Canadian Journal of Philosophy

ISSN: 0045-5091 (Print) 1911-0820 (Online) Journal homepage: http://www.tandfonline.com/loi/rcjp20

Critical Notice

PAUL RUSSELL

To cite this article: PAUL RUSSELL (2002) Critical Notice, Canadian Journal of Philosophy, 32:4, 587-606, DOI: 10.1080/00455091.2002.10716532

To link to this article: https://doi.org/10.1080/00455091.2002.10716532

CANADIAN JOURNAL OF PHILOSOPHY
Volume 32, Number 4, December 2002, pp. 587-606

Critical Notice

JOHN MARTIN FISCHER and MARK RAVIZZA, *Responsibility and Control: A Theory of Moral Responsibility*. Cambridge: Cambridge University Press 1998. Pp. viii + 277.[1]

I

According to the central tenets of classical compatibilism, the only kind of control required for agents to be free and responsible is the ability to act according to the determination of their own desires and willings. Since this condition can be satisfied without denying the thesis of determinism, it is argued, we can dismiss the pessimistic worries of the incompatibilist as unfounded.[2] While this view of things dominated compatibilism for many generations, by the end of the twentieth century it was generally regarded as too crude and simple to serve as an adequate foundation for a plausible compatibilism. Recent compatibilist work on this subject has been shaped in large part by the related debate about whether or not moral responsibility requires that an agent 'could have done otherwise' or had 'alternative possibilities.' Among the more influential contributors to these debates have been Harry Frankfurt and

1 My thoughts on this book have benefited from discussions with Bob Bunn, David Zimmerman and other members of a free will seminar given at UBC during the Winter term of 2000. A visit and pair of lectures given at UBC by John M. Fischer the same term further helped this process along. I am also grateful to colleagues who attended a symposium on this subject at the Canadian Philosophical Association meeting at Quebec City in May 2001 for their friendly and helpful comments.

2 A prominent example of the classical position can be found in R.E. Hobart, 'Free Will as Involving Determinism and Inconceivable Without It,' in Bernard Berofsky, ed., *Free Will and Determinism* (New York: Harper & Row 1966), esp. 72-3.

Daniel Dennett. Both Frankfurt and Dennett have argued that we ought to sever the (assumed) link between alternate possibilities and the kind of control required for moral responsibility.[3] Their work on this subject provides an important point of departure for the significant and valuable project that Fischer and Ravizza pursue in *Responsibility and Control*.

In 'Alternative Possibilities and Moral Responsibility' Frankfurt develops several counter-examples to the principle that a person is responsible for what he has done only if he could have done otherwise. He describes various cases that aim to show that, given the actual sequence of events, the agent's exercise of control over his action is not impaired by the lack of alternative possibilities. Dennett endorses Frankfurt's position, but goes on to argue that he is 'insufficiently ambitious' on this issue. According to Dennett, the whole question of 'alternative possibilities' is one that is neither empirically answerable nor worth caring about.[4] The kind of control that we ought to care about, Dennett maintains, does not depend on alternative possibilities, but rather on the 'power to be moved by reasons' (ER, 18-19, 25, 50-1, 98) Fischer and Ravizza introduce their work by stating that their primary aim is to advance this general line of reasoning (RC, 26-7). That is to say, they aim to 'give a comprehensive account of the kind of control that grounds moral responsibility' (RC, 14) and to show, in particular, that the relevant kind of control does not presuppose alternative possibilities (RC, 51, 54).

II

Fischer and Ravizza ground their project on the assumption that for an agent to be morally responsible for his conduct he must be in control of it in some suitable sense. They argue, however, that there are two very different accounts of control that may be suggested. The first, which they call 'regulative control,' requires 'genuinely open alternative possibilities' (RC, 30). In contrast, 'guidance control' does not require open alternatives. To explain this contrast Fischer and Ravizza cite the example of a person (Sally) who is driving a car, and turns it to the right. In this actual sequence she manifests guidance control of the vehicle. In

---

3 Harry Frankfurt, 'Alternate Possibilities and Moral Responsibility,' *Journal of Philosophy* **66** (1969) 828-39; reprinted in John M. Fischer, ed., *Moral Responsibility* (Ithaca: Cornell University Press 1986); Daniel Dennett, *Elbow Room* (Oxford: Clarendon Press 1984)

4 Dennett, *Elbow Room*, Ch. 6. Hereafter *Elbow Room* is abbreviated as ER, and *Responsibility and Control* as RC.

most cases of this kind we assume that the driver also had the power to turn the car to the left, but did not exercise this power — this is 'regulative control.' Generally we suppose regulative and guidance control come together, but this need not be the case. Imagine, for example, that Sally's car has 'instructor's (dual) controls.' If Sally shows any sign of turning left, the instructor will intervene and turn the car to the right. However, in the actual circumstances, Sally turns the car to the right and the instructor does not intervene. This is a case, Fischer and Ravizza suggest, where a person exercises guidance control even though she lacks regulative control.

Where an agent enjoys regulative control she must be able to exercise guidance control in both the actual sequence and the alternative sequence (RC, 31). As the example shows, however, guidance control does not presuppose regulative control or alternate possibilities (RC, 32). With this distinction in place, Fischer and Ravizza proceed to argue that while moral responsibility does presuppose some form of control, the relevant kind of control is guidance control, not regulative control and alternate possibilities (RC, 33-4).

The importance of 'Frankfurt-type cases,' according to Fischer and Ravizza, is that they 'invite us to look more carefully at the characteristics of the actual sequence that leads to the action' (RC, 37; emphasis in original).

> In contrast to traditional views, an actual-sequence model of moral responsibility holds that ascriptions of responsibility do not depend on whether agents are free to pursue alternative courses of action (and thus have alternative scenarios genuinely accessible to them); rather, what is important is (roughly speaking) what the agents actually do, and how their actions come to be performed. (RC, 37)

In this way, the actual-sequence approach involves a 'switch from a focus on the relevant agents and their properties, to a focus on the processes or ''mechanisms'' that actually lead to the action' (RC, 38). On this account, the issue that ought to concern us, is whether or not 'the kind of mechanism that actually operates is reason-responsive' (RC, 38). Fischer and Ravizza say that they employ the terminology of 'the mechanism' simply as a way of referring to the (actual) process that leads to the action, or the way that the action came about.[5]

---

5 Fischer and Ravizza say that they 'are not committed to any sort of 'reification' of the mechanism; that is, we are not envisaging a mechanism as like a mechanical object of some sort. The mechanism leading to an action is, intuitively, the way the action comes about; and, clearly, actions can come about in importantly different ways' (RC, 38n.).

According to Fischer and Ravizza, there are 'two distinct dimensions' to guidance control. The first is that the agent must 'own' the mechanism that issues in action (RC, 39; cp. 241). The second is that the 'mechanism' must possess the relevant 'degree of reasons-responsiveness.' Most of *Responsibility and Control* is concerned with the requirement of reasons-responsiveness and how it should be interpreted and applied. The account presented turns on a contrast between 'strong' and 'weak' interpretations of reasons-responsiveness. In the case of strong reasons-responsiveness, where the agent has sufficient reason to do otherwise, the agent would recognize the 'sufficient reason to do otherwise and thus choose to do otherwise and do otherwise' (RC, 41). In other words, the actual mechanism that operates is one that (borrowing a notion from Robert Nozick) 'tracks value.' In these circumstances the agent's actions 'would be tightly aligned with reasons' (RC, 42). The difficulty here, however, is that strong reasons-responsiveness cannot plausibly be presented as a necessary condition for moral responsibility. If this were the case, then agents could never be responsible for actions that they had (moral) reason to refrain from doing.

The challenge for Fischer and Ravizza, therefore, is to provide a 'weaker' account of reasons-responsiveness that allows for 'a looser kind of fit between reasons and action' (RC, 43). They begin with a contrasting account of 'weak reasons-responsiveness.' To understand cases of this kind we must rely on the notion of 'holding fixed the actual kind of mechanism that operates' — as we do in the case of strong reasons-responsiveness. The intuitive idea is that we can ask how the same mechanism would operate in circumstances where the sufficient reasons that are presented to us are different. To assess this we must rely on counterfactual, possible world reasoning (RC, 44n.).[6] In the case of strong reasons-responsiveness, as we have noted, the same mechanism will continue to 'track' the available sufficient reasons, and guide action accordingly. In contrast with this, a 'weak' mechanism does not always track reason in this way. All that is required of 'weak' mechanisms is 'that there exists some possible scenarios (or possible world) in which there is a sufficient reason to do otherwise, the agent recognizes this

6 A further difficulty here concerns the notion of 'mechanism individuation'; cp. RC, 40. Fischer and Ravizza state that it is simply a presupposition of their theory 'that for each act, there is an intuitively natural mechanism that issues in action, for the purposes of assessing guidance control and moral responsibility' (RC, 47). They deny, however, that this commits them to 'a stringent view of ''same kind of mechanism'' (according to which sameness requires sameness down to microdetails)' (RC, 52n.).

reason, and the agent does otherwise' (RC, 44). Unlike cases of strong reasons-responsiveness, agents operating with a 'weak' mechanism may sometimes act against good reasons, but may still be responsive to some reason.[7]

In light of this analysis, it may be argued that only weak reasons-responsiveness is needed for guidance control, and thus for moral responsibility. Fischer and Ravizza, however, reject this suggestion. The problem is that the account suggested is too weak. To explain this they draw a distinction between two ways in which a mechanism may fail to be 'responsive to reasons.' The first concerns recognition of reasons, whereas the second concerns reactivity to reasons. An agent may well be able to recognize that reasons exist, but unable to 'translate' those reasons into choices that guide his subsequent behavior (RC, 69). That is to say, the agent must not only be aware of what reasons there are, but also capable of being *moved* by them. The strategy that Fischer and Ravizza adopt is to argue that while only 'a very weak sort of reactivity' is needed for the kind of responsiveness relevant to moral responsibility, a stronger sort of receptivity to reasons is required (RC, 69). They describe this account as a 'moderate reasons-responsive' view — one that occupies the 'middle-ground' between demands on responsiveness that are either too weak or too strong.

The distinctive feature of the moderate account is the 'asymmetrical' aspect of its commitments. The recognition requirement is understood as suggesting that the mechanism in question must at least be *regularly* receptive to reasons.

> In judging a mechanism's receptivity, we are not only concerned to see that a person acting on that mechanism recognizes a sufficient reason in one instance; we also want to see that the person exhibits an appropriate *pattern* of reasons-recognition. In other words, we want to know if (when acting on the actual mechanism) he recognizes how reasons fit together, sees why one reason is stronger than another, and understands how the acceptance of one reason as sufficient implies that a stronger reason must also be sufficient. (RC, 70-1; emphasis in original)

---

7 Contrast, for example, the case of a person who is acting on the basis of an 'irresistible desire' with someone who is simply 'weak-willed.' The compulsive will continue to act in the same way no matter what reasons are provided for refraining from the conduct. The weak-willed person, although not strongly reasons-responsive, will nevertheless refrain from the conduct when presented with some relevant set of sufficient reasons (RC, 43, 48).

According to the 'moderate' account, then, the actual mechanism that issues in action must be at least 'regularly' receptive to reasons, and this implies that the mechanism exhibits some relevant pattern of reasons-recognition. This still falls short of invariably recognizing reasons, but it demands more than simply recognizing sufficient reasons in some isolated instance.

In contrast, there is no similar demand that reasons-reactivity exhibit some relevant 'pattern' or 'regularity.' Fischer and Ravizza justify this asymmetry by arguing that reactivity, unlike recognition, is 'all of a piece.'

> That is, we believe that if an agent's mechanism reacts to some incentive to (say) do other than he actually does, this shows that the mechanism can react to any incentive to do otherwise. Our contention, then, is that a mechanism's reacting differently to a sufficient reason to do otherwise in some other possible world shows that the same kind of mechanism can react differently to the actual reason to do otherwise. (RC, 73)

The crucial claim here is 'that reactivity is all of a piece in the sense that the mechanism can react to all incentives, if it can react to one' (RC, 74). This implies, as Fischer and Ravizza point out, that 'there can be considerably more idiosyncrasy in the reactivity component of the mechanism that leads to action than in the receptivity component' (RC, 74).

Fischer and Ravizza maintain that the 'refinements' involved in distinguishing 'moderate' from 'weak' and 'strong' reasons-responsiveness allow us to make better sense of several 'difficult cases' that we encounter in moral life. The most interesting of these is the psychopath. One way of interpreting the psychopath is that he is not capable of recognizing moral reasons when they are presented to him — and for this reason the psychopath cannot be morally responsible. On another interpretation, however, the psychopath is 'appropriately receptive to [moral] reasons,' but systematically fails to be reactive to them. That is, this individual never 'translates' these reasons into choices. Nevertheless, so long as this individual has shown some reactivity to [non-moral] reasons, he remains sufficiently in control to be regarded as a morally responsible agent. Although moral reasons are inert with this individual, he has reacted successfully to available reasons in other cases, which shows that the actual mechanism involved *can* react to some relevant incentive to do otherwise (RC, 78-80).

In sum, the kind of control required for moral responsibility depends on dispositional or modal properties of moderate reasons-responsive mechanisms (RC, 53). Nothing about the thesis of determinism implies that agents do not possess mechanisms of this kind. The properties that we ought to be concerned with are those involved in the actual sequence. It does not matter, for the analysis of these properties, that the possible

worlds or alternative sequences that ground our judgments about them are not '*genuinely accessible* to the agent' (RC, 53; emphasis in original). On this account, therefore, moral responsibility depends only on guidance control, which does not require alternate possibilities or the freedom to do otherwise. Fischer and Ravizza describe the position that this analysis yields as 'semicompatibilism': even if causal determinism is incompatible with alternate possibilities, it is still compatible with moral responsibility (RC, 53).

III

Fischer and Ravizza extend their analysis of guidance control and moderate reasons-responsiveness to cover both consequences and omissions. The argument for this is given in considerable detail, and relies heavily on a series of (exotic) examples that illuminate the problems involved. I will not trace their argument through all these intricacies. Instead, I want to discuss a few objections that may be raised in criticism of their account of moderate reasons-responsiveness considered as the mode of control that is required for moral responsibility.

As I have explained, Fischer and Ravizza aim to find an account of reasons-responsiveness that occupies the middle-ground between the 'strong' and 'weak' accounts. Their asymmetrical 'moderate' account allows 'a very weak sort of reactivity,' but requires a stronger sort of receptivity, one that involves 'an understandable pattern of reasons-recognition, minimally grounded in reality' (RC, 73). The difficulty with this position is that it is not clear why we should require some 'pattern' or 'regularity' for receptivity, but reject the same demand for reactivity. Clearly Fischer and Ravizza hold that strengthening the reactivity requirement, in line with the receptivity requirement, would be too demanding — since we do not want to exclude agents who have a mechanism that is regularly receptive to reasons and has shown that it can react to them. The controversial assumption that this position rests upon is that 'reactivity is all of a piece in the sense that the mechanism can react to all incentives, if it can react to one' (RC, 73-4). It may be argued, however, that this same reasoning can be applied to the receptivity requirement, which would result in a return to a 'weak reasons-responsive' view. On the face of it, therefore, the asymmetry that Fischer and Ravizza introduce, in order to arrive at a 'moderate' position, seems to depend on ad hoc adjustments, rather than principle-driven considerations.

There are, I believe, more fundamental difficulties for a reasons-responsive view of the kind that Fischer and Ravizza seek to defend — whether it takes the form of strong, weak or moderate mechanisms. In

the first place, while agents may *possess* reasons-responsive mechanisms or dispositions, it is not clear that they have control over how capacities of this kind are actually exercised in particular conditions of action. In order to have control of this sort, critics will argue, we need 'regulative control' and alternate possibilities.[8] It may argued in reply that this sort of difficulty is not troubling in the case of strong reasons-responsive mechanisms. That is, in these cases the agent possesses mechanisms that cannot fail to recognize and react to the available reasons. An ability to 'disconnect' our deliberations and choices from the reasons we are presented with is not worth wanting.[9] On this view, since the agent is guided flawlessly by reason and enjoys perfect practical reason, she may be viewed as *perfectly free*. The conclusion reached, therefore, is that at least in the case of strong reasons-responsive mechanisms, regulative control has no value.

It is not obvious that this conclusion is entirely consistent with all our intuitions on this subject. For example, it is not evident that an agent who is naturally governed by (moral) reasons, and does what is required of him effortlessly, *deserves* moral praise. Moral praise should be reserved for those who must 'struggle' to be good and do the right thing. Consider a person who is strongly reasons-responsive and another individual who had to exercise regulative control to do what is right. According to the account defended by Fischer and Ravizza, it is morally irrelevant that one individual is 'programmed' to act in accordance with moral reason while the other had to exercise regulative control and make an effort. There is, nevertheless, a clear sense in which the first individual is just 'lucky' to be strongly reasons-responsive, whereas the other person was not just lucky, but exercised a form of (moral) control not available to the strongly reasons-responsive agent. Worries of this kind capture an important strand in (neo-Kantian) incompatibilist thinking.[10] Plausible lines of reply may, of course, be suggested in support of the contrary

---

8 This general problem is discussed in some detail in R. Jay Wallace, *Responsibility and the Moral Sentiments* (Cambridge, MA: Harvard University Press 1994), esp. 183-90, 199, 201-13. Wallace considers the 'bipartite' objection that 'what matters to responsibility is not only the possession of the requisite general ability [i.e. reasons-responsiveness], but also the *opportunity* to exercise that general ability in the particular circumstances of action' (208). I am not persuaded by Wallace's effort to deal with this objection, but the difficulty is taken up and addressed.

9 Cp. Dennett, *Elbow Room*, 66.

10 See, e.g., C.A. Campbell, 'Is ''Freewill'' A Pseudo-Problem?' in Berofsky, ed., *Free Will and Determinism*, esp. 131-5; and Robert Kane, *The Significance of Free Will* (New York & Oxford: Oxford University Press 1996), Ch. 8 and 9.

position. (For example, it may be argued that moral praise should not be reserved only for cases where the agent makes some effort to do what is right, or even that it is a moral failing to need to 'struggle' to be moral.) Nevertheless, Fischer and Ravizza need to say more about the set of worries associated with this issue as it relates to the (moral) significance of regulative control.

The difficulties that we have been considering relate to the case of strong reasons-responsive mechanisms, which do not allow for the possibility that the agent may fail to respond appropriately to available reasons. Moderate reasons-responsiveness, of the kind that Fischer and Ravizza defend, must accommodate the possibility of failures. Cases of this kind, however, present their own problems. When an agent who is moderately reasons-responsive fails to respond appropriately to available reasons, the relevant disposition is (somehow) not activated in the specific circumstances. It is argued, nevertheless, that since the agent *could* recognize the reasons, and (if he did) *could* have reacted to them, there is no basis for denying that the agent had the relevant kind of (guidance) control over his conduct. The obvious worry, however, is that since the agent lacks regulative control, and the alternate possibilities associated with it, what the agent could not control is how this capacity of guidance control was *actually exercised in the specific circumstances*. On the face of it, therefore, the agent is liable to blame and retribution, on the moderate reasons-responsive account, merely for *possessing* capacities that he is not able to exercise control over. In the cases where the agent 'succeeds' in tracking reason this may not disturb us; but in the case of 'failures' the agent appears to be simply *unlucky* enough to be moderately reasons-responsive and placed in circumstances where the mechanism fails to track the reasons that were present. This is not a substantial enough foundation on which to rest our negative moral sentiments, much less the retributive practices that come with them. For this reason, I believe, incompatibilists will continue to hold that moral responsibility requires regulative control. A convincing defence of guidance control needs to argue in detail, by way of the kinds of cases that have been mentioned, why control of this (further) kind is unnecessary. Without an argument of this kind, incompatibilists will find that the model advanced leaves a significant gap in our ordinary intuitions about the sort of control that is required for moral responsibility.

IV

Another problem that is closely tied in with these difficulties concerns how the agent *acquires* her reasons-responsive mechanisms. Clearly the character of these mechanisms will determine the kind of choices and

decisions that we actually make. The question arises, therefore, whether we control the process by which these mechanisms are acquired. Suppose, for example, that a reasons-responsive mechanism might itself be 'induced neurophysiologically.'[11] The further worry that arises from this kind of situation is that the agent could be, as Robert Kane puts it, controlled in a covert, non-constraining manner by another agent.[12] In these circumstances the agent is controlled by way of manipulating his will, but without compromising the reasons-responsive mechanism that his actions issue from. Where covert, non-constraining control of this kind is present, Kane observes, we generally conclude that 'the kind of freedom that we normally associate with autonomy (or controlling our own lives)' is compromised.[13] It follows that there must be more to the sort of control required for moral responsibility than agent's simply possessing a reasons-responsive mechanism.

Fischer and Ravizza agree that the process of 'implanting' (moderate) reasons-responsive mechanisms — and the scope for covert, non-constraining control that this permits — would undermine responsibility. The problem that they face is to explain why this is so without conceding that what troubles us in these circumstances is that the agent lacks the power to be the ultimate originator of his action, which would require regulative control and alternate possibilities. The strategy that they pursue is to argue that there is a second component of guidance control, that is 'conceptually distinct' from reasons-responsiveness and missing in (troubling) cases of implantation or manipulation of the will. The missing component is that the mechanism that issues in action must be 'the agent's own' (RC, 39, 89, 170, 227, 241).

According to Fischer and Ravizza, for a mechanism to be owned by an agent, she must 'take responsibility' for it, and this involves a process that is 'essentially historical' (RC, 200).[14] There are, they maintain, three conditions that are necessary and sufficient for taking responsibility. The

---

11 This example comes from David Zimmerman, 'Acts, Omissions and ''Semicompatibilism,''' *Philosophical Studies* **73** (1994) 209-23.

12 Kane, *The Significance of Free Will*, Ch. 5. See also Richard Double, 'Puppeteers, Hypnotists, and Neurosurgeons,' *Philosophical Studies* **56** (1989) 163-73; and Ish Haji, *Moral Appraisability* (Oxford: Oxford University Press 1998), Ch. 6.

13 Kane, *The Significance of Free Will*, 67

14 The historical view of moral responsibility is analogous, Fischer and Ravizza claim (RC, 178-82), to the historical account of justice that is defended by Robert Nozick in *Anarchy, State and Utopia* (New York: Basic Books 1974). I discuss this analogy further below.

process begins with a child's moral education, as she comes to see herself 'as an agent' (RC, 208, 210-11, 238). What the child comes to see, on this account, is that certain upshots in the world are a result of her choices and actions. When this condition is satisfied, the child is then in a position to see herself as 'a fair target for the reactive attitudes as a result of how [she] exercises this agency in certain contexts' (RC, 211).[15] Finally, 'the cluster of beliefs specified by the first two conditions must be based, in an appropriate way, on the individual's evidence' (RC, 238). The third condition is important because it is in place to ensure that the first two conditions are not themselves satisfied by way of some 'inappropriate' process of manipulation or implantation. One especially important feature of this process of 'taking responsibility' is that it leads Fischer and Ravizza to endorse a 'subjectivist' approach to moral responsibility — one that requires 'an agent have a certain kind of view of himself, in order to be morally responsible for his behavior' (RC, 229). In this case, what is required is that the responsible agent must 'see himself as an agent who is an appropriate candidate for the reactive attitudes' (RC, 220-3, 229).[16]

The historical account of moral responsibility that Fischer and Ravizza defend, which they explain in terms of the process of 'taking responsibility,' should be contrasted with non-historical or 'current time-slice' accounts. The most prominent examples of this, they suggest, are the 'mesh theories' of Harry Frankfurt and Gary Watson.[17] According to Watson's theory, for example, the actions of a free agent exhibit a 'mesh' (i.e. absence of conflict) between what the agent values and what she most desires, or what she wills. The agent 'is responsible for an action insofar as there is a mesh between the valuational and motivational preference to perform the action' (RC, 185). A theory of this kind is

---

15 The reference to the role of reactive attitudes in this context draws on P.F. Strawson's 'landmark' essay 'Freedom and Resentment,' *Proceedings of the British Academy* **48** (1962), 187-211; reprinted in John M. Fischer and Mark Ravizza, eds., *Perspectives on Moral Responsibility* (Ithaca: Cornell University Press 1993). Fischer and Ravizza explain the Strawsonian view as holding that 'being morally responsible is being an *appropriate candidate* for the reactive attitudes' (RC, 7).

16 The 'subjectivist' approach to moral responsibility, as Fischer and Ravizza point out, is defended at length by Galen Strawson, *Freedom and Belief* (Oxford: Clarendon Press 1986), esp. Ch. 13 and 15.

17 Harry Frankfurt, 'Freedom of the Will and the Concept of a Person,' *Journal of Philosophy* **68** (1971) 5-20; reprinted in John M. Fischer, ed., *Moral Responsibility*. Gary Watson, 'Free Agency,' *Journal of Philosophy* **72** (1975) 205-20; reprinted in John M. Fischer, ed., *Moral Responsibility*.

'purely structural.' It does not concern itself with how the mesh came about or was produced. All that it looks for is 'a particular pattern in the relevant snapshot properties' (RC, 185). One obvious difficulty here is that the mesh may be produced by some other agent, through a process such as hypnotism or 'direct stimulation of the brain' (RC, 187, 196). Cases of this kind show, according to Fischer and Ravizza, that it is 'possible for there to be two agents who have all the same current time-slice properties but who differ as to their moral responsibility.' The difference in responsibility, they argue, 'comes from differences in the agents' histories' (RC, 202). More specifically, two agents could have the same reasons-responsive mechanisms (i.e. with the same dispositional properties), but nevertheless differ as to their moral responsibility for an action because their histories are different.

These claims about the relevance of history to moral responsibility are applied directly to the problematic cases of implanted reasons-responsiveness. In cases of this kind, Fischer and Ravizza argue, the agent has not gone through any relevant historical process of 'taking responsibility,' so the ownership condition of guidance control is not satisfied (RC, 234). It is possible, however, that there may be circumstances 'in which the agent's taking responsibility itself is somehow electronically implanted' (RC, 235). That is, the agent has gone through the first and second stages of 'taking responsibility' — and so sees himself as an agent and an appropriate target of reactive attitudes — but the whole process has been manipulated by another agent (e.g. a 'demonic neurophysiologist'). How do we handle cases of this kind?

Fischer and Ravizza appeal to the third condition of 'taking responsibility' in order to show that the agent in these circumstances is not morally responsible. Although the first and second conditions are satisfied, the agent 'has not formed his view of himself in the appropriate way' (RC, 236). At this critical juncture, however, the argument in defence of their position seems to run out of line. That is, since 'the relevant notion of appropriateness must remain unanalyzed' (RC, 236), the reader is asked to accept that there is an intuitive distinction between appropriate and inappropriate ways that an agent comes to see himself as an agent and a fair target of reactive attitudes. On one side of this boundary, we are asked to 'rule out the direct electronic implantation of the relevant cluster of beliefs [i.e. involved in satisfying the first two conditions]'. On the other side, however, we are not to reject this process when it is a result of causal determinism. The operating assumption appears to be that, so long as the process is not artificially manipulated or controlled by another agent, the fact that it is causally determined does not discredit the mechanism's credentials as 'owned by the agent.'

Fischer and Ravizza want to show that their account of guidance control cannot be assimilated to non-historical reasons-responsive views

that are unable to provide a role for the process of 'taking responsibility,' and that would otherwise be vulnerable to worries relating to implantation cases and covert, non-constraining control by other agents. The way that the non-historical theories have dealt with worries about implantation cases is to suggest either that what actually worries us in these circumstances is that the reasons-responsive mechanism is (somehow) compromised, or that suitably sensitive reasons-responsive mechanisms must be able to detect and avoid the kind of manipulation and control that seems objectionable.[18] According to Fischer and Ravizza's account, however, what ought to worry us is whether the agent *owns* the mechanism, which is a different matter. Their ownership condition serves to exclude agents who operate with reasons-responsive mechanisms that are implanted or manipulated. This is all that a plausible compatibilism needs, they maintain, to satisfy its critics.

V

Fischer and Ravizza state in several contexts that their aim is not to provide 'a knockdown argument' for the compatibility of causal determinism and moral responsibility in general, or the compatibility of causal determinism and the process of taking responsibility in particular (RC, 11, 228, 236, 251). Their aim, in the spirit of Robert Nozick's philosophical methodology, is to provide 'philosophical explanations' rather than pursue 'coercive philosophy' (RC, 11). What they aim to provide, therefore, is 'a strong plausibility argument' for their conclusions, but not 'an argument that any rational agent is compelled to accept' (RC, 11). Although this standard is weaker than some compatibilists have aimed at, it is still clear that there are two audiences who need to be 'persuaded' if this project is going to be more than an exercise of preaching to the converted. The doubters in question are, first, those with incompatibilist leanings, and, second, compatibilists who take a non-historical approach to responsibility.

Let us begin with incompatibilist doubts. I have already indicated that incompatibilists will generally be troubled by the fact that agents who operate with reasons-responsive mechanisms do not possess the kind of control that enables them to determine *how* these mechanisms operate in particular circumstances. Without regulative control, they maintain, it is not up to the agent whether the mechanism actually succeeds or fails to

---

18 Wallace, *Responsibility and the Moral Sentiments*, 197-8; Dennett, *Elbow Room*, 33-8

'track reasons.' I believe that incompatibilists will have similar worries about Fischer and Ravizza's account of the process of 'taking responsibility' as it is supposed to explain the agent's 'ownership' of mechanisms.

Considerations about 'ownership,' and the analogy of Nozick's historical entitlement conception of justice, show us where some of these problems lie. Nozick argues that individuals come to acquire property (i.e., own things) by way of two basic historical processes: either they acquire things by 'appropriating' them from things previously unheld, or they acquire something because it is transferred to them from someone else who is entitled to it (i.e., already owns it).[19] According to Nozick there is an important, although complex relationship between the process of coming to own something and the voluntary (free) activity of individuals. In some cases people may come to own things because they have control over the relevant process. For example, a person may acquire an item by mixing their own labor with something that was previously unheld; or a person may consent to an exchange of goods and acquire something through this process. In these cases, the individual concerned controls what it is that she comes to own. This is not, however, always the case. An individual may be given or inherit some item from another person. In these circumstances you may come to own something without choosing it or even accepting it. Indeed, you may not even want it. Ownership may involve, therefore, having something *imposed* upon you, independent from or even against your own will.

In the case of most items that we acquire against our own will, we can discard them as we choose — we do not have to continue to live with them. One obvious exception to this, however, is ownership of our own bodies.[20] In so far as I own my body (e.g. my brain) it is not something that I have consented to or selected, nor am I in a position to exchange it or dispose of it as I choose. The most that I can do is *acknowledge* that this is 'my own body.' It is evident, nevertheless, that acknowledging ownership of my body does not imply that I (somehow) consented to or chose it. Ownership, in this case, does not imply control over the process of acquisition.

These reflections on the historical entitlement theory of justice show that acquiring something may involve processes that the individual does not control, even though it results in ownership. This is a consideration

---

19 Nozick, *Anarchy, State and Utopia*, Ch. 7, sect. 1

20 G.A. Cohen observes that the foundation of Nozick's (Lockean) theory of justice is the 'thought that each person is the morally rightful owner of himself' (Cohen, 'Self-Ownership, World-Ownership, and Equality,' in Frank S. Lucash, ed., *Justice and Equality Here and Now* [Ithaca: Cornell University Press 1986], 109).

that is directly relevant to the process of coming to 'own' our reasons-responsive mechanisms. The question that arises for this theory is whether coming to own these mechanisms can plausibly be construed as a process that the agent controls, so that he can be said to have consented to or voluntarily acquired these mechanisms. That is to say, do agents stand to their reasons-responsive mechanisms in the way that they stand to their own bodies or in the way that they stand to external goods that they purchased or worked for?

This is a crucial question for the position that Fischer and Ravizza defend because, according to them, it is a basic presupposition of our thinking about moral responsibility that it requires some relevant form of control (RC, 13-14). Clearly, then, if we are able to speak of an agent 'taking responsibility' for his (own) mechanism, this implies some scope for control over the process of mechanism acquisition. In fact, according to Fischer and Ravizza, responsibility for actions that issue from a mechanism depends upon 'transmitting' responsibility to the behavior from the mechanism. If the agent does not control the process of mechanism acquisition in some relevant manner, then he cannot 'take responsibility' for it — whether he 'owns' it or not. If he is not responsible for the mechanism because he does not control how it was acquired, then there is no possibility of 'transmitting' responsibility from the (owned) mechanism to the action that issues from it.

Fischer and Ravizza ought to have something substantial to say about the extent to which the process of mechanism acquisition that leads to 'taking responsibility' involves control, but they have little to say about this. This is, I believe, no minor problem for their theory. In the first place, while we can make good sense of having control of our actions on the basis of possessing reasons-responsive mechanisms, it is not at all obvious what it *means* to say that an agent controls the process of acquiring such mechanisms. As Fischer and Ravizza point out, this process begins in the earliest stages of moral development, when a person is still an infant (RC, 208-10). At this stage the child does not have any (fully developed) reasons-responsive mechanism in place. The relevant mechanisms evolve and develop over time, and are continually conditioned through the influence of moral education as provided by parents and other adults. It is these individuals, therefore, rather than the child, who are able (in some degree) to control the process of mechanism acquisition and development. For this reason responsibility for the kinds of mechanisms children acquire and develop rests more plausibly on the shoulders of the adults who have raised the child. It is only at a much later stage of moral development, when a reasons-responsive mechanism of some kind is already in place, that children or adolescents are able to think critically about the way that their own deliberative capacities actually operate. Even at this stage, however, there is little or no

question of the agent being able to reform or remodel the mechanisms that he is (already) operating with at will. Control of this kind is not available even to mature adults, much less younger children.

Let us suppose, nevertheless, that some reasonable sense can be made of the suggestion that the (mature) agent has control over mechanism acquisition. This form of control must, on Fischer and Ravizza's account, be understood in terms of the agent's ability to deliberate and decide about these mechanisms by way of exercising the reasons-responsive mechanisms that they already have. The incompatibilist will argue, however, that this presents the theory with a serious regress problem. Any decision to accept or reject a given mechanism must be based on a mechanism that is itself either chosen or given. At some point in this process, the mechanism involved in the process of mechanism acquisition (i.e. selecting or choosing a mechanism) must have been unchosen or presented to the agent through (natural) causes that he did not control. Reflections of this kind erode any confidence that we have that agents are able to control the process of mechanism acquisition, since the particular *trajectory* that the process takes is structured by a mechanism that the agent has been 'given,' not one that he has 'chosen.'[21] Any choice concerning mechanism acquisition must eventually depend on unchosen mechanisms — even on the optimistic assumption that mature agents are able to make choices of this kind.

In the case of implantation of reasons-responsive mechanism what troubles Fischer and Ravizza is that some other agent controls the nature and character of the mechanism that an agent is operating with. This situation, they argue, compromises the agent's 'ownership' of the (implanted) mechanism, and so rules out moral responsibility. This is not the same worry that the incompatibilist has about this situation. What troubles the incompatibilist is that the agent has no relevant form of *control* over what mechanism it is that he comes to 'own' or operate with. It is immaterial, from this point of view, whether the agent acquires or is 'given' the mechanism through the active intervention of another agent or by way of natural, impersonal processes that he does not control.

---

21 This is a variant, I believe, of a familiar set of problems in Aristotle's ethical theory. Aristotle is committed to the view that moral agents are responsible for the character traits that they acquire, because there is a distinction to be drawn between (early) conduct that *shapes* character and (mature) conduct that *expresses* it. Critics point out, however, that a sharp distinction of this kind is hard to defend. Beyond this, during the early stages of moral development, when character is being shaped, the agent's choices and deliberations are conditioned by factors that they do not control. See, e.g., W.F.R. Hardie, *Aristotle's Ethical Theory*, 2nd ed. (Oxford: Clarendon Press 1980), 175.

In neither of these circumstances can an agent be held responsible for what he does not effectively control.

While agents may come to 'own' their reasons-responsive mechanism (i.e. in some 'appropriate' manner), the relationship is more like that which a person has to his body. Whether our body is 'given' to us by some other agent (e.g. God) or by natural causes, we have little or no control over its basic properties and qualities (height, color, etc.). Although it is ours, and as mature adults we recognize that responses to its qualities and features may be fair or unfair, we are not in a position to 'take responsibility' for something that we do not control. It follows from this that if responsibility for action depends on a prior responsibility for the mechanism that action issues from, no agent is morally responsible for their conduct.[22]

VI

Now let us consider how compatibilists who take non-historical approach to responsibility might respond to Fischer and Ravizza's account of guidance control. The obvious point of debate is the claim that there could be 'two agents who have all the same current time-slice properties but who differ as to their moral responsibility' (RC, 202). Fischer and Ravizza support this claim, as we have noted, by pointing out that two agents who share the same reasons-responsive mechanism may nevertheless differ in that one of them is being (covertly) controlled by another agent by means of implants or electronic manipulation. They deny that an agent who is being covertly controlled in this way can be morally

---

22 It is worth noting that mechanisms that provide agents with the power of 'regulative control' do not generate worries about implantation of the kind that the reasons-responsive mechanisms involved with guidance control encounter. Let us suppose that it is possible to implant a mechanism that supports regulative control in an agent (call it a 'regulative control mechanism,' or RCM). The fact that a RCM is implanted will not trouble the incompatibilist, because implantation as such will not compromise the agent's ability to exercise regulative control. Moreover, although a mechanism has been implanted, this does not make it possible to control the agent by means of this process. On the contrary, since the agent retains regulative control, the source or historical origins of the RCM is irrelevant to the way that the agent actually exercises this mechanism in the specific circumstances. What this shows is that incompatibilist worries about the implantation of reasons-responsive mechanisms are reducible to worries about the importance of being able to control how the mechanism is actually being exercised. If these concerns are satisfied — as they are in the case of RCM — incompatibiists could set aside the 'history' of mechanism acquisition as irrelevant to their concerns.

responsible — which is why the condition of reasons-responsiveness does not suffice for guidance control on their account. This attitude to the presence of covert controllers is what Kane describes as 'soft compatibilism.'[23] 'Hard compatibilists,' however, may argue that, so long as reasons-responsiveness is not compromised by covert control or the implantation process, there is no relevant basis for distinguishing these individuals.

Consider, for example, two individuals who have the same strong reasons-responsive mechanism. The only difference between them is their histories are different: one mechanism has been implanted, whereas the other has been acquired through an 'appropriate' process of 'taking responsibility.' Although neither of these agents enjoys the power of regulative control, compatibilists agree that this is not worth wanting, since this would only enable them to *fail* to track reason. It is also evident that even if there is an external controller in the implantation case, there is no possibility of getting this individual to be guided by anything other than the available reasons. Moreover, in both cases the agent's conduct issues from mechanisms that have causal origins that (ultimately) the agent has no control over. The fact that these causes include other agents in one case and not in the other does not, in itself, make a difference, hard compatibilists claim, that we should be troubled by. Both agents have all the control that is required for moral freedom and responsibility, and the ability of both to function as moral agents and participate in the moral community is unaffected by the (historical) differences in the process of mechanism acquisition.

The 'hard compatibilist' can further develop this point by asking why any one should *care* about different histories when the current time-slice properties are exactly the same. There are two relevant perspectives to consider. First, from the 'external' point of view of those who must deal with these two agents in the moral community — i.e. participate and engage with them in 'moral conversation' (RC, 212) — there is no ability that one has that the other lacks. Both deliberate and act in the same way, since their receptivity and reactivity is the same, and both are equally appropriate targets of 'moral address' (RC, 212). Beyond this, the 'internal' view of both agents is also the same. Both have a similar experience of being agents and seeing themselves as fair targets of reactive attitudes. Clearly, then, both these agents, despite their different histories, satisfy the 'subjective' requirements associated with taking responsibility. In short, from both an internal and external perspective these two individu-

---

23 Kane, *The Significance of Free Will*, 67-8

als are 'inter-changeable' in respect of all powers and abilities that matter to moral responsibility.

It is true, of course, that if both these individuals are only moderately reasons-responsive there will be circumstances in which they will both fail to track reasons. The explanation for these failures will not be the same, because the histories of their mechanism acquisition differ. In the case of the individual with the implanted mechanism, the explanation will make reference to the presence of some form of covert, non-constraining control. Whereas in the other case the explanation will refer only to impersonal, natural causes of some kind, not involving covert control by another agent. The hard compatibilist can argue, nevertheless, that it is not obvious why this difference matters to moral responsibility one way or the other. In one case the failure is 'natural' and 'normal,' in the other it is 'artificial' and 'abnormal,' but in neither case does the agent have any control over the fact that he has failed rather than succeeded to track reason. For *this* kind of control to be possible, the agents would need regulative control and alternative possibilities — something that no compatibilist theory is able to accommodate.[24]

VII

In this paper I have been concerned to examine critically Fischer and Ravizza's attempt in *Responsibility and Control* 'to give a comprehensive account of the kind of control that grounds moral responsibility' (RC, 14). The kind of control required for moral responsibility, they argue, is not some form of 'regulative control' that involves alternate possibilities. What is required is guidance control, which is compatible with causal determinism (RC, 34). Guidance control has 'two separate dimensions'

---

24 The 'hard compatibilist' can, of course, reject the claim that moral responsibility is 'essentially historical' and still accept the 'subjectivist approach' to moral responsibility. More specifically, the hard compatibilist may agree that it will not suffice for moral responsibility that the agent's actions issue from a reasons-responsive mechanism unless the agent also sees himself as agent and as a fair target of reactive attitudes (i.e. understood as time-slice properties). What the hard compatibilist will deny, however, is that an agent who satisfies these conditions through a process of implantation cannot be morally responsible because he does not 'own' the mechanism that issues in action. In general, Fischer and Ravizza fuse the (distinct) claims concerning 'subjectivism' and 'history' very tightly in order to defend their 'soft compatibilist' position regarding implantation problems. There is, nevertheless, scope for separating these two claims more sharply in a way that would allow compatibilists to accept the 'subjectivist approach' without endorsing an 'historical' account of moral responsibility.

that Fischer and Ravizza aim to articulate and defend. The first of these is that the mechanism that actually issues in the relevant conduct must be moderately reasons-responsive. The second is that the mechanism concerned must be the agent's own. The arguments that Fischer and Ravizza provide in defence of this overall position are carefully articulated and imaginatively defended. It is impossible not to admire this book. It will receive — and deserves to receive — a considerable amount of attention and discussion from all those who have an interest in the free will problem. I am confident that this work will prove to be of lasting influence and importance. My own discussion shows that I am not persuaded by the particular arguments that Fischer and Ravizza have given to support the two basic components of guidance control. However, if my own experience is anything to judge by, readers will find this book challenging, stimulating, and highly rewarding. It sets a high standard by which to evaluate and assess other work and theories in the field.

Received: June, 2001
Revised: September, 2001

PAUL RUSSELL
University of British Columbia
Vancouver, BC
Canada V6T 1Z1