



Recent Work on Free Will and Moral Responsibility

Neil Levy*

University of Melbourne

Michael McKenna

Florida State University

Abstract

In this article we survey six recent developments in the philosophical literature on free will and moral responsibility: (1) Harry Frankfurt's argument that moral responsibility does not require the freedom to do otherwise; (2) the heightened focus upon the source of free actions; (3) the debate over whether moral responsibility is an essentially historical concept; (4) recent compatibilist attempts to resurrect the thesis that moral responsibility requires the freedom to do otherwise; (5) the role of the control condition in free will and moral responsibility, and finally (6) the debate centering on luck.

1. Introduction

In earlier times, certainly in the first half of the twentieth century, the debate over free will and moral responsibility seemed to consist in nothing more than a series of mopping up exercises, aimed at overcoming the last vestiges of resistance to compatibilism. These days, things look rather different. Incompatibilism is firmly back on the agenda, in both libertarian and hard determinist guises. Moreover, the focus has decisively changed, away from analyses of the 'could' of 'could have done otherwise'. In this survey, we shall cover six of the topics which have been at the center of recent concern: (1) Frankfurt's argument against the Principle of Alternative Possibilities, aimed at showing that moral responsibility does not require the freedom to do otherwise; (2) the move from questions about the freedom to do otherwise to questions about the source of free actions; (3) the debate over whether moral responsibility is an essentially historical concept; (4) recent compatibilist attempts to resurrect the thesis that moral responsibility requires the freedom to do otherwise; (5) the role of the control condition in free will and moral responsibility, and finally (6) the debate centering on luck.

We begin by briefly sketching the consequence argument, which is widely credited with revivifying the debate (Ginet, 'In Defense of the Principle

of Alternative Possibilities'; Wiggins; van Inwagen, 'Incompatibility of Free Will'; *Essay on Free Will*). Though the consequence argument is not the focus of this survey – simply because it is the one aspect of the debate which is most likely already to be familiar to non-specialists – it is essential to sketch it, inasmuch as it is a constant reference point in what follows.

Informally, the consequence argument runs as follows (van Inwagen has suggested various formalization of the argument):

If determinism is true, then our acts are the consequences of the laws of nature and events in the remote past. But it is not up to us what went on before we were born, and neither is it up to us what the laws of nature are. Therefore, the consequences of these things (including our present acts) are not up to us. (*Essay on Free Will* 16)

The consequence argument is very powerful. It has been credited with breaking the compatibilist hegemony over the free will debate. Together with Frankfurt's argument against the Principle of Alternative Possibilities (to be explored below), it has decisively shifted the landscape of the debate, as philosophers sympathetic to the view that moral responsibility or free will is compatible with causal determinism have looked for ways to vindicate this claim without confronting the consequence argument. It has thus ushered in the era of semi-compatibilism (Fischer, *Metaphysics of Free Will*): the view that moral responsibility is compatible with determinism, whether or not free will is best understood as the ability to do otherwise, and whether or not this ability is compatible with determinism. But the consequence argument has also revived the libertarian project.

There has been, and continues to be, a great deal of productive debate over the consequence argument (for instance, Campbell, 'Free Will'; Kapitan, 'Master Argument'; McKay and Johnson; Warfield, 'Causal Determinism'). Here, however, we concentrate largely on what might be called the post-consequence debate: areas of the free will landscape which, if not directly opened up by the consequence argument, have been given great impetus by both the attempt to find an account of moral responsibility that does not confront the consequence argument – such as the source compatibilist views canvassed in section 3 and the attributionist views discussed in section 6 – and the revived debate over libertarianism, discussed in sections 2, 3 and 7. We hope thereby to give the reader a sense of the current state of play in the debate over whether human beings are free agents.

Before we turn to discussion of this debate, it is worth saying a word about why we treat together what seem to be two separate questions – the question concerning moral responsibility and the question concerning free will. Though moral responsibility and free will are different concepts, the moral responsibility debate is near co-extensive with the free will debate. This is in part because freedom is widely taken to be a necessary condition of moral responsibility, and because the strongest intuitions

about freedom of action are provoked by considering morally laden actions, for which agents are morally responsible if they acted freely. Even when philosophers distinguish between the debates, the tools they use and the argumentative moves they make in discussing whichever question they focus on tend to be the common stock of philosophers across both debates. It would therefore be cumbersome and artificial to treat them separately.

2. Frankfurt's Argument against Principle of Alternative Possibilities

Discussions about freedom and moral responsibility remain heavily influenced by Harry Frankfurt's (1969) well known argument against the Principle of Alternative Possibilities (PAP). PAP states:

A person is morally responsible for what she has done only if she could have done otherwise.

Frankfurt argued against PAP by use of a (putative) counterexample. Roughly, the case is of the following sort: Black wants Jones to shoot Smith and would prefer that Jones do so on his own. Worried, however, that Jones will decide against so acting, Black covertly arranges things so that if Jones shows any sign that he will not shoot Smith, Black will cause him to do so. As it happens, Jones shoots Smith unaided, and in ignorance of Black's plan. In this scenario, Frankfurt argues, Jones shoots Smith freely and is morally responsible for shooting Smith, and yet – thanks to Black's presence – he could not have done otherwise (of course, this case is only a loose and incomplete sketch of what has come to be known as a Frankfurt-type example; there are by now many sophisticated variants on this basic structure).

Debate over Frankfurt's argument reached a fever pitch in the mid nineties with the introduction of a powerful objection to any Frankfurt-type example (Ginet, 'In Defense of the Principle of Alternative Possibilities'; Kane, *Significance of Free Will*; Widerker, 'Libertarianism and Frankfurt's Attack'). The objection focuses on how Black is able to discern Jones's free act before Jones performs it, and sketches a dilemma: the sign Black relies on to discern how Jones will act is either infallible or it isn't. In the former case, the infallibility must be explained by some sort of deterministic relation between Jones's states prior to his action, and the action itself. On this horn, the example begs the question against the incompatibilist. In the latter case, however, the agent retains the ability to do otherwise, and there is thus no threat to PAP.

This defense of PAP (and other related PAP-type principles) has met with various objections (Haji, *Moral Appraisability*; Hunt, 'Moral Responsibility'; McKenna, 'Robustness, Control, and the Demand'; Mele and Robb, 'Rescuing Frankfurt-Style Cases'; Pereboom, 'Alternate Possibilities'; Stump, 'Libertarian Freedom').¹ The puzzle set by PAP defenders seems to consist in this: Frankfurt defenders must produce an example in which, at the

time of free action, the agent could not have done otherwise, but in which alternative courses of action are not ruled out by a sign or other ‘vehicle’ that illicitly presupposes a deterministic relation. Frankfurt defenders have attempted to satisfy this demand in various ways. For illustration, consider a Frankfurt-type example by Mele and Robb (‘Rescuing Frankfurt-Style Cases’). Black initiates a deterministic process in Bob’s brain that will bring it about at *t* that Bob decides to steal Ann’s car. Black does not need to have any idea about what Bob might plan, that is, Black does not rely upon any prior sign of what Bob will decide at *t*. But as it happens, Bob’s own indeterministic mental processes lead to the very result at *t* that Bob decides to steal Ann’s car. The catch is that Black’s device has a peculiar quirk. It is sensitive in such a way that if Bob’s indeterministic process issues at *t* in a decision to steal Ann’s car (rather than in some other decision), then the deterministic process set in motion by Black is rendered inert – it is causally preempted at *but not prior to t*. And so Bob decides to steal Ann’s car on his own at *t*, and yet due to the other (deterministic) process set in motion by Black, Bob could not do otherwise than decide at *t* to steal Ann’s car.

As with the other attempts to defend Frankfurt’s argument, Mele and Robb’s efforts have met with various objections (e.g., Kane, ‘Responsibility, Indeterminism’; Pereboom, *Living without Free Will*; Widerker, ‘Frankfurt’s Attack’). The details of the various objections need not concern us here nor the replies Mele and Robb (‘Bbs, Magnets and Seesaws’) have advanced. What is of interest is a salient concern amongst critics of Mele and Robb’s approach. The concern is that Mele and Robb’s example does not rule out all *relevant* alternative possibilities, and so does not clearly confound PAP.²

What is meant here by ‘relevant’? Notice that, on Mele and Robb’s approach, there does remain a miniscule alternative possibility. Assuming their example is coherent, it is causally undetermined whether Bob decides to steal Ann’s car on his own, or only due to Black’s deterministic device (a point Mele and Robb openly admit). But then not all alternative possibilities are ruled out in their example. Indeed, a bit of reflection suggests that, in so far as the puzzle is set in terms that assume that the agent is not causally determined, it seems that there is *no way* to rule out *all* alternative possibilities in a Frankfurt example.³ There remains at least the indeterminacy associated with whether or not the agent acts on his own. But Mele and Robb are not moved by (what they reasonably regard as) their harmless concession. This is because, as popular phrasing now has it, the alternatives left in their Frankfurt example are not *robust*. The causal openness of these alternatives cannot offer the sort of openness that is of interest to those who seek to show that the freedom pertaining to moral responsibility is a freedom over alternative courses of action. Some of the objections to Mele and Robb’s approach (which shall not be canvassed here) turn on the objection that their examples do not rule out all robust

alternative courses of action (see, e.g., O'Connor; Widerker, 'Frankfurt's Attack'; Mele and Robb's reply, 'Bbs, Magnets and Seesaws').

Robustness, a term first introduced by John Martin Fischer (*Metaphysics of Free Will*), is one of the central contested issues in the debate over Frankfurt's argument. Various writers have developed different notions of it.⁴ The central matter of contention, as first identified by Fischer ('Responsibility and Control'), is what an adequate sort of alternative would be such that a defender of PAP (or a related principle) could say that an agent's possession of such an alternative at the time of action aids in accounting for or explaining her freedom as it bears on her moral responsibility. For example, a causally open alternative that the agent could not voluntarily control, one that was a product of mere randomness standing outside the relevant etiological machinery of action would not be an alternative that would aid one who wanted to defend a conception of freedom in terms of control over alternative courses of action. Attention to what various writers count as a robust as opposed to a non-robust alternative is one key to understanding the (sometimes unrecognized) differences amongst various writers in this arena.

It is fair to say that the central line of debate over Frankfurt's argument concerns the attempt to produce examples to get around the puzzle set out above, examples that do not illicitly presuppose a deterministic relation between prior states of the agent and action, and yet rule out all robust alternative possibilities. But there have been other contentious issues. We shall consider three here.

First, one way to untangle the puzzle set out above is to resist the assumption that the agent in a Frankfurt example cannot be determined. Fischer ('Recent Work on Moral Responsibility') first argued for this strategy; Haji and McKenna have defended it as well ('Dialectical Delicacies'; 'Defending Frankfurt's Argument'). Briefly, the assumption that one must work with an agent who is not determined issues from the assumption that the Frankfurt Defender must convince *the incompatibilist* that the freedom to do otherwise is not relevant to moral responsibility (and to the freedom that moral responsibility requires). Hence, the Frankfurt Defender, it is supposed, must work with a libertarian free agent. But suppose that the necessity of freedom to do otherwise for moral responsibility is regarded as in dispute *prior to* the debate between compatibilists and incompatibilists. On this approach, it is an open question whether anyone is free or responsible at all (perhaps the hard determinist or the hard incompatibilist is correct). Suppose, rather, that whether an account of freedom should meet the desiderata set forth by PAP (or some PAP-type principle) were open to dispute in the first place. *Then* there would be no obligation on behalf of the Frankfurt Defender to begin with the case of a libertarian free agent. Instead, the Frankfurt Defender could proceed by simply considering a Frankfurt example making no particular assumptions about the truth or falsity of determinism, and asking whether, *if* the agent

is not free and responsible, is this due to the fact that he has no (relevant) alternative possibilities? Since one manner of ruling out these alternatives (the presence of a counterfactual intervener such as Black) seems irrelevant to whether the agent is free and responsible, perhaps other manners of ruling them out are equally harmless. Hence, even if determinism also rules out these contested alternative possibilities, certain features of Frankfurt examples *suggest* that this has no bearing on whether that agent is free and responsible. Clearly, with the hedge 'suggest', this line of argumentation is not decisive. Others have offered interesting objections to it (e.g., Goetz, 'Frankfurt-Style Compatibilism'; Palmer).

Second, another way to frame the debate over PAP in light of Frankfurt's attack is by way of deontic principles one step removed from considerations of moral responsibility. Consider just one facet of moral responsibility, blameworthiness. Suppose it is granted that one is blameworthy for an action only if she ought not to have done what she did. But now, if we are in the business of ought judgments, then relevant deontic principles about abilities come along for the ride. If 'ought' implies 'can' (OIC), then blameworthiness for what one has done implies that one could have refrained from acting as one did. Given this, some have argued that if, in a Frankfurt scenario, the agent genuinely could not have done otherwise, then the agent is not blameworthy (or not morally responsible in a relevant way). Hence, PAP is not refuted after all (see e.g., Copp, 'Defending the Principle of Alternative Possibilities'; "Ought" Implies "Can"; Widerker, 'Frankfurt on "Ought implies Can"'). Naturally, others have resisted (e.g., Fischer, 'Recent Work on Moral Responsibility'; Haji, 'Alternative Possibilities'; Yaffe, "Ought" Implies "Can"). Here too the debate is not settled, and much turns on whether the 'can' at issue in OIC is the same as the one that pertains to PAP. Another issue, of course, is whether Frankfurt-type examples should simply be regarded as counter-examples to OIC along with PAP.

Third, Widerker has sketched a novel response to Frankfurt Defenders. He questions whether, in a Frankfurt scenario in which all robust alternative possibilities are ruled out, the agent is morally responsible ('Frankfurt's Attack on Alternative Possibilities'). That is, he questions the moral assumption while granting for the sake of argument the metaphysical assumption that such Frankfurt scenarios are possible. Widerker asks, if you regard the agent as blameworthy for what he did in a Frankfurt scenario, *what would you have had him do?* Widerker has dubbed this the W-defense, and he challenges any Frankfurt defender to give content to a proper answer. The trouble is that either the Frankfurt defender cannot give a positive answer (which invites charges of unfairness), or in giving one, an alternative is identified (which appears to play into the PAP defenders' hand). Frankfurt himself opts for the latter strategy ('Some Thoughts Concerning PAP') (see also Zimmerman, 'Moral Significance'). But others have resisted giving a positive answer, and have replied to Widerker by inviting him and other

PAP defenders to consider the moral quality of what the agent has done, thereby offering a positive basis for a judgment of moral responsibility (e.g., McKenna, 'Relationship'; Pereboom, 'Reasons Responsiveness').

3. Source Compatibilism and Source Incompatibilism

It is often assumed that Frankfurt's argument against PAP is an argument for the truth of compatibilism. But this is a mistake. Frankfurt's argument, if sound, does no more than aid the compatibilist in defeating one type of argument for incompatibilism. If freedom over (relevant) alternatives is not necessary for moral responsibility (or the sort of freedom that moral responsibility requires), then, as the Consequence Argument contends, the incompatibilist may well be right that determinism rules out the freedom to do otherwise: $\sim\langle\rangle(D\&FO)$. But as this freedom is not necessary for moral responsibility, $\sim\langle\rangle(D\&FO)$ cannot figure as a premise in an argument for the incompatibility of determinism and moral responsibility (or the sort of freedom required for moral responsibility).

In light of the above, Fischer first suggested that an incompatibilist may accept Frankfurt's argument but yet object that determinism rules out moral responsibility due to the manner in which it brings about an agent's actions ('Responsibility and Control'). Of course, it is open to the compatibilist who accepts Frankfurt's argument to contend that nothing in the actual causal sequence giving rise to non-deviant human actions is threatened (regarding responsibility) under the assumption of determinism. Here the controversy is over the source of agency, rather than with freedom involving alternatives to what one actually does. Call all those theorist who agree with the conclusion of Frankfurt's argument *Source Theorists*, who then can be parsed into the camps of *Source Compatibilists* and *Source Incompatibilists*.⁵ Call those who reject Frankfurt's conclusion *Leeway Theorists*, who then can be parsed into the camps of *Leeway Compatibilists* and *Leeway Incompatibilists*.⁶

For the most part, Source Compatibilists have divided into two camps as regards the proper account of free agency as it bears on moral responsibility. One camp explains the freedom at issue in terms of a mesh between structural aspects of a well-functioning agent's psychology as they issue in non-deviant exercises of agency. The view is most often associated with Frankfurt's hierarchical model ('Freedom of the Will'). Acting freely in the needed way is a matter of an agent's action-producing desires not being in conflict with the higher order desires with which she identifies. She acts on the desires that she wants to act upon, where her wanting reveals what she as an agent identifies with. Another camp explains the freedom at issue in terms of responsiveness to reasons. An agent acts freely in the needed way if she was, in some manner (perhaps by way of certain features or mechanisms of her own agency) responsive to reasons. When she acts freely she acts in response to reasons, and is so disposed that, were

other relevant reasons present, she would have adjusted her conduct in ways that reveal her moral competence. The view is most often associated with Fischer (*Metaphysics of Free Will*) and Fischer and Ravizza, but various others have also developed the idea (e.g., Haji, *Moral Appraisability*).⁷

Assuming the soundness of Frankfurt's argument – which of course is certainly far from a settled matter – perhaps the most lively issue currently concerns the arguments left to Source Incompatibilists. These incompatibilists cannot rely upon that traditional incompatibilist staple, the Consequence Argument, and so they must produce arguments to establish the incompatibility of determinism with freedom and moral responsibility due to considerations about the sources of agency. In our estimation, there are three.⁸ One is the Direct Argument (e.g., van Inwagen, 'Incompatibility of Responsibility'; *Essay on Free Will*; Warfield, 'Determinism and Moral Responsibility'). Another is the Ultimacy Argument (e.g., Kane, *Significance of Free Will*; Smilansky; G. Strawson). A third is the Manipulation Argument (e.g., Kane, *Significance of Free Will*; Pereboom, 'Determinism *Al Dente*'; 'Alternate Possibilities'; Taylor, *Metaphysics*).⁹ In the remainder of this section, we shall consider the first two of these arguments, in the next, we shall consider the third.

A proper formulation of the Direct Argument is beyond the scope of this paper (see McKenna, 'Hard-Line Reply' for a fuller presentation). We opt here for an informal gesture at the main idea behind it. The Direct Argument, like its cousin the Consequence Argument, makes use of a modal principle that transfers a certain modal status regarding one proposition to other propositions that are entailed by it. The Consequence Argument works with some notion such as *inability to render false*. The Direct Argument works with the notion *non-responsibility for*, or as van Inwagen puts it, *not even partly morally responsible for* ('Incompatibility of Responsibility'; *Essay on Free Will*). So suppose that p and that no one is even partly morally responsible for the fact that p , which we can formulate as 'NR p ', and suppose that p entails q , and no one is even partly morally responsible for the fact that p entails q , that is, NR($p \rightarrow q$). According to the Transfer of Non-responsibility rule of inference TNR, from NR p and NR($p \rightarrow q$) we can conclude that NR q .

Given TNR, the Direct Argument runs as follows. Determinism can be characterized as the thesis that the facts of the past and the laws of nature ($p \& l$) jointly entail one unique future, f . This might be put as follows: $\Box (p \& l \rightarrow f)$. Granting that no one is even partly morally responsible for the conjunct $p \& l$, that is, NR($p \& l$); and no one is even partly morally responsible for that fact that determinism is true, that is, NR($p \& l \rightarrow f$); it follows that no one is even partly morally responsible for any fact about the future, that is NR(f). Hence, if determinism is true, no one is even partly morally responsible for the fact that, for example, she chooses to lie, or to act cruelly towards an innocent, and so on. According to the Direct Argument, if determinism is true, non-responsibility for the facts

of the remote past and the laws of nature transfers along relations of entailment at deterministic worlds to non-responsibility for anything that an agent does.

Van Inwagen labels this the Direct Argument since, according to him, it establishes an incompatibility between moral responsibility and determinism directly. It does not attempt to show that some condition necessary for moral responsibility (free will or the ability to do otherwise) is incompatible with determinism. Hence, even if it could not be established beyond dispute that, for instance, the ability to do otherwise is incompatible with determinism (as compatibilists such as David Lewis contend), that argumentative limitation for the incompatibilist would have no bearing on the Direct Argument, which sidesteps this controversy.

Van Inwagen's ingenious dialectical maneuver has drawn much attention. Mark Ravizza and then Ravizza along with John Martin Fischer (*Responsibility and Control*) attempted to produce counterexamples to the Transfer of Non-responsibility principle at work in the Direct Argument. Ravizza, and then Fischer and Ravizza, showed that one could construct 'two path' cases to a certain outcome such that a person might not be responsible for the facts in virtue of which one path leads to an outcome, but she could be morally responsible for the outcome insofar as she brought it about by way of a different path. Here is a quick formulation of their best known example, Avalanche. Conditions of a glacier at t_1 are such that an Avalanche will occur at t_2 , which will result in the destruction of a camp at t_3 . Betty is not even partly morally responsible for the facts in question. But as it turns out, unaware of this, Betty sets off dynamite at t_2 with the intention of causing an Avalanche that will destroy the camp, and as it happens, the camp is destroyed at t_3 due to the Avalanche Betty set off at t_2 . This is a counterexample, Fischer and Ravizza hold, to TNR.

But others have contended that TNR simply needs to be reformulated so as to rule out cases like Fischer and Ravizza's while still capturing what is intuitive and argumentatively forceful in this incompatibilist insight. Thus, various writers have offered reformulations that restrict Transfer NR to one path cases (Ginet, 'Libertarianism'; McKenna, 'Source Incompatibilism'; Stump, 'Direct Argument'). Fischer ('Frankfurt Style Compatibilism'; *My Way*) has countered that these reformulations come closer and closer simply to expressing the incompatibilist intuition that determined agents cannot be morally responsible, which is a patently question begging move. The controversy is far from settled.

David Widerker ('Farewell to the Direct Argument') has taken a different approach in resisting the Direct Argument. He contends that van Inwagen is wrong to think that the argument can stand independently of settling the question as to whether determinism is incompatible with the ability to do otherwise. The Direct Argument, Widerker contends, relies upon the assumption that determinism and the ability to do otherwise are incompatible. For instance, Widerker asks, in the absence of the assumption

that the laws of nature are unavoidable for a person, what reason would there be to think that one is not even in part morally responsible for them? Furthermore, Widerker argues, the compatibilist who holds fast to her preferred analysis of the ability to do otherwise can argue that, at a determined world, it might be true that no agent is morally responsible even in part for the facts of the remote past or the laws of nature, nor for the fact that determinism is true at that world. But if she can do otherwise, then she might well be morally responsible for what she does. Ergo, until the incompatibilist has refuted the compatibilist on the battleground of 'ability to do otherwise' there will always be a clear counterexample available to the critic of the Direct Argument.

Widerker's assault on the Direct Argument is compelling, but readers will notice what a crucial role the ability to do otherwise plays for Widerker. Perhaps Widerker asks it to do more work than it can. For example, without it, he seems to suggest, one could not make the case for non-responsibility for the laws of nature. But mightn't there be other sufficient grounds for non-responsibility for the laws irrespective of considerations of unavoidability? Similarly, Widerker assumes that some compatibilists, Leeway Compatibilists, could rely upon their account of ability to do otherwise in producing a counterexample to the Direct Argument. But this strategy will not be open to the Source Compatibilist who finds that the incompatibilist makes a compelling case for the claim that determinism rules out the ability to do otherwise (e.g., Fischer, *Metaphysics of Free Will*). At a determined world, this Source Compatibilist will not find succor in a counterexample of Widerker's sort, and so will still feel the sting of the Direct Argument. Insofar as Widerker has not offered a refutation of the Direct Argument that could be endorsed by the range of contestants in the debate, his effort remains incomplete, though, as always, intriguing and forceful (see McKenna, 'Saying Goodbye' for a new attempt to resist the Direct Argument that builds on both Widerker's and Fischer's views).

Consider now, if only briefly, a distinct argument that can be used in the service of the Source Incompatibilist, the Ultimacy Argument, which can be formulated either in terms of free action or instead in terms of moral responsibility. Note that, drawing upon van Inwagen's terminology, the argument could therefore be formulated either as an indirect or a direct argument for the incompatibility of determinism and moral responsibility. We shall work with a direct formulation.

The Ultimacy Argument has importantly different variations, which turn on how ultimacy is to be understood. Galen Strawson adopts the extreme view that ultimacy requires being the ultimate source or cause of oneself, a *causa sui*, which is a metaphysical impossibility. According to Strawson, our ordinary concept of moral responsibility has this exceedingly high standard built right into it, a standard that reveals an incoherence in so far as it is thought to apply to finite creatures like ourselves. The argument goes roughly like this:

A person is morally responsible for what she does only if she is ultimately morally responsible for the way that she is mentally as this bears on what she does. No person can be ultimately morally responsible for the way she is mentally. Therefore, no person is morally responsible for what she does.

Notice that Strawson's version does not turn on the thought that determinism is true and that due to this we are not ultimate sources of our actions and therefore not morally responsible. Rather, irrespective of determinism, what moral responsibility demands is simply incoherent – that we be ultimate creators of our very selves, he claims. For if we are not, then we act from reasons that ultimately we did not fashion, but that were fashioned for us. In the latter case, we are not responsible for acting on these reasons. Rather, we are simply the hapless victims (or beneficiaries) of the contingencies that led to our possession of them.

Recently, there have been several replies to Strawson. Some work on crucial inferences that he makes, such as the assumption that one is morally responsible for acting on certain reasons only if one is morally responsible for the reasons themselves (Mele, *Autonomous Agents*; Clarke, 'On an Argument'). Others have pointed out that, in modest ways, we can fashion ourselves (Vihvelin, 'Compatibilism'). Others have questioned the exceedingly high standard Strawson seems to find in our ordinary concept of moral responsibility (Fischer, *My Way*). Is this the correct characterization of the concept to begin with? Yet others argue, even if this notion of ultimacy is integral to our ordinary concept of moral responsibility, perhaps it is eliminable on a cautious revisionary program (Vargas, 'Responsibility').

Another way to formulate the Ultimacy Argument is to make the role of determinism central, so that the worry is clearly about how determinism pollutes the proper sources of agency:

A person is directly (and not just derivatively) morally responsible for what she has done only if she is the ultimate source of her action. If determinism is true, no one is the ultimate source of her action. Hence, if determinism is true, no one is morally responsible for her actions.

Here ultimacy is characterized in something like the following manner. If an agent is the ultimate source of her action, then she contributes some necessary condition for it, NC, such that there are no sufficient conditions for NC that obtain independently of her (e.g., Kane, *Significance of Free Will*; Smilansky). Of course, if determinism is true, ultimacy, construed as such, is unattainable.

A delicate question for the advocate of this Ultimacy Argument is what further content can be given to the demand for ultimacy. As it is, the condition that it identifies is a negative one, namely, the absence of a certain condition. If this is all that ultimacy comes to, it looks as if it involves merely the expression of the incompatibilist demand, and not something that should have an intuitive pull on us in the absence of our theoretical convictions. Kane (*Significance of Free Will*), for instance, argues

that without ultimacy, we cannot have creativity, or lead meaningful lives. But is this true? Others have simply denied assertions of this sort (e.g., Fischer, *Metaphysics of Free Will*).

Now it might be that some would find the world a preferable place were they to know that they contribute to it something that is not contained in the past. And there is no reason known to us why such a preference would be *irrational*. If they prefer it, then they might prefer to be indeterministic initiators and hence, in this anemic sense, ultimate initiators of their actions. But demonstrating the rational permissibility of such a preference is a far cry from an argument that, absent such indeterministic initiation, an agent's actions cannot be free and she cannot be morally responsible for them. This very modest type of source incompatibilism has been sketched by Mele ("Soft Libertarianism"), who found a precursor to it in a suggestion made by Dennett (*Elbow Room*). For the Source Incompatibilist who hopes to get more from the notion of ultimacy than Mele has suggested she might, it is incumbent on her to give some *positive* (and not merely a negative) account of what ultimacy is, why it matters, and why compatibilists can't have it too.

4. Manipulation and History

The Manipulation Argument is a powerful resource for incompatibilists of both the leeway and source variety. Richard Taylor offered a formulation of it, and it has been improved upon recently by Pereboom (*Living without Free Will*) and Mele (*Free Will and Luck*). Space does not permit a full treatment of the recent and expansive debate on this topic. A bit of simplification will have to do.

The Manipulation Argument relies upon our intuitive reactions to what is supposed to be *objectionable* manipulation of an otherwise normally functioning agent. It is important to keep the emphasis 'objectionable' in mind, since often morally responsible agents are assaulted with and in fact influenced by manipulation, and in many cases no one would regard this as a reason to withdraw judgments of freedom and moral responsibility. Advertisements, coaxing from friends, or simple deceptions of various sorts all involve fairly innocuous cases of manipulations of agency. How does the argument proceed? When manipulation *is* objectionable (say in manner X), we feel that it undermines the proper operation of our capacities as morally responsible agents. Call this the Manipulation Premise. The Manipulation Argument feeds off such cases (involving X) by attempting to show, typically in cases of global rather than local manipulation¹⁰, that an agent so manipulated (in manner X) differs in no relevant respect from a normally functioning agent brought into the same state through a normal history of causal determination. But if being so manipulated is freedom and responsibility undermining, given a No Relevant Difference Premise, so is causal determination.

How might the compatibilist respond? Consider Frankfurt's compatibilist position, which holds that it is sufficient for an agent's acting freely and being morally responsible for her conduct that she acts on the desires that she wants to act upon. Kane (*Significance of Free Will*), for one, has objected that agents with such mental states could be the product of massive social engineering, as is imagined in Skinner's *Walden Two*. A similar charge is made by Don Locke, who speculates that the Devil or a neurologist could fashion a person to be as rich an agent as Frankfurt might imagine. To this, famously, Frankfurt has simply replied that any agent that truly is fashioned so as to satisfy everything Frankfurt demands acts freely and is morally responsible ('Three Concepts'). So Frankfurt's reply to relevant formulations of the Manipulation Argument is to resist the Manipulation Premise according to which agents so manipulated are not free and responsible. More recently, some have attempted to defend Frankfurt's striking position (e.g., Arpaly; Kapitan, 'Autonomy'; McKenna, 'Responsibility'; Vargas, 'On the Importance of History').

Others have resisted the Manipulation Argument as, for instance, Kane or Locke has fashioned it, in a different way (e.g. Fischer and Ravizza; Haji, *Moral Appraisability*; Mele, *Autonomous Agents; Free Will and Luck*). They argue that there is a relevant difference between an agent coming to be a certain way through a non-deviant deterministic process, and being 'zapped' into that state by a mad neurologist. Hence, in response to pertinent cases of global manipulation, these theorists resist the No Relevant Difference Premise. How so? They argue that history matters. Distinguishing between nonhistorical (or current-timeslice) and historical accounts of the conditions for responsible agency, they argue that moral responsibility is an historical concept.¹¹ How an agent comes to be in the relevant state she is at a time can have a bearing on whether she is free and responsible in acting as she does at that time. Hence, two agents that do not differ in terms of their ahistorical properties could nevertheless differ as regards their freedom and responsibility in so far as one came to satisfy the ahistorical conditions by way of a freedom and responsibility conferring process, whereas another came to satisfy the very same ahistorical properties by way of a freedom and responsibility undermining process. So, for instance, the person who comes to have white supremacist convictions through a process of critical reflection differs in important respects from one who came to these convictions through a history of psychologically crippling abuse and indoctrination. Call compatibilists who argue that moral responsibility is a historical concept Historical Compatibilists, and call those who resist the need for a historical requirement Ahistorical Compatibilists.

How the Historical Compatibilist cashes out the historical requirement varies. Fischer and Ravizza, for instance, require that an agent go through a process whereby she comes to take responsibility for certain springs of her actions. On their account, an agent must adopt certain subjective states

in order to be a morally responsible agent. On Haji's (*Moral Appraisability*) view, an agent must come to acquire an authentic evaluate scheme by first being instilled by others in her community with morally stable beliefs and rational principles of critical evaluation, and then by being given the chance to form her own scheme arising from this one. Mele (*Autonomous Agents; Free Will and Luck*) offers a negative historical account whereby an agent acts freely and as a morally responsible agent (Mele has also spoken in terms of autonomy), if, with other conditions satisfied, she acts from values and principles that she has neither acquired nor sustains through any process that bypasses her capacities for critical evaluation with respect to them. Regardless, all of these Historical Compatibilists believe that in a case like the one Locke or the one Kane put to Frankfurt, they can simply deny that the respective agents had properly freedom and responsibility conferring histories. So these Historical Compatibilists thereby deny the No Relevant Difference Premise of the Manipulation Argument.

But how well does the Historical Compatibilist Strategy stand up to recent attempts to advance the Manipulation Argument? Pereboom (*Living without Free Will*) rests his Source Incompatibilist position on a version of the Manipulation Argument, one that has a few amendments that mark a genuine advance in the debate. Pereboom adds to his argument a set of cases in which the manipulation in question is meant to come closer, step by step, to a case of a normal agent acting at a determined world. At each step Pereboom employs a principle of treating like cases alike, and so moves from an outlandish case of manipulation to a more mundane one, and then simply to a case in which determinism is true. His argument, dubbed the Four Case Argument, works from three cases to the fourth, which is just the case of a normal agent acting at a deterministic world. The principle of treating like cases alike, and the march from the extreme to the more mild case is one major way of advancing the debate, since it aids in establishing the No Relevant Difference Premise.

In very short order, here are how the cases go. In Case 1, Professor Plum is manipulated from moment to moment by a team of neuroscientists to kill Ms White. Plum is so manipulated that he satisfies a conjunct of differing compatibilist accounts of free agency. Plum kills Ms. White for selfish reasons. But, Pereboom contends, no one would say that Plum is free or morally responsible for killing Ms White. In Case 2, the team of neuroscientists manipulate Plum from a temporal distance, installing the right sort of program so that at the relevant time in his life, Plum kills Ms White, in the very same fashion and for the very same reasons as in Case 1. Surely, in this case as well, Pereboom reasons, Plum is not morally responsible. In Case 3, Plum is brought up in an environment with a strict indoctrination in favor of often acting rationally egoistically. All unfolds just like it does in Case 2: again Plum kills Ms White for the same reasons. Pereboom draws the same conclusions about Plum in this case; he is not morally responsible. Finally, in Case 4 Pereboom simply considers a

deterministic setting in which, by happenstance, the very same causes that were instilled moment by moment in Case 1, that were set in motion from a temporal distance in Case 2, and that were settled by a strict environment in Case 3, simply unfold in the normal way. In this case as well, Pereboom argues, treating like cases alike, we should conclude that Plum is not morally responsible.

Pereboom adds one further twist to the standard Manipulation Argument as it had been previously advanced. He includes a *best explanation argument*. The best explanation, Pereboom contends, of our intuition in Cases 1 through 3 that the agent is not morally responsible in each is that in each the agent's conduct was the product of deterministic processes that were beyond his control. That is why we should have the same response to Plum in Case 4.

Now return to the Historical Compatibilists. How might they respond? Crucially, they might, and indeed, some have, argued that in some or all of Pereboom's three cases of manipulation, the agent does not satisfy the relevant historical conditions. It looks as if, for example, in Case 1, Plum's agency is in question. It is not even clear that he is an integrated agent at all, and so certainly does not take responsibility for any feature of his own agency at any point in his development (Fischer, *My Way*). And in Case 2 it looks as if the agent's process of development is straight-jacketed in such a way that she does not have the capacity for critical reflection (Mele, 'Critique of Pereboom's Four-Case Argument'). To these worries Mele has also argued, contrary to Pereboom, that the best explanation of our intuitions in Cases 1 through 3 is the presence of the manipulation and not the fact of determinism.

But Pereboom's basic response is seductive. Whatever historical demands the Historical Compatibilist wishes to add, he will simply include those conditions in his cases of manipulation ('Defending Hard Incompatibilism'). Even in Case 1, he argues, he can include the historical development required by any of his adversaries by seeing to it that the moment by moment manipulators go through the proper routine.¹² It seems that Pereboom has a point. Try as they might, compatibilists of any stripe will eventually have to admit that it is at least in principle possible to imagine a case of manipulation that includes everything a compatibilist would demand (barring of course a mere *ad hoc* amendment that no global manipulation is allowed at all). If so, eventually, the compatibilist will be forced to a point at which she can no longer respond to the advocate of the Manipulation Argument by denying the No Relevant Difference Premise. Instead, she will either have to resist the Manipulation Premise or concede defeat (see also McKenna, 'Responsibility').¹³

It might be thought that, in the end, this shows that the incompatibilist wins the day. For if Historical Compatibilists shied away from resisting the No Relevant Difference Premise in the face of Frankfurt's Nonhistorical Compatibilist position, why would they now think it better to resist in

light of newly refashioned cases of manipulation, ones that have all the juicy historical stuff included? The answer might well be that in such cases, it is now no longer so intuitively clear that the manipulation imagined is all that objectionable. The real test, then, for both the compatibilists and the incompatibilists might simply be to have an argument over a manipulation case that does as much as any could to install all that a compatibilist – even a Historical Compatibilist – might want.

This is just what Mele does in the most recent rendition of the Manipulation Argument. Mele simply cuts out all possibilities where a compatibilist might press on the No Relevant Difference Premise. To do this, he considers a creation case. The goddess Diana creates a zygote that will become Ernie, and she does so in such a fashion that one day Ernie performs some act, for instance, poisoning his aunt. Filling in relevant versions of the premises we have identified earlier, Mele labels this the Zygote Argument, and he asks how a compatibilist should respond to this argument. He contends both that it would not be reasonable for the compatibilist to resist the No Relevant Difference Premise, and that the compatibilist should resist the Manipulation Premise. The agent, Ernie, in this case, is free and morally responsible, despite the manner in which he was created. Of course, Mele contends, the incompatibilist should not agree on this point. And so here we have what looks like a stalemate.

But appearances are deceiving. For there remain the undecided, the agnostics, like Mele, who want to hear more from either camp before deciding which position is more compelling. In this way, the force of the Manipulation Argument is to be evaluated in terms of what further considerations either side can marshal for the attractiveness of their respective positions. For instance, can the incompatibilist get around the luck problem (as presented in section 7 below)? In the absence of a satisfactory response to that, and with a willingness to believe more strongly that there are morally responsible agents than that there are not, the agnostic might have good reason to lean towards the compatibilist camp. But were the incompatibilist able to address the luck problem, then perhaps the incompatibilist could say what more there is to indeterministic initiation of action that makes it preferable to any compatibilist variety. Then the agnostic might have reasons to lean in the other direction. In conclusion, Mele leaves the Manipulation Argument (or rather, as he might call his transformed version, the Creation Argument) as a way to help us diagnose the bases of our disagreements as seen through the eyes of those no side has yet convinced.

5. The Compatibilist Rejection of Frankfurt's Argument against PAP

Prior to the emergence of Frankfurt's argument against PAP, the free will debate was dominated by rival analyses of the 'could' of 'could have done otherwise', where this was understood as entailing an agential ability. Many compatibilists welcomed Frankfurt's attack on PAP, inasmuch as it

allowed them to abandon a debate which had bogged down into a dialectical stalemate (to use Fischer's useful phrase). However, not all compatibilists accept the soundness of Frankfurt's argument against PAP. Some continue to believe that the incompatibilist must, and can, be fought on the traditional terrain of agential abilities (e.g., Berofsky, 'Ifs, Cans, and Free Will'; Campbell, 'Compatibilist Theory'; M. Smith; Vihvelin, 'Free Will Demystified'). In this section, we will sketch what we call the dispositionalist argument against Frankfurt.

Frankfurt's argument is designed to show that agents can be responsible despite lacking the freedom to do otherwise. The dispositionalist argument aims to demonstrate that agents in Frankfurt-type examples do have relevant alternative possibilities after all, and so retain the freedom to do otherwise. Michael Smith has advanced a possible-worlds analysis of agential abilities. He suggests that we answer questions about whether agents are able to do otherwise by asking whether they would do otherwise in nearby possible worlds. If, in the right set of possible worlds, the agent would do otherwise, he possesses the ability to do otherwise in the actual world, and it is therefore false that the agent lacks relevant alternative possibilities.

Of course, the plausibility of this claim depends, in important part, on Smith providing us with a principled way of settling which possible worlds are relevant for fixing agential abilities. Smith argues that agential abilities are constituted by their intrinsic properties. Hence, only possible worlds in which the agent's intrinsic properties are held constant are relevant. But since the counterfactual intervener of Frankfurt-type examples is not an intrinsic property of the agent, we are entitled to bracket the intervener. Thus, the relevant possible worlds are those in which the intervener is absent (or dormant). But without the counterfactual intervener, the agent *can* do otherwise (in some suitable compatibilist sense of 'can'). Hence, the agent has relevant alternative possibilities in the actual world; she possesses the abilities that make it true that she can do otherwise. Smith motivates this claim by comparing agential abilities to intrinsic properties, and counterfactual interveners to 'finks' which prevent them from being manifested. Just as an object continues to possess a disposition, even in the presence of an extrinsic fink which will prevent that disposition being manifested, so it is true that agents possess the relevant abilities to do otherwise, despite the presence of counterfactual interveners.

On this view, agential abilities are held to consist in bundles of dispositions. The view is further developed by Vihvelin ('Free Will Demystified'). For Vihvelin, the dispositionalist argument gives the compatibilist more than merely a means of replying to the proponents of Frankfurt-type examples, thereby returning the debate to the dialectical stalemate that prevailed before their emergence. They also provide us with a diagnosis of that stalemate, and a way of moving beyond it. The stalemate arose, she argues, from the fact that compatibilists attempted to analyse the relevant abilities in terms of simple conditionals (e.g., an agent can ϕ in C just in case the

agent *would* ϕ in C, if she chose to ϕ in C). Incompatibilists were right in rejecting such analyses as inadequate, and compatibilists had no adequate reply. But the development of possible worlds semantics for conditionals, to which Smith appeals, also gives us a means of resurrecting the project of giving a conditional analysis of ability, roughly along the lines of the conditional analysis of finkish dispositions (Campbell 1997 presents views that prefigure the dispositionalist argument).

Does the dispositionalist argument succeed? Cohen and Handfield argue that it does not. Cohen and Handfield contend that abilities can be *radically* finked, and abilities that are radically finked are lost. Frankfurt-type examples can be constructed in which agential abilities to actualise alternative possibilities are radically finked, and in which therefore the agent cannot do otherwise, but in which they are nevertheless morally responsible. Finks are radical, they argue, when they are intrinsic to the object. They suggest that a case that Frankfurt advanced in another context, the case of the willing addict (Frankfurt 'Freedom of the Will') can be seen as an example of an agent whose abilities are radically finked. Due to his addiction, the willing addict has a craving for drugs sufficient to force him to ingest them. However, because he consumes drugs willingly, his addiction plays no role in his actual drug-taking behaviour. It is false that the addict could refrain from taking drugs, but he is, intuitively, responsible for his drug taking behavior (since he takes the drugs willingly, and his addiction plays no role in his behaviour). Since the willing addict's addiction is constituted by intrinsic properties of his, we cannot abstract away from it in asking whether he retains the ability to do otherwise. Hence if we apply the dispositional analysis, we get the result that the willing addict has lost the ability to do otherwise.

Cohen and Handfield's counterexample to the dispositionalist is intriguing, but it is far from clear that it succeeds. Attention to the logic of Frankfurt-type examples, at least as understood by Fischer and Ravizza, suggests that the apparent success of their counterexample rests on a slide from the *agent* to the *mechanism* upon which the agent acts. To see the point, consider how Fischer and Ravizza's understand the moral responsibility of agents in Frankfurt-type examples. Fischer and Ravizza argue that agents are morally responsible for their actions when the mechanisms upon which they act are moderately reasons-responsive, where a mechanism is moderately reasons-responsive if it is capable of tracking reasons in some kind of patterned way, and in addition there is at least one possible world in which it would cause a different action in response to a reason. Fischer and Ravizza suggest that we test whether a mechanism is moderately reasons-responsive by asking how it would respond in various counterfactual scenarios. They argue that considering how the agent would respond in various counterfactual scenarios reveals the properties of the mechanism upon which the agent *actually* acts. If it is the case that the mechanism would cause the agent to act otherwise in some possible world in which she was presented

with a sufficient reason to do otherwise *and* in which the counterfactual intervener was not present, then the mechanism is properly reactive to reasons; similar tests would tell us whether it was properly responsive.

The similarities between the test Fischer and Ravizza use for moral responsibility and the question that Smith and Vihvelin urge us to ask to discover whether the agent possesses the ability to do otherwise should be apparent. If the mechanism upon which the agent acts is moderately reasons-responsive, where moderate reasons-responsiveness is manifested by appropriate reactions and responses in nearby possible-worlds, then the agent is morally responsible, Fischer and Ravizza argue. Similarly, Smith and Vihvelin test for the ability to do otherwise by asking whether the agent *would* do otherwise in nearby possible worlds. There are two crucial differences between the approaches: first, Fischer and Ravizza *deny* that the agent can do otherwise, since their test for this ability differs from that suggested by Smith and Vihvelin. Whereas Smith and Vihvelin test for the ability to do otherwise by asking whether the agent has the appropriate dispositions, as she is actually constituted, Fischer and Ravizza hold that the relevant question is whether the circumstances she is in allow for her to do otherwise, given her constitution.

Second, and more relevantly here, Fischer and Ravizza focus on the mechanism upon which the agent acts, whereas Smith and Vihvelin focus on the agent himself. That is, the latter argue that the ability to do otherwise is retained due to intrinsic properties of the agent *other than* those manifested in the actual sequence, hence broadening the scope beyond the actual-sequence mechanism.

Now, it is a large and difficult question whether the focus of debates over moral responsibility is appropriately on the actual-sequence mechanisms, or whether we should follow Smith and Vihvelin in broadening the focus to the agent as a whole. This is not a question we can hope to address here. However, we suspect that Cohen and Handfield's argument moves illegitimately from one focus to the other. That is, it may be the case that the intuition that agents are morally responsible for actions in Frankfurt-type examples is the product of focusing upon the actual-sequence mechanism. It is because the mechanism which actually issues in the action is appropriately reasons-responsive and reactive that we feel confident in holding the agent responsible. But if that's the case, then broadening the focus to include *other* mechanisms, mechanisms which would trump or preempt the actual-sequence mechanism, is illegitimate in the dialectical context. If we fail to bracket these other mechanisms when we ask whether the agent has the ability to do otherwise, we fail to focus on the mechanism which underwrites the ascription of moral responsibility.

If it is appropriate to broaden the focus from the actual-sequence mechanism to the entire agent in the case of the willing addict, we should conclude that the agent is not responsible: it is false that the *agent* is moderately reasons-responsive. It is only if we narrow the focus to the

actual-sequence mechanism that we get the result that the agent is morally responsible. But if we so narrow the focus then *other* intrinsic properties of the agent are no longer relevant to the questions we ask. So, Cohen and Handfield's results only go through if they commit to a mechanism-based and not an agent-based view of reasons-responsiveness.

This is not to say that the Frankfurt argument against PAP emerges unscathed from this debate. The dispositionalists have done us a service by drawing to our attention the close analogy between counterfactual interveners and finks. They have shown us that familiar arguments (such as that of Fischer and Ravizza) for the conclusion that agents in Frankfurt-type examples are morally responsible rest upon internalist grounds; that is, we conclude that the agent is responsible because their intrinsic properties are unaltered by counterfactual interveners. But internalism is a controversial thesis. Moreover, it may not prove stable: it may be that though sometimes our intuitions about moral responsibility track properties internal to agents, at other times they seem to track properties that cross the boundary between agents and the world (Levy, 'Agents and Mechanisms'; 'Counterfactual Intervention'). It may therefore be that settling the question whether agents in such cases retain the ability to do otherwise requires us to move to the terrain of the debate over internalism.

6. *The Control Condition*

It is generally (though not universally) supposed that agents are only (directly) responsible for actions that they perform freely. Given this assumption, intuitions about moral responsibility are commonly taken to be reliable guides to the freedom of agents. Now, there are two kinds of conditions traditionally taken to excuse agents of responsibility: ignorance of the nature of one's actions, and lack of control over one's actions. Recently, however, an influential group of theorists have begun to question this latter assumption. They argue that responsibility for an action or omission does not require control: it requires only that the action be appropriately attributable to the agent (Scanlon; A. Smith, 'Responsibility for Attitudes'; 'Control'; Arpaly, 'Merit, Meaning, and Human Bondage').

An action is, in the sense at issue here, attributable to the agent if it is expressive of her attitudes and values. Actions and omissions are expressive of the agent when they are (non-deviantly) the product of where she stands as a practical agent. Attitudes are thus expressive of who I am if they belong to the class of *judgment-sensitive attitudes*. Judgment-sensitive attitudes are attitudes that, in ideally rational agents, are sensitive to reasons, such that these agents have them when, and only when, they judge there to be sufficient reason for them (Scanlon 20). Insofar as we are rational agents, we are not simply 'stuck' with our judgment-sensitive attitudes. Instead, they are the product and the expression of ourselves as agents. We can therefore appropriately be asked to justify them. It makes

no sense to ask someone to justify their height, their skin color or their compulsions, simply because none of these aspects of the person are sensitive to their judgments. But someone can be asked to justify their political views, their fundamental values and their sense of what is important and what trivial. Their judgment-sensitive attitudes reveal where they stand on questions of value.

It is obvious that attributability does not require much in the way of control. Consider an omission (omissions cases play a significant role in motivating attributionist accounts of moral responsibility). A person's failure, say, to recall their spouse's birthday does not seem to be a lapse over which they exercise direct control, of the kind that most theories require for moral responsibility, inasmuch as the lapse is not the product of any intentional action of theirs (that is, any action which is intentionally aimed at bringing about their forgetting). Such theories hold the person responsible for the lapse only if some directly controlled act or omission in the causal chain leading to it can be identified; say an overt decision not to write the date down in their diary. Sometimes such a directly controlled act or omission will figure in the causal history, but it seems very likely that often there will no such action or omission. Nevertheless the lapse is plausibly taken to be expressive of the person's attitudes, inasmuch as it is plausibly true that were they to care more for their spouse, they would not have forgotten her birthday. If we are responsible for our lapses, then responsibility cannot require control; attributability is sufficient.¹⁴

Of course, it would be question-begging to *assume* that an agent is responsible for such a lapse; nevertheless, the intuition that we are sometimes responsible for such lapses is powerful. Volitionists, as Levy ("The Good, the Bad and the Blameworthy") has dubbed theorists who maintain that control is a necessary condition for moral responsibility, can hold agents responsible for some lapses – those which are the product of willed actions, over which the agent exercised the relevant kind of control – but the epistemic conditions on moral responsibility ensure that the proportion of such lapses which can be sheeted home to agents is likely to be small. Only if a lapse is a product of a controlled action *and* the agent foresaw, or could reasonably have been expected to foresee, that the lapse might result from the action can the volitionist hold the agent responsible for it (Vargas, "Trouble with Tracing").

Given that we frequently blame agents for lapses, attributionism is a powerful and plausible view. However, it has been accused of failing to be able to make a sense of a distinction that is itself intuitive; the distinction between *bad* agents and *blameworthy* agents (Levy, "The Good, the Bad and the Blameworthy"). Gary Watson's 1987 discussion of the case of the ruthless murderer Robert Harris illustrates the distinction nicely ("Responsibility and the Limits of Evil").¹⁵ Harris's actions were appropriately expressive of his identity as a practical agent; his total insensitivity to the value of human life. Reading Watson's lengthy description of his crimes certainly

triggers outrage at his actions and our indignation at him. But when we read his sad biography, our reactions are tempered. A volitionist might explain our responses in the following manner: we blamed Harris so long as we thought that his actions were controlled by him, in the appropriate manner (where appropriate control has demanding epistemic conditions, such that an agent is responsible for an action only if he grasps the moral norms that apply to it). When we learn of Harris's abusive upbringing we are inclined to think that it left him unable to grasp the relevant moral norms. If so, we reason that we should retract the judgment that he exercised relevant control over his actions. Hence, we conclude that he is a bad agent, but not a blameworthy one.

Angela Smith ('Control') responds that the bad/blameworthy distinction is one that we should be extremely reluctant to invoke. To regard an agent as bad but not blameworthy is to regard them as passive victims of their character and circumstances. It is to adopt what P. F. Strawson called 'the objective attitude' towards them. In Smith's words, 'it is, in effect, to say that he is not to be regarded as someone to be reasoned with, but merely as someone to be understood, treated, managed, or controlled' ('Control' 388). It is thus to exclude them from the circle of moral agents.

But this claim is subject to dispute. Consider, first, our appraisal of some agents from different cultures and different times. It is at least possible that these agents acted conscientiously, in the light of the moral and nonmoral norms as they saw them. Nevertheless, we want to say that their actions were wrong, and that the attitudes they expressed were bad. We do not, for that reason, regard them as irrational. To the contrary, we might see them as rationally responding to faulty norms. If we go this route, then we may withhold condemnation of them. Many people have the intuition that Aristotle was not blameworthy for keeping slaves or for his sexism; that his actions and attitudes were wrong, but not blameworthy. If we find this intuition compelling, then we shall want to retain a badness/blameworthiness distinction.

Closer to home, we might want to say the same thing about many entirely ordinary agents. It is, surely, a relatively common experience to meet a rational agent who nevertheless exhibits islands of irrationality: discrete beliefs or values which do not cohere well with their other beliefs or which fail to meet evidential standards to which they otherwise adhere, but regarding which they remain committed. The explanation of these islands of irrationality often involves some biographical fact; for instance, the agent may have simply inherited some prejudice. For a shorter or longer period of time, the agent may be unable to recognize that the prejudice is unjustified; moreover, if the prejudice goes unchallenged (perhaps because it is widely shared in her social circle) she may lack the opportunity to reassess it. Naturally it would be question-begging to claim that she is not responsible for the attitude. But the volitionist need not make this claim, to argue against the attributionist reply to the bad/

blameworthy distinction. The volitionist need only to point out that it is – apparently – coherent to think that the agent is not responsible *for that attitude*, but remains a globally responsible and rational agent. It seems false to think that we must choose between global objective and reactive attitudes; instead, it seems plausible to think that we can take the objective attitude toward *some* attitudes of others – indeed, even toward some attitudes of ourselves – while retaining the reactive attitudes toward the person.

Attributionists often distinguish between two kinds of responses to wrongdoing: holding responsible and blaming (or imposing sanctions). The latter, but not the former, require the satisfaction of control conditions, they argue (Scanlon; A. Smith, ‘Control’). This fact suggests the possibility that the debate between attributionists and volitionists is as much verbal as substantive: it may be that volitionists identify responsibility with the attributionist’s blaming, whereas what attributionists call responsibility the volitionist might call (following Watson, ‘Two Faces of Responsibility’) areatic criticism. Making progress on the debate therefore requires arguments for identifying responsibility with one or other notion.

7. *The Luck Argument*

Considerations of luck and chance have long played a central role in debates over free will. They have often been at the heart of compatibilist responses to libertarianism. These responses come in two versions, one metaphysical and one focused on the agent’s reasons. The metaphysical version runs, roughly, as follows. Suppose that an agent’s actions are indeterministically caused by beliefs and desires. In that case, it seems a matter of chance how she acts; given precisely the same beliefs and desires, she might have acted otherwise. But it seems false to think that a chance action is free. This, very roughly, is the *Mind* argument, so-called because several versions appeared in that journal during the heyday of the compatibilist consensus (e.g., Nowell-Smith; Smart).

The recent debate has been dominated by a successor to the *Mind* argument, which differs from it in two ways. First, it does not aim to establish the incompatibility of luck and free will directly; second it focuses more on the agent’s reasons for actions than upon the metaphysics of causation. Once again, the target is libertarian accounts of free will.

Consider, first, an event-causal libertarianism, such as that influentially defended by Robert Kane (*Significance of Free Will*). Kane argues that agents act with (direct) free will when they exercise *dual rational control*, which is the ability to choose, for reasons, from (at least) two options. We possess such an ability only when we have very strong reasons for incompatible actions, and no decisive reasons against at least two of our options. In such a case, Kane argues, we may simultaneously try to make two choices or decisions. The resulting psychological conflict can disrupt the thermodynamic equilibrium in our brains, with the result that the quantum-level indeter-

minacies, usually too insignificant to have any effect at the macroscopic level of neural firings, are magnified. This magnification begins an indeterministic process in our brains, and as a result, it becomes genuinely undetermined which of the options we shall ultimately select.

Kane argues that if, under these conditions, an agent tries to ϕ and to ψ (where ϕ -ing and ψ -ing are incompatible), whichever option she ultimately selects, she chooses freely. He points out that whichever option she selects, she chooses because she was *trying* to select it, she had strong *reasons* in favor of it, and (he claims) whichever option she selects she will endorse as what she was trying to do. In other words, whatever option she selects plausible compatibilist conditions on freedom of action will be satisfied for her with regard to that choice. All Kane has added to an adequate compatibilist picture is the requirement that it be genuinely undetermined which option she selects. But indeterminism in the causal path cannot rule out moral responsibility, and (therefore) freedom of action. We cannot say that because it was undetermined which option an agent selects, she fails to select it freely. Kane presents us with the following analogy. Suppose a man strikes a glass table, intending to break it. But suppose it is genuinely undetermined whether he succeeds in breaking it, due to the structure of the glass. Finally, suppose that he does break the table. Clearly, the man is responsible not merely for *trying* to break the glass table, but also for actually breaking it. Similarly, Kane claims, in cases where his conditions on direct freedom of action are met, and it is genuinely undetermined whether an agent (say) ϕ s, if she ϕ s she is responsible for ϕ -ing, and therefore ϕ s freely (Kane, 'Responsibility, Luck and Chance').

Kane's claim that because the agent endorses either option, however she chooses she acts freely appears designed to block the *Mind* argument and related objections from luck. Critics – compatibilist and libertarian – have been unimpressed. They continue to claim that it is simply a matter of *luck* which action she performs. We can drive the point home utilizing van Inwagen's ('Free Will Remains a Mystery') replay argument. Consider an agent who satisfies Kane's conditions for direct freedom, who is faced with an undetermined choice and eventually selects one of her options. Now, suppose that we replay the relevant stretch of the history of the universe, where the relevant stretch includes the agent's deciding whether to ϕ or to ψ . Let this stretch of history be replayed multiple times. On some of these replays the choice would go one way, and in others it would go the other (if the agent was equally inclined toward each option, we can expect that in about half of these worlds she would ϕ , ψ -ing in the rest). Now, what brings it about that she ϕ s or ψ s? Nothing, apparently – nothing about her, her reasons, desires, volitions, or tryings – except luck. She does not control *which* alternative she settles for. The indeterministic causal process settles that. So long as our actions or decisions are the products of indeterministic event-causal processes, *we* do not exercise any greater control over them than we would in a deterministic world. *We* do

not control the indeterministic process that settles how we decide or act (Mele, 'Ultimate Responsibility'; Strawson, 'Unhelpfulness'; Haji, 'Compatibilist Views').

Kane's reply may well be successful for some pairs of actions. Suppose it is undetermined precisely *how* an agent performs an action. In that case, it seems that the fact that she does not control whether she performs it in one way or another does not seem to detract from her responsibility for it. There may even be cases in which it is undetermined which of two incompatible actions an agent performs, yet in which she is intuitively responsible for the action performed: consider cases in which it is a matter of chance whether an agent performs one equally blameworthy action or another. However, Kane requires something stronger. Since he believes that agents are only responsible for their actions if they are the ultimate sources of their actions, where someone is the ultimate source of their action when they play an irreducible role in forming the character which they express in action, he needs it to be the case that, at least sometimes, the incompatible actions between which the agent chooses, and which are causally open to her, differ in moral valence. That is, the options must differ substantively from each other in their implications for the goodness of the agent's moral character. But if it is causally open to an agent to choose between two or more options which differ in moral valence, and it is a matter of luck which she chooses, it seems false to say that she is responsible for the option she actually selects. Suppose, for instance, she performs a morally good action. We would withhold praise if we believed that she might nevertheless have chosen a morally bad action, and that nothing about her – her character, efforts, attitudes, and so on – explain why she chose the first rather than the second.

Now, something akin to this argument has been utilized by prominent defenders of agent-causal libertarianism to motivate their own view (O'Connor; Clarke, *Libertarian Accounts*; Pereboom, 'Is Our Conception of Agent Causation Coherent?').¹⁶ What is missing from event-causal libertarianism, they claim, is the ability by the agent *herself* to exercise freedom level control over which option she selects. But, they argue, agent-causation is itself such an ability: the power agents possess, as substances, to initiate new causal chains *ab initio*. Hence agent-causal libertarianism is uniquely satisfactory as an account of free will. However, some philosophers have argued that the agent-causal theory is itself vulnerable to the luck objection. That is, they have argued that just as it is merely a matter of chance whether, given event-causal libertarianism, an agent ϕ or ψ , inasmuch as if we replayed the scenario in identical worlds, sometimes the agent would ϕ and sometimes she would ψ , and nothing about *her* would account for the difference between worlds, so too an agent-cause would sometimes ϕ and sometimes ψ , and nothing about *her* would explain which she does. Roll back history and sometimes the agent chooses one way, sometimes another. Given that there is no difference in the two

worlds, and no difference, especially, in the agent and her reasons, before the moment of choice, it is hard to see what can account for the difference in the choice other than chance (Haji, 'Active Control'; Mele, 'Critique of Pereboom's Four-Case Argument').

Proponents of agent-causal libertarianism (of its coherence, if not of its existence) reject this argument. Indeed, they hold that it begs the question against them. The agent-causal power, they argue, just *is* the power to choose between alternative courses of action, everything else, including the agent's reasons, held fixed. Thus, demanding that the agent-causalist provide an explanation of the decision actually made in terms of the agent's reasons begs the question against the very possibility of agent-causation. By hypothesis, the agent-causalist holds, the agent's reasons are precisely the same across worlds in which she chooses different courses of action. But it doesn't follow that mere chance explains the agent's ϕ -ing in some worlds and ψ -ing in others. Instead, the cross-world difference is to be explained by *the agent's exercise of her agent-causal power*. The difference is a difference in the way the agent exerted her free will (Clarke, 'Agent Causation'). Proponents of the luck objection are asking for further elucidation where, in principle, there is none to be had.

Suppose, however, that the agent-causal view of human action is correct; that is, that we each possess an irreducible power to choose between causally open alternatives. In that case, it seems, the *Mind* argument would fail, or straightforwardly beg the question. Nevertheless, the argument from luck remains powerful against this view. If there is an agent-causal power, it seems that this is a power that cannot be exercised for reasons. Since (by hypothesis) the options between which she chooses are causally open to her with all her reasons held fixed, the final choice between these options cannot be made *for reasons*. Her reasons may constrain the options between which she chooses, but – on pain of double counting of reasons – she cannot make the choice for those reasons. Hence, she makes it blindly. Once again, she may count as morally responsible for the choice if the difference between the options causally open to her are relatively trivial, or even if they have much the same moral valence. But if the agent-causal libertarian is to hold an agent responsible for choosing a wrongful action, on the grounds that it was causally open to her to choose a better alternative, she must explain how she can make *that* choice for reasons.

One interesting question concerns the seriousness of the luck objection for libertarianism. The debate often seems to presuppose that if the luck objection succeeds, libertarian agents exercise less control over their (putatively) directly free actions than compatibilist agents, inasmuch as their actions are subject to luck. But a different conclusion might be drawn: that libertarian agents are shown, by the luck objection, to exercise *no more control* than compatibilist agents.

Mele (*Autonomous Agents*; 'Soft Libertarianism') has developed, though not endorsed, a libertarian position upon which (putatively) directly

free actions are undetermined inasmuch as it is undetermined what considerations occur to agents in the course of deliberation. To the extent to which there is genuine indeterminism, what considerations come to agents' minds is subject to chance, but, Mele points out, compatibilist agents do not control what considerations come to mind during deliberation either; hence this kind of libertarian agent is *no more* subject to luck than is a compatibilist agent. If it is rational (or at least not irrational) to prefer that one's actions do not have sufficient conditions which predate their proximal deliberative springs, then it might be rational (or at least not irrational) to prefer to be a libertarian agent of the kind Mele describes than a compatibilist agent, despite the fact that the libertarian agent exercises no more control than the compatibilist agent.

Mele has labeled this position Modest Soft Libertarianism (MSL). Its modesty is due to the place at which Mele suggests that the libertarian locate the indeterminacy, further back in the causal chain leading to action rather than, for instance, at the time of formation of the judgment, intention formation, choice, or decision. Its softness is due to the fact that, in the form Mele presents it, it does not *entail* the falsity of compatibilism. Rather, it is claimed that libertarianism is a sort of agency that one might reasonably prefer. (Of course, it is open to hard incompatibilists to keep the modest part and drop the soft bit, but it is difficult to see what argument one might offer for this move, if it is granted that a libertarianism of this sort offers no more control than a compatibilist can provide.)

Given that Mele's MSL agent is no more subject to luck than is a compatibilist agent, one pressing question is whether *the compatibilist agent is unacceptably subject to luck*. One way to understand the problem raised by manipulation cases is as a problem of luck: not what Mele calls *present* luck, luck at or around the time of (putatively) directly free action, but *constitutive* luck. A manipulated agent is subject to a bizarre kind of constitutive luck; luck in the traits and dispositions with which one finds oneself. Mele (*Free Will and Luck*) offers a compatibilist solution to the problem of manipulation which also serves as a solution to the problem of constitutive luck: he sets down historical conditions. When an agent acquires and sustains various character traits, such as the values and principle informing their actions, by a process that does not undermine their capacities to evaluate them critically, they become partially responsible for their traits and dispositions.¹⁷

Will this answer the problem of constitutive luck for the compatibilist? It might be protested that, if there is a problem of constitutive luck to be solved in the first place, an obvious regress worry arises: won't the very actions through which agents supposedly take responsibility for their traits and dispositions *import* this same constitutive luck? How Mele replies to this worry is revealed in his response to another pressing question that naturally arises in reflecting on MSL.

The more ardent incompatibilist will find MSL unsatisfying in that the loci of indeterminacy are in places in which the libertarian element in the

etiology of action is too modest. *Real* libertarians, it might be declared, want their indeterminacies placed at moments that play a more central role in the actions for which we are directly responsible. Mele recently has introduced a more daring libertarianism, designed to satisfy this demand. In Mele's Daring Soft Libertarianism (DSL), the indeterminacy is placed right at the time of the decision. DSL attempts to face down, rather than avoid, the problem of present luck. Thus, in response to the replay argument sketched above, in certain scenarios, Mele holds that the DSL should simply accept that there is luck in relevant outcomes – in directly free actions – but hold that the luck is not problematic. When is luck not problematic? When the relative undetermined probabilities of incompatible decisions (such as acting continently and resisting the extra shot of whiskey one desires, or instead acting weakly and taking it against one's own best judgment) are themselves shaped by an agent's own prior history. Sometimes agents determine the relative probabilities by prior self-shaping decisions, from processes of reflecting on and thereby learning from past experience, and so on. When the luck that is left is significantly influenced in this way, the agent can still decide freely and be morally responsible for these decisions.

Still, one can press Mele's proposed DSL. The undetermined free decisions later in an agent's life are themselves free on this view only due to the earlier probability shaping decisions giving rise to subsequent character traits. But what about these earlier decisions; were they not subject to luck? In other words, how does DSL avoid the regress problem mentioned above? Mele's responds by pointing out that agents start as little agents, neonates, and gradually become morally responsible as adults. The actions performed by little agents are generally not very morally significant, and therefore involve only very attenuated degrees of moral responsibility. When much smaller degrees of moral responsibility are at stake, the threshold for responsibility is much lower; hence little agents can be responsible, to that degree, despite the fact that the luck to which they are subject is not the product of previous (free) actions or decisions of theirs. We can be fully morally responsible agents because luck is constrained by our character traits, which were formed by ourselves when we were only somewhat morally responsible. Constitutive luck never gets washed out entirely. But just as DSL insists that some degree of present luck is compatible with moral responsibility, it can live with a small degree of constitutive luck. Once again, the parallels between Mele's compatibilist account of free agency and his libertarian account are worth noting: if this reply to the problem of constitutive luck succeeds, on behalf of DSL, it seems equally available to the compatibilist.

Mele's proposed resolution to the luck problem – which is to accept some sorts of luck rather than to try and expunge it – is the most direct and innovative attempt to solve the luck problem, and to unify thinking about both incompatibilist and compatibilist approaches. Its influence is

bound to be extensive. However, we register two reservations about it. First, Mele's claim that the threshold for holding little agents morally responsible is low because their actions are relatively trivial can be disputed. It might be claimed that the threshold for attributing responsibility to little agents is exactly the same, and that Mele is misled into thinking it is lower only because, since the stakes are relatively low, the costs of mistakes are small. In general, so this objection goes, the claim that the moral significance of an action affects the threshold for attributing responsibility to it is not compelling.

Second, Mele's project seems to us subject to an important qualification. Mele's proposal should be easier for the compatibilist to embrace than the incompatibilist. This is because, in Mele's defence of DSL, the actional machinery that does the character shaping work of increasing certain probabilities and reducing others is largely compatibilist machinery. It is an agent's ability to restrict certain probabilities by securing more or less effectively some traits over others that gives the agent the control over her character shaping that, according to Mele, should assuage the libertarian. But these agential abilities are bought mostly with compatibilist dollars, not libertarian ones. All the libertarian buys is the mere fact of the indeterministic occasions for luck, whereas the compatibilist machinery is what secures the stability of the character traits. Naturally, if the compatibilist is picking up the better part of the check, they will want to stake a greater claim on what is purchased. Nevertheless, assuming Mele is right, if the price is high, the good purchased is valuable: an adequate reply to the luck problem.¹⁸

8. Concluding Remarks

The six issues we have canvassed by no means exhaust the new developments that have unfolded in recent debates about free will and moral responsibility. As we noted in opening this piece, we have set aside any discussion of the Consequence Argument, which remains an area of intense controversy. Nor have we addressed the equally interesting recent developments surrounding debates about agent causation.¹⁹ Furthermore, we have said nothing of the threats to moral responsibility resulting from resultant or constitutive luck. For example, Michael Zimmerman ('Taking Luck Seriously') has recently argued that taking luck seriously and attempting to expunge its influence from our moral responsibility judgments has the (surprising) result of showing that the importance of the debate over Frankfurt's attack on PAP is very much overinflated. In fact, the importance of what we are morally responsible *for*, he contends, is overestimated. What is of deeper importance is that we are morally responsible, *tout court*, as he puts it, given what we might or might not have done had luck not played its part. Or consider instead recent discussions of situationism. Dana Nelkin has recently taken up the question as to whether there is

any stable notion of moral character at all that can be the subject for our theorizing about morally responsible agency. Contra Nelkin (2005), others, such as John Doris (2002), in his highly influential *Lack of Character*, offer startling evidence that the very assumption of a core moral agent who can have control and be the center of moral knowledge and justified belief is at bottom unsupportable. Furthermore, there remains a lively debate about the best way to advance a Strawsonian approach to moral responsibility as grounded in the reactive attitudes. Some, like Manuel Vargas ('Responsibility and the Aims of Theory'), argue that we have good reason to sustain large swaths of the reactive attitudes, but only at the price of revising our concept of moral responsibility, a view Vargas labels revisionism. Others, such as Erin Kelly, maintain that we cannot rationally support the aspects of the reactive attitudes that involve the retributive sentiments, which is similar to a point defended recently by Pereboom (*Living without Free Will*). By contrast, Paul Russell holds that the central Strawsonian position can be defended, when it is emphasized that there is a crucial sentimental aspect to responsibility – a moral sense – that requires an agent's capacity for relevant moral emotions.

Also unmentioned here is the large and rapidly growing empirical literature on free will-related topics. One focus of the new field of experimental philosophy has been folk intuitions about responsibility and freedom, especially the degree to which the folk are intuitively disposed towards incompatibilism (Nahmias et al.; Woolfolk, Doris, and Darley; Nichols and Knobe). Another focus of empirically engaged literature has been the extent to which findings in neuroscience and social and cognitive psychology threaten the exercise of free agency. In addition to the literature of situationism, there is a large literature – generally but not exclusively deflationary of scientific threats – concentrating on neuroscientific results on the timing of conscious intentions (Dennett, *Freedom Evolves*; Levy, 'Libet's Impossible Demand'; Mele, *Free Will and Luck*) and on experimental evidence for epiphenomenalism (Metzinger; Bayne). There is also a large and growing literature on abnormal psychological conditions, and whether, and how, these limit freedom. These debates have focused especially on addiction (Wallace, 'Addiction as Defect of the Will'; Watson, 'Excusing Addiction'; Yaffe, 'Recent Work on Addiction'; Levy, 'Autonomy and Addiction') and psychopathy (Haji, 'On Psychopaths and Culpability'; Greenspan; Levy, 'Responsibility of the Psychopath Revisited').

The above list is far from exhaustive. In any event, we believe that the topics that we have focused upon represent core controversies in this lively philosophical subfield. Throughout, we have attempted to retain a critical focus, and yet treat the controversies in a largely non-partisan fashion. We believe that we will have succeeded if we have provided even a partial road map for others who wish to advance these issues still further.

Acknowledgments

We would like to thank Randolph Clarke, Daniel Cohen, John Martin Fischer, Alfred Mele, Derk Pereboom, Manuel Vargas, and Brian Weatherson for thoughtful comments on an earlier draft.

Short Biographies

Neil Levy is a Principal Research Fellow at the Centre for Applied Philosophy and Public Ethics, University of Melbourne, and Director of Research at the Oxford Centre for Neuroethics. His research has two main thrusts: ethics from a naturalistic standpoint, especially neuroethics, and free will and moral responsibility, which is where he turns when he is missing his armchair. He holds two PhDs, both from Monash University (Australia). He is the author of five monographs, including, most recently, *Neuroethics* (Cambridge University Press, 2007). His work on free will and moral responsibility has appeared in *Philosophical Quarterly*, *Philosophical Studies*, *The Monist*, and the *Journal of Philosophy*, as well as in book chapters and other journals.

Michael McKenna writes primarily on the topics of free will and moral responsibility; he has authored or co-authored papers appearing in, among other places, *Journal of Philosophy*, *Philosophical Review*, *Nous*, *Philosophy and Phenomenological Research*, *Philosophical Studies*, *Midwest Studies in Philosophy*, *Philosophical Topics*, and *American Philosophical Quarterly*. He has also co-edited with David Widerker *Moral Responsibility and Alternative Possibilities* (Ashgate, 2003), and another forthcoming collection, co-edited with Paul Russell, *Free Will and Reactive Attitudes* (Ashgate, 2008). He is currently working on a book on the nature of moral responsibility, as well as another, under contract with Routledge Press, *A Contemporary Introduction to Free Will*. McKenna defends the thesis that both free will and moral responsibility are compatible with the metaphysical thesis of determinism, as well as with the more inclusive thesis of naturalism. McKenna is currently professor of Philosophy at Florida State University. Previously he was assistant and then associate professor and chair of the Department of Philosophy and Religion at Ithaca College. He has held visiting positions at Bryn Mawr College and University of Colorado, Boulder. He received his Ph.D. in philosophy from University of Virginia in 1993, and his B.A. from Thiel College in 1984.

Notes

* Correspondence Address: Neil Levy, Centre for Applied Philosophy and Public Ethics, University of Melbourne, Parkville 3010 Australia. Email: nlevy@unimelb.edu.au.

¹ For a lively debate over this controversy, see the essays in the collection edited by Widerker and McKenna.

² Alternatively, some might argue that even if Mele and Robb's example proves PAP false, there are other, more precise PAP-like principles that are *not* discredited by the example in question. For instance, it might be argued that an agent is morally responsible for performing a particular action, A, only if she could have avoided performing that very action, A. Perhaps the alternatives (allegedly) remaining in Mele and Robb's examples are adequate to satisfy this principle.

³ It should be noted that not all Frankfurt defenders agree. Hunt ('Moral Responsibility'), for instance, argues that even in relevant indeterministic settings, all alternative possibilities can be expunged.

⁴ For example, note the similarity of Pereboom's treatment of robustness (*Living without Free Will*) with McKenna's ('Robustness, Control, and the Demand'). Contrast these with Mele (*Free Will and Luck*). Fischer's formulations differ slightly from any of these (*Metaphysics of Free Will*; 'Responsibility and Alternative Possibilities'; *My Way*).

⁵ Among the many Source Compatibilists are Fischer (*Metaphysics of Free Will*); Fischer and Ravizza; Haji, *Moral Appraisability*; 'On Psychopaths and Culpability'; Frankfurt, 'Freedom of the Will'; McKenna, 'Responsibility'; 'Relationship'; Wallace, *Responsibility and the Moral Sentiments*. Mele articulates but remains agnostic about such a view (*Autonomous Agents*; *Free Will and Luck*). Watson's ('Free Agency') position is consistent with such a view, and Wolf ('Asymmetrical Freedom'; *Freedom within Reason*) holds a hybrid view in which she agrees with the Source Compatibilist regarding praiseworthiness. (It should be noted that Wallace comes to the same conclusion as Frankfurt regarding PAP, but by a different route.) Dennett is hard to pin down. Earlier (*Elbow Room*) he seemed to be in the Source Compatibilist camp. But more recently (*Freedom Evolves*), it seems he has withdrawn his commitment to a source theorist approach.

Among the growing number of Source Incompatibilists are Hunt ('Moral Responsibility'); Pereboom ('Determinism *Al Dente*'; *Living without Free Will*); Stump ('Libertarian Freedom'); Zagzebski ('Does Libertarian Freedom Require Alternate Possibilities?'). Mele articulates but remains agnostic about such a view (*Autonomous Agents*; 'Soft Libertarianism'; *Free Will and Luck*). Recently, Widerker seems to have joined their ranks ('Libertarianism').

⁶ The term 'Leeway' was first introduced by Pereboom ('Alternate Possibilities'). Among the recent Leeway Compatibilists, whose ranks had dwindled but are again growing, one finds Berofsky, 'Ifs, Cans, and Free Will'; 'Classical Compatibilism'; Campbell, 'Compatibilist Theory'; Kapitan, 'Master Argument'; Lewis; M. Smith; Vihvelin, 'Free Will Demystified'. Wolf ('Asymmetrical Freedom'; *Freedom within Reason*) holds a hybrid view in which she agrees with the Leeway Compatibilist regarding blameworthiness. (Of course, traditionally, Leeway Compatibilists have a long list of contributors in the history of philosophy from Hume through the likes of Ayer, Hobart, Moore, Schlick, and Smart.)

Among the many recent Leeway Incompatibilists are Chisholm; Kane, *Significance of Free Will*; Ginet, *On Action*; O'Connor; Rowe; R. Taylor; van Inwagen, *Essay on Free Will*; Warfield, 'Causal Determinism'; Wiggins. (Traditionally, Leeway Incompatibilists have a long list of contributors in the history of philosophy from Kant through the likes of C. A. Campbell and Paul Edwards.)

⁷ Other compatibilists, such as the attributivists (discussed in section 6 below) seem committed to a source theorist approach, though perhaps some would resist.

⁸ These three arguments are distinguished and discussed in McKenna, 'Source Incompatibilism'.

⁹ So as to avoid any misunderstanding, a cautionary note is in order: Many of the authors identified here are not Source Incompatibilists but rather Leeway Incompatibilists. But this only shows that these clear-sighted incompatibilists recognize that there might be paths to an incompatibilist conclusion other than by way of showing that determinism is incompatible with the ability to do otherwise. Hence, they offer argumentative resources to those other incompatibilists who disagree with them about the soundness of Frankfurt's argument.

¹⁰ 'Global' and 'local' in these contexts are comparative notions. The idea is that local manipulation proceeds by manipulating only a relatively small range of an agent's mental states (for example, by electronically stimulating an agent's brain in order to add an extra desire into the mix of an agent's motivational state). Global manipulation is, by contrast, a matter of massively revising an agent's psychological constitution (but typically not so much that there is a change in identity). An example might be the real life case of Patty Hearst, who was kidnapped and brainwashed into joining a revolutionary movement. She then came to participate in bank robberies and other activities far removed from anything that one would have expected from her pre-brainwashed self.

¹¹ Some adopt the language of ‘internalism’ and ‘externalism’. But we find the terminology misleading. The natural association of externalism in this context concerns the claim that, for instance, mental states include conditions that reach beyond the agent. (Think of Putnam’s claim that ‘meaning ain’t all in the head’.) Internalist approaches do not require this. The distinction between historical versus nonhistorical concepts does not map on well here. An exercise of morally responsible agency might require an extended stretch of time, and yet happen only internally to the agent; hence it might be both internalist (in the sense we are using it) and historical. Or it might require only current timeslice features, and thus be nonhistorical, but still be externalist in that it involves facts obtaining, so to speak, outside the skin of the agent.

¹² There might be some worries as to how Pereboom can pull this off in Case 1. If the causal connections between the agent’s mental states arise externally, then it seems that there is not the proper sort of causal integration required for agency. If Case 1 is to be saved, it must face this challenge. The most likely prospect seems to be by way of the manipulation producing the same neural realizers as are produced in the normal way. These realizers should be designed to sustain as much as possible the internal causal integration present in normal cases of agency (McKenna, ‘Hard-Line Reply’; Pereboom, ‘Hard-Line Reply’). But perhaps this is not a coherent possibility (Demetriou). If not, it seems that Pereboom’s Case 1 fails to satisfy the No Relevant Difference Premise. (We are indebted to Brain Weatherson for help on this point.)

¹³ The dialectic at this point is delicate, and perhaps the preceding paragraph moves along a bit too quickly. It might be a simple matter for Pereboom to build in the historical stuff. But the question is, once he has done that, is it so clear that such an agent does not act freely and is not morally responsible? Regardless, if it is not clear, if it becomes open to the compatibilist to maintain that an agent manipulated in this way *is* free and responsible, then our basic point remains: the compatibilist will have to resist Pereboom here by way of the Manipulation Premise. (We owe Alfred Mele for helping us clarify our position here.)

¹⁴ Sher (*In Praise of Blame*) defends an interesting related view. He argues that responsibility is a causal notion, and that we are therefore responsible only for what we control. But *blame*, he claims, is not causal; hence agents can be blameworthy for traits (at least) for which they are not responsible. His ‘Out of Control’ seems to expand the scope of blameworthiness to actions which are not controlled.

¹⁵ To avoid any misunderstanding, we do not claim that Watson would interpret the case as we do here. We only call upon the example as a familiar one that could be understood as we shall present it.

¹⁶ These philosophers are defenders of the unique *adequacy* of agent-causal libertarianism, not of its *actuality*. Pereboom (*Living without Free Will*) believes that though agent-causal libertarianism is the correct account of free will, it is not a feature of the actual world, and Clarke (*Libertarian Accounts of Free Will*) has grave doubts about even its possibility. O’Connor is convinced that human beings are agent-causes.

¹⁷ For an interesting critical discussion of Mele’s historical commitments, see Vargas, ‘On the Importance of History’.

¹⁸ One way to reinforce the point is to see that the libertarian might, with some justice, accuse Mele of changing the subject. Libertarians demand indeterminism because they hold that it is a necessary condition of free will and/or moral responsibility. Since on Mele’s DSL account indeterminism is seen as something that might rationally be preferred by agents, but which is *not* necessary for free will or moral responsibility, Mele is committed to denying this, and therefore to denying the very heart of the libertarian claim. He *makes room for* indeterminism, but in a way and a place that is orthogonal to the central issues. Hence all the work in ensuring that agents are free and responsible is done by Mele’s compatibilist conditions. (We are thankful to Manuel Vargas for suggesting this way of expressing our concern.)

¹⁹ For instance, O’Connor offered a thorough defense of this libertarian strategy, and then Clarke (*Libertarian Accounts of Free Will*) offered yet another, distinct examination of the best case for agent causation. Pereboom (*Living without Free Will*; ‘Is Our Conception of Agent Causation Coherent?’) defends agent causation in so far as he wishes to show that it is at least a coherent possibility (though one we have good empirical reason to reject at our world). In response, various objections to agent causation have been put forward, such as those by Ginet (‘Freedom’), Mele (*Free Will and Luck*), and Haji (‘Dialectical Delicacies’).

Works Cited

- Arpaly, Nomy. *Merit, Meaning and Human Bondage: An Essay on Free Will*. Princeton, NJ: Princeton UP, 2006.
- . *Unprincipled Virtue*. New York, NY: Oxford, UP, 2003.
- Bayne, Tim. 'Phenomenology and the Feeling of Doing: Wegner on the Conscious Will'. *Does Consciousness Cause Behavior? An Investigation of the Nature of Volition*. Eds. S. Pockett, W. P. Banks, and S. Gallagher. Cambridge, MA: MIT Press, 2006. 169–86.
- Berofsky, Bernard. 'Classical Compatibilism: Not Dead Yet'. *Moral Responsibility and Alternative Possibilities*. Eds. David Widerker and Michael McKenna. Aldershot: Ashgate Press, 2003. 107–26.
- . 'Ifs, Cans, and Free Will: The Issues'. *The Oxford Handbook of Free Will*. Ed. Robert Kane. Oxford and New York, NY: Oxford UP, 2002. 181–201.
- Buss, Sarah and Lee Overton. *Contours of Agency: Essays on Themes from Harry Frankfurt*. Cambridge, MA: MIT Press, 2002.
- Campbell, Joseph. 'A Compatibilist Theory of Alternative Possibilities'. *Philosophical Studies* 88 (1997): 319–30.
- . 'Free Will and the Necessity of the Past'. *Analysis* 67 (2007): 105–11.
- Campbell, Joseph Keim, Michael O'Rourke, and David Sheir, eds. *Freedom and Determinism*. Cambridge, MA: MIT Press, 2004.
- Chisholm, Roderick. 'He Could Have Done Otherwise'. *Journal of Philosophy* 64 (1967): 409–18.
- Clarke, Randolph. 'Agent Causation and the Problem of Luck'. *Pacific Philosophical Quarterly* 86 (2005): 408–21.
- . *Libertarian Accounts of Free Will*. New York, NY: Oxford UP, 2003.
- . 'On an Argument for the Impossibility of Moral Responsibility'. *Midwest Studies in Philosophy* 29 (2005): 13–24.
- Cohen, Daniel and Toby Handfield. 'Finking Frankfurt'. *Philosophical Studies* 135 (2007): 363–74.
- Copp, David. 'Defending the Principle of Alternative Possibilities: Blameworthiness and Moral Responsibility'. *Nous* 31 (1997): 441–56.
- . "'Ought" Implies "Can", Blameworthiness, and the Principle of Alternate Possibilities'. *Moral Responsibility and Alternative Possibilities*. Eds. David Widerker and Michael McKenna. Aldershot: Ashgate Press, 2003. 265–300.
- Demetriou, Kristin. 'The Soft-Line Solution to Pereboom's Four-Case Argument'. 2008. Unpublished manuscript.
- Dennett, Daniel. *Elbow Room: Varieties of Free Will Worth Wanting*. Cambridge, MA: MIT Press, 1984.
- . *Freedom Evolves*. London: Penguin Books, 2003.
- Doris, John. *Lack of Character*. Cambridge, MA.: Cambridge UP, 2002.
- Fischer, John Martin. 'Frankfurt Style Compatibilism'. *Contours of Agency: Essays on Themes from Harry Frankfurt*. Eds. Sarah Buss and Lee Overton. Cambridge, MA: MIT Press, 2002. 1–26.
- . *The Metaphysics of Free Will*. Oxford: Blackwell Publishers, 1994.
- . *My Way*. New York, NY: Oxford UP, 2006.
- . 'Recent Work on Moral Responsibility'. *Ethics* 110 (1999): 93–139.
- . 'Responsibility and Alternative Possibilities'. *Moral Responsibility and Alternative Possibilities*. Eds. David Widerker and Michael McKenna. Aldershot: Ashgate Press, 2003. 37–52.
- . 'Responsibility and Control'. *Journal of Philosophy* 89 (1982): 24–40.
- . 'The Transfer of Nonresponsibility'. *Freedom and Determinism*. Eds. Joseph Keim Campbell, Michael O'Rourke, and David Sheir. Cambridge, MA: MIT Press, 2004. 189–209.
- and Mark Ravizza. *Responsibility and Control: An Essay on Moral Responsibility*. Cambridge: Cambridge UP, 1998.
- Frankfurt, Harry. 'Alternate Possibilities and Moral Responsibility'. *Journal of Philosophy* 66 (1969): 829–39. Reprinted in *Moral Responsibility*. Ed. John Martin Fischer. Ithaca, NY: Cornell UP, 1986. 143–52; and *The Importance of What We Care About*. Ed. Harry Frankfurt. Cambridge, UK: Cambridge UP, 1988. 1–10.
- . 'Freedom of the Will and the Concept of a Person'. *Journal of Philosophy* 68 (1971): 5–20. Reprinted in *Moral Responsibility*. Ed. John Martin Fischer. Ithaca, NY: Cornell UP,

- 1986; *The Importance of What We Care About*. Ed. Harry Frankfurt. Cambridge, UK: Cambridge UP, 1988. 65–80; and *Free Will*. Ed. Gary Watson. New York, NY: Oxford UP, 1982. 81–95.
- . ‘Some Thoughts Concerning PAP’. *Moral Responsibility and Alternative Possibilities*. Eds. David Widerker and Michael McKenna. Aldershot: Ashgate Press, 2003. 339–48.
- . ‘Three Concepts of Free Action II’. *Proceedings of the Aristotelian Society* suppl. 49 (1975): 113–25.
- Ginet, Carl. ‘Freedom, Responsibility, and Agency’. *Journal of Ethics* 1 (1997): 85–98.
- . ‘In Defense of the Principle of Alternative Possibilities: Why I Don’t Find Frankfurt’s Argument Convincing’. *Philosophical Perspectives* 10 (1996): 403–17. Reprinted with an added addendum in *Moral Responsibility and Alternative Possibilities*. Eds. David Widerker and Michael McKenna. Aldershot: Ashgate Press, 2003. 75–90.
- . ‘Libertarianism’. *Oxford Handbook of Metaphysics*. Eds. Michael Loux and Dean Zimmerman. Oxford: Oxford UP, 2003. 587–612.
- . ‘Might We Have No Choice?’ *Freedom and Determinism*. Ed. Keith Lehrer. New York, NY: Random House, 1966. 87–104.
- . *On Action*. Cambridge: Cambridge UP, 1990.
- Goetz, Stewart. ‘Frankfurt-Style Compatibilism and Begging the Question’. *Midwest Studies in Philosophy* 29 (2005): 83–105.
- Greenspan, Patricia S. ‘Responsible Psychopaths’. *Philosophical Psychology* 16 (2003): 417–29.
- Haji, Ishtiyaque. ‘Active Control, Agent Causation, and Free Action’. *Philosophical Explorations* 7 (2004): 131–48.
- . ‘Alternative Possibilities, Moral Obligation, and Moral Responsibility’. *Philosophical Papers* 22 (1993): 41–50.
- . ‘Compatibilist Views of Freedom and Responsibility’. *The Oxford Handbook of Free Will*. Ed. Robert Kane. Oxford and New York, NY: Oxford UP, 2002. 202–28.
- . *Moral Appraisability*. New York, NY: Oxford UP, 1998.
- . ‘On Psychopaths and Culpability’. *Law and Philosophy* 17 (1998): 117–40.
- and Michael McKenna. ‘Defending Frankfurt’s Argument in Deterministic Contexts: A Reply to Palmer’. *Journal of Philosophy* 103 (2006): 363–72.
- . ‘Dialectical Delicacies in the Debate about Freedom and Alternative Possibilities’. *Journal of Philosophy* 101 (2004): 299–314.
- Howard-Snyder, Daniel and Jeff Jordan, eds. *Faith, Freedom, and Rationality*. Lanham, MD: Rowman and Littlefield, 1996.
- Hunt, David P. ‘Moral Responsibility and Buffered Alternatives’. *Midwest Studies in Philosophy* 29 (2005): 126–45.
- . ‘Moral Responsibility and Unavoidable Action’. *Philosophical Studies* 97 (2000): 195–227.
- van Inwagen, Peter. *An Essay on Free Will*. Oxford: Clarendon Press, 1983.
- . ‘Free Will Remains a Mystery’. *Philosophical Perspectives* 14 (2000): 1–19.
- . ‘The Incompatibility of Free Will and Determinism’. *Philosophical Studies* 27 (1975): 185–99.
- . ‘The Incompatibility of Responsibility and Determinism’. Eds. M. Brady and M. Brand. *Bowling Green Studies in Applied Philosophy* 2 (1980): 30–7.
- Kane, Robert, ed. *The Oxford Handbook of Free Will*. Oxford and New York, NY: Oxford UP, 2002.
- . ‘Responsibility, Indeterminism, and Frankfurt-style Cases’. *Moral Responsibility and Alternative Possibilities*. Eds. David Widerker and Michael McKenna. Aldershot: Ashgate Press, 2003. 91–106.
- . ‘Responsibility, Luck, and Chance: Reflections on Free Will and Indeterminism’. *Journal of Philosophy* 96 (1999): 217–40.
- . *The Significance of Free Will*. Oxford: Oxford UP, 1996.
- Kapitan, Tomis. ‘Autonomy and Manipulated Freedom’. *Philosophical Perspectives* 14 (2000): 81–104.
- . ‘A Master Argument for Incompatibilism?’ *The Oxford Handbook of Free Will*. Ed. Robert Kane. New York: Oxford UP, 2002. 127–57.
- Kelly, Erin. ‘Doing Without Desert’. *Pacific Philosophical Quarterly* 83 (2002): 180–25.

- Levy, Neil. 'Agents and Mechanisms: Fischer's Way'. *Philosophical Quarterly* 57 (2007): 123–30.
- . 'Autonomy and Addiction'. *Canadian Journal of Philosophy* 36 (2006): 427–48.
- . 'Counterfactual Intervention and Agents' Capacities'. *Journal of Philosophy* 105 (2008): 223–39.
- . 'The Good, the Bad and the Blameworthy'. *Journal of Ethics and Social Philosophy* 1 (2005): 1–16.
- . 'Libet's Impossible Demand'. *Journal of Consciousness Studies* 12 (2005): 67–76.
- . 'The Responsibility of the Psychopath Revisited'. *Philosophy, Psychiatry & Psychology* 14 (2007): 129–38.
- Lewis, David. 'Are We Free to Break the Laws?' *Theoria* 47 (1981): 113–21.
- Locke, Don. 'Three Concepts of Free Action I'. *Proceedings of Aristotelian Society* suppl. 49 (1975): 95–112.
- McKay, Thomas and David Johnson. 'A Reconsideration of an Argument against Incompatibilism'. *Philosophical Topics* 24 (1996): 113–22.
- McKenna, Michael. 'A Hard-Line Reply to Pereboom's Four-Case Manipulation Argument'. *Philosophy and Phenomenological Research* 77 (2008): 142–59.
- . 'The Relationship between Autonomous and Morally Responsible Agency'. *Personal Autonomy*. Ed. James Stacey Taylor. New York, NY: Cambridge UP, 2004.
- . 'Responsibility and Globally Manipulated Agents'. *Philosophical Topics* 32 (2004): 169–82.
- . 'Robustness, Control, and the Demand for Morally Significant Alternatives'. *Moral Responsibility and Alternative Possibilities*. Eds. David Widerker and Michael McKenna. Aldershot: Ashgate Press, 2003. 201–17.
- . 'Saying Goodbye to the Direct Argument the Right Way'. *Philosophical Review* 117 (2008): 349–83.
- . 'Source Incompatibilism, Ultimacy, and the Transfer of Non-responsibility'. *American Philosophical Quarterly* 38 (2001): 37–51.
- Mele, Alfred. *Autonomous Agents*. New York, NY: Oxford UP, 1995.
- . 'A Critique of Pereboom's Four-Case Argument for Incompatibilism'. *Analysis* 65 (2005): 75–80.
- . *Free Will and Luck*. New York, NY: Oxford UP, 2006.
- . 'Soft Libertarianism and Frankfurt-Style Scenarios'. *Philosophical Topics* 24 (1996): 123–41.
- . 'Ultimate Responsibility and Dumb Luck'. *Social Philosophy and Policy* 16 (1999): 274–93.
- and David Robb. 'Bbs, Magnets and Seesaws: The Metaphysics of Frankfurt-style Cases'. *Moral Responsibility and Alternative Possibilities*. Eds. David Widerker and Michael McKenna. Aldershot: Ashgate Press, 2003. 127–38.
- . 'Rescuing Frankfurt-Style Cases'. *Philosophical Review* 107 (1998): 97–112.
- Metzinger, Thomas. 'Inferences are Just Folk Psychology'. *Behavioral and Brain Sciences* 27 (2004): 670.
- Nahmias, Eddy, et al. 'Is Incompatibilism Intuitive?' *Philosophy and Phenomenological Research* 73 (2007): 28–53.
- Nelkin, Dana. 'Freedom, Responsibility, and the Challenge of Situationism'. *Midwest Studies in Philosophy* 29 (2005): 181–206.
- Nichols, S. and J. Knobe. 'Moral Responsibility and Determinism: The Cognitive Science of Folk Intuitions'. *Nous* 41 (2007): 663–85.
- Nowell-Smith, P. 'Free Will and Moral Responsibility'. *Mind* 55 (1948): 45–61.
- O'Connor, Timothy. *Persons and Causes*. New York, NY: Oxford UP, 2000.
- Palmer, David. 'New Distinctions, Same Troubles: A Reply to Haji and McKenna'. *Journal of Philosophy* 102 (2005): 474–82.
- Pereboom, Derk. 'Alternate Possibilities and Causal Histories'. *Philosophical Perspectives* 14 (2000): 119–38.
- . 'Defending Hard Incompatibilism'. *Midwest Studies in Philosophy* 29 (2005): 228–47.
- . 'Determinism *Al Dente*'. *Nous* 29 (1995): 21–45.
- . 'The Hard-Line Reply to the Multiple-Case Manipulation Argument'. *Philosophy and Phenomenological Research* 77 (2008): 160–70.
- . 'Is Our Conception of Agent Causation Coherent?'. *Philosophical Topics* 32 (2004): 275–85.

- . *Living without Free Will*. Cambridge: Cambridge UP, 2001.
- . 'Reasons Responsiveness, Alternative Possibilities, and Manipulation Arguments against Compatibilism: Reflections on John Martin Fischer's *My Way*'. *Philosophical Books* 47 (2006): 198–212.
- Ravizza, Mark. 'Semicompatibilism and the Transfer of Non-Responsibility'. *Philosophical Studies* 75 (1994): 61–94.
- Rowe, William. *Thomas Reid on Freedom and Morality*. Ithaca, NY: Cornell UP, 1991.
- Russell, Paul. 'Responsibility and the Condition of Moral Sense'. *Philosophical Topics* 32 (2004): 287–306.
- Scanlon, T. M. *What We Owe to Each Other*. Cambridge, MA: Harvard UP, 1998.
- Schoeman, Ferdinand, ed. *Responsibility, Character, and the Emotions: New Essays in Moral Psychology*. Cambridge: Cambridge UP, 1987.
- Sher, George. In *Praise of Blame*. New York, NY: Oxford UP, 2006.
- . 'Out of Control'. *Ethics* 116 (2006): 285–301.
- Smart, J. J. C. 'Free-Will, Praise and Blame'. *Mind* 70 (1961): 291–306.
- Smilansky, Saul. *Free Will and Illusion*. Oxford: Clarendon Press, 2000.
- Smith, Angela. 'Control, Responsibility, and Moral Assessment'. *Philosophical Studies* 138 (2008): 367–92.
- . 'Responsibility for Attitudes: Activity and Passivity in Mental Life'. *Ethics* 115 (2005): 236–71.
- Smith, Michael. 'Rational Capacities, or: How to Distinguish Recklessness, Weakness, and Compulsion'. *Weakness of Will and Practical Irrationality*. Eds. Sarah Stroud and Christine Tappolet. New York, NY: Oxford UP, 2003. 17–38.
- Speak, Dan. 'The Impertinence of Frankfurt-Style Argument'. *Philosophical Quarterly* 57 (2007): 76–95.
- Strawson, Galen. *Freedom and Belief*. Oxford: Clarendon Press, 1986.
- . 'The Unhelpfulness of Indeterminism'. *Philosophy and Phenomenological Research* 60 (2000): 149–56.
- Strawson, P. F. 'Freedom and Resentment'. *Proceedings of the British Academy* 48 (1962): 187–211.
- Stroud, Sarah and Christine Tappolet, eds. *Weakness of Will and Practical Irrationality*. New York, NY: Oxford UP, 2003.
- Stump, Eleonore. 'The Direct Argument for Incompatibilism'. *Philosophy & Phenomenological Research* 61 (2000): 459–66.
- . 'Libertarian Freedom and the Principle of Alternative Possibilities'. *Faith, Freedom, and Rationality*. Eds. Daniel Howard-Snyder and Jeff Jordan. Lanham, MD: Rowman and Littlefield, 1996. 73–88.
- Taylor, James Stacey, ed. *Personal Autonomy*. New York, NY: Cambridge UP, 2004.
- Taylor, Richard. *Metaphysics*. Englewood Cliffs, NJ: Prentice Hall, 1974.
- Vargas, Manuel. 'On the Importance of History for Responsible Agency'. *Philosophical Studies* 127 (2006): 351–82.
- . 'Responsibility and the Aims of Theory: Strawson and Revisionism'. *Pacific Philosophical Quarterly* 85 (2004): 218–41.
- . 'The Trouble with Tracing'. *Midwest Studies in Philosophy* 29 (2005): 269–91.
- Vihvelin, Kadri. 'Compatibilism, Incompatibilism, and Impossibilism'. *Contemporary Debates in Metaphysics*. Eds. John Hawthorne, Theodore Sider and Dean Zimmerman. Malden, MA: Blackwell, 2007. 303–18.
- . 'Free Will Demystified: A Dispositional Account'. *Philosophical Topics* 32 (2004): 427–50.
- . 'Freedom, Foreknowledge, and the Principle of Alternate Possibilities'. *Canadian Journal of Philosophy* 30 (2000): 1–24.
- Wallace, R. Jay. 'Addiction as Defect of the Will: Some Philosophical Reflections'. *Law and Philosophy* 18 (1999): 621–54.
- . *Responsibility and the Moral Sentiments*. Cambridge, MA: Harvard UP, 1994.
- Warfield, Ted. 'Causal Determinism and Human Freedom are Incompatible'. *Philosophical Perspectives* 14 (2000): 167–80.
- . 'Determinism and Moral Responsibility are Incompatible'. *Philosophical Topics* 24 (1996): 215–26.

- Watson, Gary. 'Excusing Addiction'. *Law and Philosophy* 18 (1999): 589–619.
- . 'Free Agency'. *Journal of Philosophy* 72 (1975): 205–20.
- . 'Responsibility and the Limits of Evil: Variations on a Strawsonian Theme'. *Responsibility, Character, and the Emotions: New Essays in Moral Psychology*. Ed. Ferdinand Schoeman. Cambridge: Cambridge UP, 1987. 256–86.
- . 'Two Faces of Responsibility'. *Philosophical Topics* 24 (1996): 227–48.
- Widerker, David. 'Blameworthiness, and Frankfurt's Argument against the Principle of Alternative Possibilities'. *Moral Responsibility and Alternative Possibilities*. Eds. David Widerker and Michael McKenna. Aldershot: Ashgate Press, 2003. 53–74.
- . 'Farewell to the Direct Argument'. *Journal of Philosophy* 99 (2002): 316–24.
- . 'Frankfurt on "Ought Implies Can" and Alternative Possibilities'. *Analysis* 51 (1991): 222–4.
- . 'Frankfurt's Attack on Alternative Possibilities: A Further Look'. *Philosophical Perspectives* 14 (2000): 181–201.
- . 'Libertarianism and Frankfurt's Attack on the Principle of Alternative Possibilities'. *Philosophical Review* 104 (1995): 247–61.
- . 'Libertarianism and the Philosophical Significance of Frankfurt Scenarios'. *Journal of Philosophy* 103 (2006): 163–87.
- and Michael McKenna, eds. *Moral Responsibility and Alternative Possibilities*. Aldershot: Ashgate Press, 2003.
- Wiggins, David. 'Towards a Reasonable Libertarianism'. *Essays on Freedom and Action*. Ed. T. Honderich. London: Routledge & Kegan Paul, 1973. 31–62.
- Wolf, Susan. 'Asymmetrical Freedom'. *Journal of Philosophy* 77 (1980): 157–66.
- . *Freedom within Reason*. Oxford: Oxford UP, 1990.
- Woolfolk, R., J. Doris, and J. Darley. 'Identification, Situational Constraint, and Social Cognition: Studies in the Attribution of Moral Responsibility Cognition'. *Cognition* 100 (2006): 283–301.
- Yaffe, Gideon. '"Ought" Implies "Can" and the Principle of Alternate Possibilities'. *Analysis* 59 (1999): 218–22.
- . 'Recent Work on Addiction and Responsible Agency'. *Philosophy and Public Affairs* 30 (2002): 178–221.
- Zagzebski, Linda. 'Does Libertarian Freedom Require Alternate Possibilities?' *Philosophical Perspectives* 14 (2000): 231–48.
- Zimmerman, Michael J. 'The Moral Significance of Alternate Possibilities'. *Moral Responsibility and Alternative Possibilities*. Eds. David Widerker and Michael McKenna. Aldershot: Ashgate Press, 2003. 301–26.
- . 'Taking Luck Seriously'. *Journal of Philosophy* 99 (2002): 553–76.