to Autonomy. *Nous*, 33, 133–43.
- Haji, I. (1999). Moral Anchors and Control. *Canadian Journal of Philosophy*, 29, 175–203.
- Haji, I. (2002). *Deontic Morality and Control*. New York: Cambridge University Press.
- Harrison, G. (2005a). Moral Luck and the Case for Hyper-Libertarianism. *Sorites*, 16.
- Harrison, G. (2005b). Free Will and Luck. Ph.D. Dissertation, University of Durham.
- Harrison, G. (2006). The Case for Hyper-Libertarianism. *Kriterion*, 20, 1–6, forthcoming.
- Hume, D. (1955). *An Enquiry Concerning Human Understanding*. Indianapolis: Bobs-Merrill.
- Kane, R. (1988). Libertarianism and Rationality Revisited. *The Southern Journal of Philosophy*, 26, 441–60.

- Kane, R. (1996). *The Significance of Free Will*. New York: Oxford University Press.
- Kane, R. (1999). On Free Will, Responsibility, and Indeterminism: Responses to Clarke, Haji, and Mele. *Philosophical Explorations*, 2, 105–21.
- Kane, R. (2002). Some Neglected Pathways in the Free Will Labyrinth. In Kane (Ed.), 2002.
- Kane, R. (Ed.) (2002). *The Oxford Handbook of Free Will*. Oxford: Oxford University Press.
- Mele, A. (1995). *Autonomous Agents: From Self-Control to Autonomy*. Oxford: Oxford University Press.
- Nozick, R. (1981). *Philosophical Explanations*. Cambridge, Mss: Harvard University Press.
- Strawson, G. (2003). The Impossibility of oral Responsibility. In Watson (Ed.), 2003.
- Van Inwagen, P. (2002). Free Will Remains a Mystery. In Kane (Ed.), 2002.
- Waller, B. (1988). Free Will Gone Out of Control: A Review of R. Kane's *Free Will and Values*. *Behaviourism*, 16, 149–62.
- Watson, G. (1996). Two Faces of Responsibility. *Philosophical Topics*, 24, 227–48.
- Watson, G. (Ed.) (2003). *Free Will*. Oxford: Oxford University Press.
- Wiggins, D. (2003). "Towards a Reasonable Libertarianism." In Watson (Ed.), 2003.

