

The Manipulation Argument, At the Very Least, Undermines Classical Compatibilism

Abstract The compatibility of determinism and the ability to do otherwise has been implicitly assumed by many to be irrelevant to the viability of compatibilist responses to the manipulation argument for incompatibilism. I argue that this assumption is mistaken. The manipulation argument may be unsound. But even so, the manipulation argument, at the very least, undermines classical compatibilism, the view that free will requires the ability to do otherwise, and having that ability is compatible with determinism. This is because classical compatibilism, in conjunction with any type of reply to the manipulation argument, has counterintuitive implications. In order to avoid such implications, we need not hold that determinism is incompatible with moral responsibility. But we must hold that determinism is incompatible with the ability to do otherwise.

Keywords Manipulation · Determinism · Compatibilism · Alternative possibilities · Responsibility · Frankfurt-style cases · Derk Pereboom · Kadri Vihvelin

Introduction

Classical compatibilism is the view that free will requires the ability to do otherwise, and having that ability is compatible with determinism.¹ This view is often paired with the principle of alternative possibilities (PAP), the view that an agent is morally responsible for what she has done only if she could have done otherwise. By contrast, according to source compatibilism—a view that is often motivated by Frankfurt-style cases (Frankfurt 1969)—an action that is free and for which one is morally responsible is one that issues from oneself in an appropriate way which in turn does *not* require the ability to do otherwise.² So, unlike classical compatibilism, subscribing to source compatibilism does not commit one to the compatibility of determinism and the ability to do otherwise.

Two prominent arguments that have been employed in favor of the incompatibility of determinism and moral responsibility are the consequence argument and the manipulation argument. However, the consequence argument (at best) establishes only that

¹ For a defense of classical compatibilism, see Saunders (1968), Lehrer (1976), Horgan (1979), Lewis (1981), Campbell (1997), Smith (2003), Fara (2008), Perry (2008), Kapitan (2011), Berofsky (2012), and Vihvelin (2013). Wolf (1990) and Nelkin (2011) do not hold that free will requires the ability to do otherwise. But they do hold that an agent is blameworthy for what she has done only if she could have done otherwise, and, crucially, having the ability to do otherwise is compatible with determinism. So Wolf and Nelkin are likewise subject to the challenge of this paper.

² For a defense of source compatibilism, see Frankfurt (1971), Watson (1975), Wallace (1994), McKenna (2005), and Sartorio (2011). Fischer and Ravizza (1998) who are semi-compatibilists similarly reject PAP and maintain that determinism is compatible with moral responsibility.

determinism is incompatible with the ability to do otherwise (Ginet 1966; van Inwagen 1983). The consequence argument *by itself* does not show that an action that is free and for which one is morally responsible requires the ability to do otherwise. Thus, source compatibilists and semi-compatibilists are, as Fischer (2010: 316) puts it, thereby able to *side-step* the consequence argument.

In contrast to the consequence argument, the manipulation argument has been perceived by many to be an equally serious challenge to all forms of compatibilism. More specifically, the compatibility of determinism and the ability to do otherwise has been implicitly assumed to be irrelevant to the viability of compatibilist responses to the manipulation argument. I argue that this assumption is mistaken. The manipulation argument may be unsound. But even so, the manipulation argument, at the very least, undermines classical compatibilism. This is because classical compatibilism, in conjunction with any type of reply to the manipulation argument, has counterintuitive implications. In order to avoid such implications, we need not hold that determinism is incompatible with moral responsibility. But we must hold that determinism is incompatible with the ability to do otherwise.

The paper is structured as follows. In section 1, I discuss Derk Pereboom's (2014) four-case manipulation argument (FCMA) which will take the center stage of this paper. In section 2, I summarize Kadri Vihvelin's (2013) version of classical compatibilism. While my argument is aimed at classical compatibilism *tout court*, it is nevertheless beneficial to focus on Vihvelin's view since it is a recent, well-developed defense of classical compatibilism. After discussing Vihvelin's view, I will then be in a position to present my distinctive case against classical compatibilism. Accordingly, in section 3 I consider the counterintuitive implications of the conjunction of classical compatibilism and a so-called soft-line reply to the FCMA. In section 4 I consider the counterintuitive implications of the conjunction of classical compatibilism and a so-called hard-line reply to the FCMA. As we will see, no matter how the source compatibilist replies to the FCMA, she need not be committed to any such counterintuitive implications.

I will now proceed to discuss Pereboom's latest defense of his FCMA. I will not attempt to cover every aspect of the argument. Rather, I intend to flesh out aspects of the FCMA that bear directly upon my distinctive case against classical compatibilism.³

1. Pereboom's Four-case Manipulation Argument

Pereboom (2014: ch. 4) presents four cases in which Professor Plum is causally determined to kill White. In each case, Plum satisfies a medley of compatibilist

³ As will become evident, I largely ignore *Cases 2* and *3*, and instead focus on comparing *Case 1* with *Case 4*. The reason why I still focus on Pereboom's argument rather than, e.g., Mele's (2006) manipulation argument which only compares two cases is that the kind of direct manipulation present in *Case 1* better serves my aim of drawing out certain implications from classical compatibilism.

conditions for responsibility (henceforth the ‘*common compatibilist conditions*’).⁴ However, the way in which Plum is causally determined to kill White varies in each case. In *Case 1* Plum kills White due to Plum’s neural states being manipulated by neuroscientists through radio-like technology. In *Case 2* Plum kills White due to Plum’s being programmed by neuroscientists at the beginning of his life. In *Case 3* Plum kills White due to Plum’s undergoing certain training practices within a community. In *Case 4* Plum kills White simply due to his strongly egoistic but reasons-responsive process of deliberation.

Pereboom claims the following: Plum₁ (of *Case 1*) is not blameworthy for deciding to kill White. So satisfying the *common compatibilist conditions* for responsibility is not sufficient for being responsible for one’s action.⁵ Moreover, there is no relevant difference among *Cases 1–4* with respect to Plum’s responsibility. So Plum₄ (of *Case 4*) is likewise not blameworthy for deciding to kill White. The best explanation for Plum’s blamelessness in *Cases 1–4* is that Plum is causally determined by factors beyond his control to decide to kill White (Pereboom 2014: 79). So compatibilism is false. We can formulate the argument as follows:

1. Plum₁ is not blameworthy for deciding to kill White.
2. There is no relevant difference among *Cases 1–4* with respect to Plum’s responsibility for deciding to kill White.
3. So, Plum is not blameworthy for deciding to kill White in *Cases 1–4*.
4. So, compatibilism is false.

Pereboom’s (2001) formulation of *Case 1* involved Plum₁’s neural states being induced from moment to moment by neuroscientists. This gave rise to the concern that Plum₁ lacks certain conditions for being an agent which may be understood as a *pre-condition* for satisfying the *common compatibilist conditions* for responsibility (Fischer 2004: 156; Mele 2005: 78; Baker 2006: 320; Demetriou 2010). In light of this concern, Pereboom (2014: 76) amends *Case 1* in such a way that the neuroscientists only exert a momentary egoism-enhancing influence upon Plum₁. Surely agency can be preserved in the face of momentary external influences (Shabo 2010: 75–77). So the compatibilist must, it seems, concede that Plum₁ is an agent in the new *Case 1*.⁶ If, however, Plum₁ satisfies this pre-condition for satisfying the *common compatibilist conditions* for responsibility, then it should be safe to stipulate that Plum₁ *does* satisfy the *common compatibilist conditions* for responsibility, as I will now explain.

According to the new FCMA, the neural realization of Plum’s reasoning process and of his decision is exactly the same in *Cases 1–4* (Pereboom 2014: 76–79). Exact

⁴ Some of these conditions include, but are not limited to, having the appropriate set of higher-order desires (Frankfurt 1971), being reasons-responsive (Fischer and Ravizza 1998), having the ability to regulate one’s actions by moral reasons (Wallace 1994), and having the ability to develop one’s moral character over time (Mele 1995; Haji 1998).

⁵ Throughout the paper, I ignore the issue of satisfying epistemic conditions for responsibility since such conditions cut across the free will debate.

⁶ All references to *Case 1* and the FCMA henceforth are references to the new *Case 1* and the new FCMA unless explicitly noted otherwise.

sameness could be interpreted in different manners. However, the intuitively correct interpretation of exact sameness—and the one I will in fact make—is that there are no *intrinsic* differences between Plum’s reasoning process or his decision in *Cases 1–4*. In other words, any differences between Plum’s reasoning process and his decision in *Cases 1–4* concern features that are external to Plum’s reasoning process and his decision.

Pereboom does not explicitly state that the strength of Plum₁’s egoistic tendencies just prior to Plum₁’s decision is part of Plum₁’s reasoning process. I, however, will assume this to be the case, and thus will assume that the strength of the egoistic tendencies of Plum just prior to his decision is the exact same (intrinsically indistinguishable) in *Cases 1–4*, irrespective of the different causal *origins* of the exact strength of these egoistic tendencies. So, given that Plum₁ is an agent, Plum’s reasoning process and his decision are intrinsically indistinguishable in *Cases 1–4*, and given that it is uncontroversial to stipulate that Plum₄ satisfies the *common compatibilist conditions* for responsibility, it should be uncontroversial to stipulate that Plum₁ likewise satisfies the *common compatibilist conditions* for responsibility. I now turn to the compatibilist’s options for replying to the FCMA.

A proponent of a soft-line reply rejects premise (2). A soft-liner might, e.g., claim that, unlike Plum₄, Plum₁ is not blameworthy for deciding to kill White because Plum₁’s decision to kill White was causally determined by factors that, crucially, include the intentional actions of other agents (Lycan 1997). Must a soft-liner adopt a historical condition for responsibility?⁷ Not necessarily. If simultaneous causation is possible (causation between events that occur at the same time), then one could maintain that while the relevant difference between Plum₁ and Plum₄ is a matter of something external to their reasoning process and decision, this difference does not concern what has or has not occurred temporally prior to their reasoning process or decision.⁸ For, if simultaneous causation is possible, then it may be possible for the neuroscientists’ exertion of the momentary egoism-enhancing influence to be simultaneous with Plum₁’s decision to kill White, whereby Plum₁’s deciding to kill White rather than refraining from deciding to kill White counterfactually depends upon this simultaneous egoism-enhancing influence. Still, even if the soft-liner does not adopt a historical condition for responsibility, the soft-liner arguably must commit to an external condition for responsibility that Plum₁ lacks, i.e. a condition that involves factors that are external to the Plum₁’s reasoning process and decision.

In contrast to a soft-line reply, a proponent of the aptly named hard-line reply bites the bullet and rejects premise (1) (McKenna 2008). In other words, despite initial appearances, Plum₁ *is* blameworthy for his decision to kill White. So, a hard-liner does

⁷ For the purposes of this paper, I understand a historical condition for responsibility to include what I will call both a positive-historical condition *and* a negative-historical condition. A positive-historical condition concerns *having* a certain history. By contrast, a negative-historical condition concerns merely *not having* a certain kind of defective history (for further discussion, see e.g. McKenna (2012), Haji (2013), and Mele (2013)). This distinction is irrelevant to my argument below. For, I will be concerned with *extrinsic* differences between Plum₁ and Plum₄ which *include* (but are not necessarily limited to) both positive-historical conditions and negative-historical conditions for responsibility.

⁸ This sort of consideration seems to have been largely overlooked in the literature.

not affirm any such external condition for moral responsibility that the soft-liner arguably must adopt. Rather, the hard-liner maintains that satisfying the *common compatibilist conditions* for responsibility (or some subset thereof) are in fact sufficient for moral responsibility, even in the face of the kind of manipulation at play in *Case 1*.

With this basic understanding of the two possible replies to the FCMA in place, I want to clarify my own assumptions about the FCMA that underlie my case against classical compatibilism. I of course do not assume that the FCMA is sound. Additionally, I don't assume that any reply to the FCMA is implausible, so long as the reply does *not* also assume that determinism is compatible with the ability to do otherwise. Instead, the only assumption I share with Pereboom (2014: 75) is that Plum₁ is an agent and satisfies the *common compatibilist conditions* for responsibility. For reasons given above, this assumption should not be controversial. I now turn to Vihvelin's version of classical compatibilism and its relationship to the FCMA.

2. Vihvelin's Compatibilism

In an important contribution to a distinctive revival of classical compatibilism (which is now referred to as 'the new dispositionalism'),⁹ Vihvelin (2004: 429) claimed that to have free will is "to have the ability to make choices on the basis of reasons and to have this ability is to have a bundle of dispositions." Additionally, "objects have dispositions by having intrinsic properties which are the causal basis of the disposition" (Vihvelin 2004: 436). Now, in her recent book, Vihvelin (2013) clarifies that this notion of free will concerns what she calls a narrow ability. For *S* to have a narrow ability to φ is for *S* to have what it takes to φ : "she's got the necessary skills and the psychological and physical capacity to use those skills" (Vihvelin 2013: 11). Vihvelin (2013: 13) similarly claims that "[n]ecessarily, if two persons are intrinsic duplicates governed by the same laws, they have exactly the same *narrow abilities*" [emphasis added].

In contrast to the narrow ability, *S* has the wide ability to φ iff *S* has the narrow ability to φ , and *S* has "the means and the opportunity and nothing external stands in her way" (Vihvelin 2013: 11). To illustrate, suppose the talented cellist Jacqueline du Pré is on an airplane in which there is no cello. Call this the *airplane scenario*. In the *airplane scenario*, du Pré has the narrow ability to play cello—she's got what it takes to do so. However, du Pré obviously lacks the wide ability to play the cello in the *airplane scenario*. By contrast, if du Pré were sitting in a room with a cello by her side, du Pré would have the narrow ability and the opportunity to play the cello. In other words, du Pré would have the wide ability to play the cello.

Since a simple conditional analysis of dispositions and of abilities, respectively, are each riddled with counterexamples, Vihvelin adopts a more nuanced understanding of dispositions (and of abilities). More specifically, Vihvelin (2013: 187) accepts the following principle:

⁹ Smith (2003) and Fara (2008) are also so-called new dispositionalists.

LCA-PROP-ABILITY S has the narrow ability at time t to do R as the result of trying iff, for *some* intrinsic property B that S has at t, and for some time t' after t, if S had the opportunity at t to do R *and* S tried to do R while retaining property B until time t', then in a *suitable proportion of these cases*, S's trying to do R and S's having of B would be an S-complete cause of S's doing R.¹⁰

Here is an illustration of the basic idea behind LCA-PROP-ABILITY which will suffice for my purposes here. Holding fixed all of my relevant intrinsic properties, if I tried to play the cello, I would fail to do so in a suitable proportion of these cases (I never learned how to play the cello). So, LCA-PROP-ABILITY renders the correct verdict that I lack the narrow ability (and thus also the wide ability) to play the cello. Next, holding fixed all of du Pré's relevant intrinsic properties, if du Pré tried to play the cello, she would succeed in doing so in a suitable proportion of such cases. So, LCA-PROP-ABILITY renders the correct verdict that du Pré has the narrow ability to play the cello.

Vihvelin's distinction between a narrow and wide ability reveals that the notion of having the ability to do otherwise at issue in PAP intuitively concerns (at least) having the wide ability to do otherwise as opposed to merely having the narrow ability to do otherwise (cf. Franklin 2011). For, *if* PAP is true, then surely du Pré cannot be (non-derivatively) morally responsible for not playing the cello in the *airplane scenario*. Indeed, Vihvelin (2013) does seem to accept this sensible position (see footnotes 20 and 21). At any rate, I now stipulate that the notion of having an ability to do otherwise at issue in this paper entails having the wide ability to do otherwise.¹¹ So PAP, as I understand the principle, entails that an agent is responsible for what she has done only if she has the wide ability to do otherwise. To be clear, the classical compatibilist is free to reject PAP as I understand the principle. This does not imply, however, that the classical compatibilist can thereby escape my argument to which I now turn.

3. A Dilemma for Soft-line Classical Compatibilism

A classical compatibilist who adopts a soft-line reply maintains that, unlike Plum₄, Plum₁ is not blameworthy for his decision to kill White. I'll call this view soft-line classical compatibilism (SLCC). I'm going to propose a dilemma for SLCC. This dilemma assumes that a soft-liner must adopt an external condition for responsibility that Plum₁ fails to satisfy. So, before presenting the dilemma, I want to defend this assumption against an important objection.

The SLCC theorist who accepts Vihvelin's framework might mistakenly object to the stipulation that there is no relevant intrinsic difference between Plum₁ and Plum₄ as follows. According to LCA-PROP-ABILITY, in order to determine Plum₁'s intrinsic

¹⁰ LCA-PROP-ABILITY draws its resources from the work of Lewis (1997) and Manley and Wasserman (2008).

¹¹ Notice that I am not committed to the claim that the ability to do otherwise is identical to the wide ability to do otherwise.

properties, we need to hold fixed the fact that Plum₁ is being manipulated when considering a suitable proportion of hypothetical cases. For, being manipulated is an “actual sequence property” of Plum₁. However, holding fixed the fact that Plum₁ is being manipulated, Plum₁ does *not* refrain from deciding to kill White in a suitable proportion of cases. So, LCA-PROP-ABILITY renders the verdict that Plum₁ in fact *lacks* the narrow ability to refrain from deciding to kill White. In other words, it’s not the case that Plum₁’s narrow ability to refrain from deciding to kill White is simply finkish or masked.¹² By contrast, when we hold fixed Plum₄’s intrinsic properties, in a suitable proportion of cases Plum₄ does, we may suppose, refrain from deciding to kill White. So Plum₁ and Plum₄ *cannot* be intrinsic duplicates of one another in all relevant respects after all. More specifically, *unlike* Plum₁, Plum₄ has the narrow ability to refrain from deciding to kill White.¹³ For this reason, one cannot consistently maintain that Plum₁ and Plum₄ are intrinsic duplicates of one another in all relevant respects, and also that only Plum₁ is being manipulated by neuroscientists.

This objection has one critical flaw. There is no inconsistency in supposing that Plum₁ and Plum₄ are intrinsic duplicates of one another in all relevant respects, and that only Plum₁ is being manipulated by neuroscientists. For, in order to determine Plum₁’s narrow abilities according to LCA-PROP-ABILITY, we do *not* hold fixed the fact that Plum₁ is being manipulated when considering a suitable proportion of hypothetical cases. Let me explain.

The manipulation of Plum₁ is nothing over and above the neuroscientists exerting a momentary egoism enhancing tendency in Plum₁. So, the crucial difference between Plum₁ and Plum₄ is, as previously stressed, the causal *origin* of the exact strength of the egoistic tendencies in Plum₁ and Plum₄, respectively. In other words, the difference consists in factors that are *external* to Plum₁ and Plum₄, respectively. So, being manipulated is no doubt an “actual sequence property” of Plum₁. Nevertheless, the neuroscientists’ intentional action is not intrinsic to Plum₁’s reasoning process or Plum₁’s decision. In that case, the neuroscientists’ intentional action *at best* robs Plum₁ of the *opportunity* to decide to refrain from killing White. However, according to LCA-PROP-ABILITY, in order to determine Plum₁’s narrow abilities we need to consider hypothetical cases in which Plum₁ *has* the opportunity to decide to kill White. For this reason, in order to determine Plum₁’s narrow abilities, we do *not* hold fixed the fact that Plum₁ is being manipulated when considering a suitable proportion of hypothetical cases. So, no inconsistency results from stipulating that Plum₁ and Plum₄ are intrinsic duplicates of one another in all relevant respects, and that only Plum₁ is being manipulated by neuroscientists.

¹² A disposition (which, according to Vihvelin, is what an ability reduces to) is finkish when there is something that would remove the disposition precisely under conditions in which the disposition would ordinarily be manifested. A masked disposition is one in which, under conditions in which it would ordinarily be manifested, the disposition fails to be manifested without the disposition being removed.

¹³ In order for LCA-PROP-ABILITY to apply to mental actions such as Plum’s decision to kill White, I am charitably assuming that it makes sense to try to decide to do something, and that there is a way around objections embodied by Lehrer’s (1968) red candy case. See Clarke (2009: 328–329) and Vihvelin (2013: 196–208).

So, *if* there is a morally relevant difference between the actions of Plum₁ and Plum₄ it may be due to the fact only Plum₄ has the wide ability to decide to refrain from killing White. Notice, however, that this position is consistent with my assumption that a soft-liner must adopt an external condition for responsibility that Plum₁ fails to satisfy.¹⁴ Moreover, this position does not escape my dilemma which I will now present.

The SLCC theorist must hold that there is an external condition for responsibility that Plum₁ fails to satisfy. In that case, this external condition either is or is not necessary for an agent to have the ability to do otherwise. In other words, if SLCC is true, then either (i) or (ii) is true:

- (i) The external condition c that is necessary for an agent S to be blameworthy for φ -ing (which Plum₁ fails to satisfy) is *also* necessary for S to be able to do other than φ .
- (ii) The external condition c that is necessary for an agent S to be blameworthy for φ -ing (which Plum₁ fails to satisfy) is *not* necessary for S to be able to do other than φ .

The conjunction of SLCC and (i) entails that Plum₁ is not blameworthy for his decision to kill White, and Plum₁ was not able to refrain from deciding to kill White.¹⁵ Moreover, Plum₁'s inability to refrain from deciding to kill White is necessitated by the intentional actions of the neuroscientists which are part of the factors beyond Plum₁'s control that causally determined Plum₁'s decision to kill White.

To be clear, there may be a *deeper* explanation as to why Plum₁ lacks the ability to do otherwise. In other words, the external condition that is necessary for responsibility could in principle not appeal to the notion of agency or action at all. Still, the SLCC theorist who accepts (i) *is* committed to the claim that the activities of the neuroscientists necessitate Plum₁'s inability to do otherwise, irrespective of whether Plum₁ lacks the ability to do otherwise *because* of the activities of the neuroscientists. So, the conjunction of SLCC and (i) entails the following:

- (A) Possibly, an agent S who φ -s retains the ability to do other than φ even when S 's φ -ing is causally determined by factors beyond S 's control. However, necessarily, S does not retain the ability to do other than φ if φ -ing is causally determined by factors beyond S 's control and those factors include the intentional actions of other agents.

Since the source compatibilist is not committed to the compatibility of determinism and the ability to do otherwise, which is affirmed in the first sentence of (A), the source

¹⁴ Perhaps a compatibilist can construct an alternative account of abilities whereby it *is* objectionable to stipulate that Plum₁ and Plum₄ are intrinsic duplicates of one another in all relevant respects and yet only Plum₁ is being manipulated. However, unless such an alternative account of abilities is offered, I contend that such a stipulation about *Cases 1–4* remains unproblematic.

¹⁵ The SLCC theorist who accepts (i) and PAP would presumably say that Plum₁ is not blameworthy for deciding to kill White at least partly *because* Plum₁ was not able to do other than decide to kill White.

compatibilist is not committed to (A). This is good news for the source compatibilist since, as I will now explain, (A) is implausible.

Consider a case just like *Case 1* except that, instead of neuroscientists, it is a machine with no intelligent designer or a spontaneously generated electromagnetic field that exerts a momentary egoism-enhancing influence upon Plum (compare Pereboom (2014: 79) and Mele (2006: 141)). In these alternative cases, Plum presumably *does* satisfy the external condition *c*, and thus *is* able to refrain from deciding to kill White. However, *under the assumption that determinism is compatible with the ability to do otherwise*, it is difficult to see why the intentional actions of another agent that are part of antecedent causal factors beyond Plum₁'s control are relevant to Plum₁'s ability to do otherwise.

I want to emphasize that a commitment to (A) does not commit one to the claim that *S* lacks the ability to do otherwise *in virtue of* the fact that the relevant factors beyond *S*'s control include the intentional actions of other agents. As previously noted, there may be a *deeper* explanation as to why Plum₁ lacks the ability to do otherwise. Even so, the burden is on one who accepts (A) to specify what this deeper explanation amounts to.

To be fair, both classical compatibilists and source compatibilists that accept a soft-line reply must accept the similar yet distinct claim that the intentional actions of another agent that are part of the relevant antecedent causal factors make a difference with respect to one's *responsibility*. However implausible this claim is, however, the crucial point for my argument is that it is *more* plausible than (A), as I will now explain.

There is no clear consensus on how to understand notions of collective responsibility, or how responsibility distributes to individuals (cf. French and Wettstein 2006). As a result, a soft-liner should consider the viable response that there is an *erosion* of Plum₁'s responsibility in light of the fact that the intentions of the neuroscientists are salient to White's death (Cushman 2008; Levy 2013; Phillips and Shaw 2014; Murray and Lombrozo ms.). More specifically, one might suggest that Plum₁'s responsibility traces back to the *neuroscientists*, but that Plum's responsibility cannot trace back to, say, electromagnetic fields. By contrast, it seems less compelling to suppose that there is an erosion of an ability to do otherwise. The idea of Plum₁'s ability to do otherwise dissolving into the hands of the neuroscientists, but *not* dissolving into the hands of, say, an electromagnetic field (whatever that means) is surely harder to swallow. So I conclude that (A) is implausible, and, crucially, that (A) is *less* plausible than the claim that the intentional actions of other agents can bear upon an agent's *responsibility*.

Let's now consider the second horn of the dilemma for SLCC. According to (ii), since Plum₁'s decision to kill White was causally determined by factors beyond Plum₁'s control, and, crucially, such factors include the intentional actions of other agents, Plum₁ fails to satisfy some external condition for responsibility. So Plum₁ is not blameworthy for his decision to kill White. However, according to (ii), this external condition is not necessary for having the ability to do otherwise. So Plum₁ retains the ability to refrain from deciding to kill White. Additionally, as indicated in the first section, I assume that Plum₁ satisfies the *common compatibilist conditions* for responsibility.

So, the conjunction of SLCC and (ii) entails the following: even though Plum₁ retains the ability to refrain from deciding to kill White and Plum₁ satisfies the *common*

compatibilist conditions for responsibility, Plum₁ is nevertheless *not* blameworthy for his decision to kill White because Plum₁ fails to satisfy some external condition for responsibility in light of the activities of the neuroscientists. Here is the more general claim that is entailed by the conjunction of SLCC and (ii):

- (B) Possibly, an agent *S* who φ -s retains the ability to do other than φ and satisfies the *common compatibilist conditions* for responsibility, even when *S*'s φ -ing is causally determined by factors beyond *S*'s control. However, necessarily, *S* is not blameworthy for φ -ing if *S*'s φ -ing was causally determined by factors beyond *S*'s control, and such factors include the intentional actions of other agents.

Since the source compatibilist is not committed to the compatibility of determinism and the ability to do otherwise, which is affirmed in the first sentence of (B), the source compatibilist is not committed to (B) which is good news. For, as I will now attempt to show, it is difficult to see why the intentional actions of other agents should undermine *S*'s blameworthiness *if S* satisfies the *common compatibilist conditions* for responsibility and *S* retains the ability to do otherwise.

Anyone who thinks that we are free and morally responsible must hold that, contra Strawson (1994), we are free and morally responsible despite not being *causa sui*, i.e. despite not being the cause of ourselves. In other words, acting freely and being morally responsible is compatible with one's action being *influenced* by external factors, including the intentional actions of other agents.¹⁶ For instance, a manager may be held responsible for harshly insulting her employee, *despite* the fact that the employee was not following the manager's directions, i.e. despite the fact that the manager's insult was *influenced* by her employee.

The point of highlighting the fact that we may be responsible for actions that are influenced by other agents is that we should similarly think that Plum₁ is blameworthy for his decision to kill White *if* we assume that Plum₁ retains the ability to refrain from deciding to kill White and Plum₁ satisfies the *common compatibilist conditions* for responsibility. For, under these assumptions, Plum₁ seems to be merely influenced by the neuroscientists. So, the fact that freedom and responsibility are compatible with a failure to be *causa sui* should render (B) untenable.

To further illustrate this point, if Vihvelin were to accept a soft-line reply and (ii), she would have to maintain that although Plum₁ retains the wide ability to refrain from deciding to kill White, and Plum₁ satisfies the *common compatibilist conditions* for responsibility, Plum₁ is nevertheless blameless for deciding to kill White. But what could possibly be missing here that renders Plum₁ blameless? Barring the impossible demand for being *causa sui*, it seems that there is nothing further that *could* be demanded in order to render Plum₁ blameworthy.

¹⁶ Perhaps we lack ultimate 'heaven and hell' responsibility for our actions given that we are not *causa sui*. As Clarke (2005: 20–3) notes, however, it does not follow that we lack a weaker form of basic desert responsibility.

In an attempt to push back against my attack on (B), one might reply as follows. We must distinguish between a weak and strong form of influence. The weak form of influence is compatible with the ability to do otherwise and responsibility. By contrast, the strong form of influence is compatible with the ability to do otherwise as well as the *common compatibilist conditions* for responsibility, but is *not* compatible with responsibility. Additionally, the manager's insult was influenced in the weak form whereas Plum₁'s decision to kill White was influenced in the strong.

This reply is unpromising. We can assume in the above example that the manager who insults her employee retains the *common compatibilist conditions* for responsibility and also retains the ability to do otherwise. Moreover, remember that Plum₁ has not been given an irresistible urge to decide to kill White. In that case, it is difficult to see how the compatibilist could distinguish in a non-*ad hoc* manner between the two aforementioned kinds of influence, such that the manager is influenced only in the weaker form but Plum₁ is influenced in the stronger form.¹⁷ At any rate, the burden seems to be squarely upon the SLCC theorist who accepts (ii)—and who is thereby rationally committed to (B)—to supply us with such a non-*ad hoc* distinction. Unless this is done, (B) should be rejected.

Given the above dilemma for SLCC, it appears that the classical compatibilist has good reason to adopt a hard-line reply. However, as I will now argue, affirming a hard-line reply likewise results in counterintuitive implications.

4. A Dilemma for Hard-line Classical Compatibilism

The classical compatibilist who adopts a hard-line reply maintains that Plum₁ is blameworthy for deciding to kill White. I'll call this view 'hard-line classical compatibilism' (HLCC). My case against HLCC is likewise embodied in the form of a dilemma:

(iii) PAP is true.

(iv) PAP is false.

Recall my stipulation that the notion of having an ability to do otherwise at issue in this paper entails having the wide ability to do otherwise. As a result, PAP, as I understand the principle, entails that an agent is responsible for what she has done only if she has the wide ability to do otherwise. Now, the conjunction of HLCC and (iii) entails that since Plum₁ is blameworthy for deciding to kill White, Plum₁ must have the ability to refrain

¹⁷ One difference between Plum₁ and the manager is that Plum₁ is *unaware* of the manipulation, whereas the manager is, we may suppose, aware of the fact that the employee influenced the manager's action of insulting the employee. But this difference is surely irrelevant. For, I know of no one who thinks that we must be aware of all of the factors that are influencing our action in order to be responsible for that action. Moreover, such a condition on responsibility would render virtually everyone blameless for their actions since virtually no one is aware of *all* of the factors that influence their actions.

from deciding to kill White. More generally, the conjunction of HLCC and (iii) entails the following:

- (C) Possibly, an agent S who φ -s retains the ability to do other than φ even when S 's φ -ing is a result of the direct¹⁸ manipulation by another agent.

Fortunately, since the source compatibilist is not committed to the compatibility of determinism and the ability to do otherwise, she is not committed to (C).

I suspect that a number of compatibilists who offered a soft-line reply to the old FCMA will affirm a hard-line reply to the new FCMA given that the kind of manipulation at play in the new *Case 1* is less invasive. Similarly, a classical compatibilist might be less reluctant to accept (C) than they would be otherwise given that the notion of direct manipulation mentioned in (C) now includes less invasive activities, such as the exertion of a momentary egoism-enhancing influence. While I grant that (C) would be even more implausible if the notion of direct manipulation mentioned in (C) concerned nothing less invasive than the kind of manipulation at play in the old *Case 1*, the classical compatibilist should not rest easy.

Notice that *regardless* of whether the neuroscientists manipulate Plum₁ according to the old *Case 1* or the new *Case 1*, the scientists know (we may suppose) that either form of manipulation guarantees—indeed, nomologically necessitates—that Plum₁ will do precisely what the neuroscientists want Plum₁ to do. So, if you agree that Plum₁ cannot do otherwise in the old *Case 1* just because the neuroscientists act on what they know will nomologically necessitate Plum₁'s decision to kill White, then you should agree that Plum₁ likewise cannot do otherwise in the new *Case 1*, and thus reject (C).

Besides the above line of reasoning against (C), there is, I think, internal pressure on compatibilists such as Vihvelin who accept PAP to reject (C). In brief, a Frankfurt-style case (FSC) turns on the distinction between an actual intervener and a merely counterfactual intervener. According to the Frankfurtian, a merely counterfactual intervener removes an agent's ability to do otherwise. Vihvelin disagrees. However, Vihvelin and other classical compatibilists seem to agree that an *actual* intervener removes an agent's ability to do otherwise. So they should reject (C); they should say that it is *not* possible for an agent to retain the ability to do otherwise in the midst of an *actual* intervener. I will now expound on the above remarks by drawing out an intriguing connection between FSCs and the FCMA.

In a traditional deterministic FSC the following is stipulated: Black wants Jones to φ . Jones ends up φ -ing on her own (though Jones' φ -ing is causally determined by factors beyond her control),¹⁹ and thus Black does not actually intervene. However, Black would have intervened if Jones were about to $\sim\varphi$ (in the absence of Black's intervention). So,

¹⁸ I understand the kind of manipulation in *Case 1* to be direct, whereas the kind of manipulation in *Case 2* (being programmed at the beginning of one's life) fails to be direct.

¹⁹ The libertarian would of course dispute the claim that Jones φ -s on her own if Jones' φ -ing is causally determined by factors beyond her control. But here we are interested in what classical compatibilists such as Vihvelin say about traditional deterministic FSCs.

the fact that Black would have intervened if Jones were about to $\sim\phi$ (in the absence of Black's intervention) implies that Jones couldn't refrain from ϕ -ing. In other words, Jones couldn't have done otherwise. However, even though Jones could not have done otherwise, Jones is still blameworthy for ϕ -ing (let's assume that ϕ -ing is a wrong action rather than a right action).

In defense of PAP, Vihvelin seems to hold that Jones can (has the *wide* ability to) $\sim\phi$ in a deterministic FSC if the method that Black is employing is only the *Preemptor* method.²⁰ According to the *Preemptor* method, Black's intervention in the counterfactual scenario is causally triggered by an event that is not an action by Jones, but nevertheless reliably indicates that Jones will $\sim\phi$ (barring an intervention by Black) (Vihvelin 2013: 97–99). Vihvelin's position that Jones can do otherwise in certain deterministic FSCs is not without adherents (Smith 2003: 25; Fara 2008: 854–855; Nelkin 2011: 64–71). This position is, however, controversial (Fischer 2008; Clarke 2009: 339–342; Franklin 2011; Kittle 2014; Sartorio Forthcoming). Nevertheless, I intend to sidestep this controversial issue entirely. For, *even if* Vihvelin has plausibly argued that Jones can do otherwise in a deterministic FSC, I am interested in what Vihvelin and other compatibilist defenders of PAP would say if the counterfactual scenario of a deterministic FSC were instead actual. In other words, I am interested in what Vihvelin and others would say about a case in which Black *actually* intervenes. Let's call such a case '*Actual Intervention*'.

A libertarian defender of PAP would of course hold that Jones is blameless (and lacks the ability to do otherwise) in *Actual Intervention* because Jones' ϕ -ing was causally determined by factors beyond her control. But what a compatibilist defender of PAP would say about *Actual Intervention* is less obvious. Still, Vihvelin seems to agree that Jones lacks the wide ability to do otherwise in *Actual Intervention*. There are two reasons for attributing this view to Vihvelin.

First, it would be odd for Vihvelin (2013: 108) to emphasize that, when Black is employing the *Preemptor* method, Jones retains the wide ability to do otherwise “on those occasions when Black *doesn't* intervene” (emphasis in the original) if Vihvelin thought that Jones retains the wide ability to do otherwise, *irrespective* of whether Black intervenes when employing the *Preemptor* method.

The second reason for attributing to Vihvelin the position that Jones lacks the wide ability to do otherwise in *Actual Intervention* stems from her thoughts about the other method Black may employ in a deterministic FSC, *viz.* the *Bodyguard* method. If Black is employing the *Bodyguard* method in a deterministic FSC (a case in which Black does not actually intervene), then Black's intervention would have been causally triggered by the beginning of any action by Jones that is contrary to Black's plan. For example, Black's intervention might have been causally triggered by Jones' trying to $\sim\phi$. Now, according to Vihvelin, if Black is only employing the *Bodyguard* method in a

²⁰ Although Vihvelin does not explicitly say that Jones has the wide ability to do otherwise in a deterministic FSC in which Black doesn't intervene but employs the *Preemptor* method, this is, I think, the correct and charitable interpretation of Vihvelin's position given her use of an analogy of a coin toss in the context of Jones' abilities in a deterministic FSC. More specifically, Vihvelin (2013: 104) says that a coin which comes up heads could have come up tails in light of not only the coin's intrinsic properties, but also given “facts about its [the coin's] environment”, *i.e.* given the coin's *opportunity* to come up tails.

deterministic FSC (a case in which Black does not actually intervene), then Jones lacks the wide ability to $\sim\phi$, although Jones retains the wide ability to *choose* to $\sim\phi$ (or at least try or begin to choose to $\sim\phi$) (Vihvelin 2013: 100).²¹ In that case, Vihvelin would surely agree that in *Actual Intervention* according to which Black *actually* intervenes by employing the *Bodyguard* method, Jones *still* lacks the wide ability to $\sim\phi$ (even if Jones retains the wide ability to, e.g., try to $\sim\phi$). After all, it would be extremely odd to suppose that if Black doesn't actually intervene, then Jones lacks the wide ability to do otherwise, but if Black actually intervenes, then Jones retains the wide ability to do otherwise. This concludes my two reasons for attributing to Vihvelin the view that Jones lacks the wide ability to do otherwise in *Actual Intervention*.

We can now begin to see how the position I am attributing to Vihvelin is relevant to (C). Vihvelin's position on deterministic FSCs seems to be as follows: irrespective of which method Black employs, so long as Black *actually* intervenes in such a way that the activities of Black guarantee—indeed, nomologically necessitate—that Jones ϕ -s, then Jones lacks the ability to do otherwise. *A fortiori*, even if Jones retains the *common compatibilist conditions* for responsibility in *Actual Intervention*, Black's *actually* intervening still implies that Jones lacks the ability to do otherwise. Let's call such a case in which Jones retains the *common compatibilist conditions* for responsibility in *Actual Intervention* '*Case Actual Intervention*'.

Alas, notice that we have stumbled upon a case that is structurally similar to Pereboom's *Case I*. More specifically, both *Case Actual Intervention* and *Case I* involve

²¹ Vihvelin holds that a deterministic FSC in which Black (who does not actually intervene) employs the *Bodyguard* method does not refute PAP even though Jones who ϕ -s does not retain the wide ability to $\sim\phi$ but only retains the wide ability to, say, try to $\sim\phi$. Consider what Vihvelin's (2013: 101) says:

Insofar as Black is a Bodyguard [...] he succeeds in depriving Jones of the alternative required for *freedom of action*, but he *necessarily* fails to deprive Jones of the alternatives required for *freedom of will*. And, of course, it is freedom of will that is the classical locus of moral responsibility.

I understand Vihvelin's remarks here not to concern the distinction between having a narrow ability to ϕ and a having wide ability to ϕ . Instead, I understand these remarks to concern the distinction between the performance of an overt bodily movement and the mere performance of a mental action. Applying Vihvelin's remarks to the FCMA is complicated by the fact that the kind of action Plum performs is a decision—a mental action. Still, let's continue to assume the coherence of trying to perform a decision. Now, suppose Vihvelin were to hold (as she very well may) that in order for Plum to be responsible for deciding to kill White, Plum only needs the wide ability to *try* to decide to kill White. Might Vihvelin avoid being committed to (C) even if, as a proponent of PAP, she were to endorse a hard-line reply? In other words, can Vihvelin commit to HLLC and (iii) without being committed to (C)? No. Instead of focusing on Plum₁'s decision, we can simply shift our attention to Plum₁'s *attempt* to make his decision. If Plum₁ is responsible for his attempt to make his decision, and PAP is true, then Plum₁ could have refrained from trying to make his decision, *despite* the fact that Plum₁'s attempt to make his decision was the result of the direct manipulation of the neuroscientists.

What if Vihvelin were to accept a hard-line reply, reject PAP as I understand the principle, and hold instead that an agent is responsible for what she has done only if she has the narrow (but not the wide) ability to do otherwise, or at least has the narrow (but not the wide) ability to try to do otherwise? If Vihvelin went this route, she would be subject to the second horn of my dilemma for the HLLC theorist which I will discuss shortly. Moreover, going this route seems to have the implausible implication that in the *airplane scenario* discussed in section 2, Jacqueline du Pré *can* be (non-derivatively) morally responsible for not playing the cello.

an agent who is manipulated to perform some action while simultaneously retaining the *common compatibilist conditions* for responsibility. As a result, since Vihvelin (and other compatibilists who defend PAP) seem to implicitly hold to the sensible position that Jones lacks the ability to do otherwise in *Case Actual Intervention*, Vihvelin and her cohorts should similarly hold that Plum₁ lacks the ability to refrain from deciding to kill White, and thus they should reject (C). This concludes my attempt to show that, by Vihvelin's own lights, (C) is false. Moreover, the burden appears to be on one who wishes to uphold (C) to offer some explanation as to how an agent can do otherwise when she is being manipulated. Unless this is done, (C) should be rejected.²²

The HLCC theorist might therefore think that the best route to go is to accept (iv) instead of (iii). Accepting (iv), however, similarly results in one of the aforementioned counterintuitive implications. More specifically, if the HLCC theorist wishes to reject (C), she must deny that Plum₁ can refrain from deciding to kill White, despite the fact that determinism is compatible with the ability to do otherwise. So, the HLCC theorist that goes this route is committed to (A) which I'll remind the readers is the following:

- (A) Possibly, an agent *S* who φ -s retains the ability to do other than φ even when *S*'s φ -ing is causally determined by factors beyond *S*'s control. However, necessarily, *S* does not retain the ability to do other than φ if φ -ing is causally determined by factors beyond *S*'s control and those factors include the intentional actions of other agents.

As I have already argued, (A) is implausible. And, to repeat, since the source compatibilist is not committed to the view that determinism is compatible with the ability to do otherwise, the source compatibilist is not committed to (A).

Conclusion

I have argued that classical compatibilism, in conjunction with either a soft-line reply or a hard-line reply to Pereboom's FCMA, has counterintuitive implications. This is because whichever reply the classical compatibilist endorses, she will be committed to one of the following implausible positions:

²² Among the very few philosophers that have explicitly considered the relationship between classical compatibilism and the manipulation argument, Berofsky (2006: 435–9) claims that it is question-begging against the classical compatibilist to assume that an agent cannot do otherwise in the case of manipulation. My response is threefold. First, classical compatibilism does not by itself entail (C). So, to say that (C) is implausible does not amount to saying that classical compatibilism is implausible. Rather, the point of highlighting (C)'s implausibility is to show that the conjunction of HLCC and PAP is implausible. Second, as already noted, in light of what classical compatibilists such as Vihvelin seem to implicitly hold regarding *Actual Intervention*, there is indeed internal pressure upon such classical compatibilists to reject (C). Third, even if my case against (C) won't convince a classical compatibilist, a group of agnostics about classical compatibilism would arguably be inclined to reject (C). This in turn should count as a cost to those who accept HLCC and PAP. In other words, my attack on (C) is not a so-called philosophical failure. For more on what does and does not constitute a philosophical failure, see Fischer and Tognazzini (2007).

- (A) Possibly, an agent S who φ -s retains the ability to do other than φ even when S 's φ -ing is causally determined by factors beyond S 's control. However, necessarily, S does not retain the ability to do other than φ if φ -ing is causally determined by factors beyond S 's control *and those factors include the intentional actions of other agents*.
- (B) Possibly, an agent S who φ -s retains the ability to do other than φ and satisfies the *common compatibilist conditions* for responsibility, even when S 's φ -ing is causally determined by factors beyond S 's control. However, necessarily, S is not blameworthy for φ -ing if S 's φ -ing was causally determined by factors beyond S 's control, and such factors include the intentional actions of other agents.
- (C) Possibly, an agent S who φ -s retains the ability to do other than φ even when S 's φ -ing is a result of the direct manipulation by another agent.

By contrast, whichever response to the FCMA the source compatibilist adopts, she will not be committed to any of these implausible positions. For, in order to avoid all three positions, we need not hold that determinism is incompatible with moral responsibility. But we must hold that determinism is incompatible with the ability to do otherwise. I conclude that Pereboom's FCMA, at the very least, undermines classical compatibilism.

Acknowledgments An early version of this paper was presented at the "Free Will and Moral Responsibility" summer school, hosted by the Moscow Center for Consciousness Studies; thanks to all participants in this group for helpful comments. Additionally, for helpful comments and discussion, I am grateful to Lorenza D'Angelo, John Fischer, Mark Heller, Sofia Jeppsson, Simon Kittle, Benjamin Matheson, Dan Miller, Matthew Talbert, Travis Timmerman, and an anonymous referee.

References

- Baker, L. R. (2006). Moral responsibility without libertarianism. *Noûs*, 40, 307–330.
- Berofsky, B. (2006). Global control and freedom. *Philosophical Studies*, 131, 419–445.
- Berofsky, B. (2012). *Nature's challenge to free will*. New York: Oxford University Press.
- Campbell, Joseph. (1997). A compatibilist theory of alternative possibilities. *Philosophical Studies*, 88, 319–330.
- Clarke, R. (2005). On an argument for the impossibility of moral responsibility. *Midwest Studies in Philosophy*, 29, 13–24.
- Clarke, R. (2009). Dispositions, abilities to act, and free will: the new dispositionalism. *Mind*, 118, 323–351.
- Cushman, F. (2008). Crime and punishment: distinguishing the roles of causal and intentional analyses in moral judgment. *Cognition*, 108, 353–380.
- Demetriou, K. (2010). The soft-line solution to Pereboom's four-case argument. *Australasian Journal of Philosophy*, 88, 595–617.
- Fara, M. (2008). Masked abilities and compatibilism. *Mind*, 117, 843–865.
- Fischer, J. M. (2004). Responsibility and manipulation. *The Journal of Ethics*, 8, 145–177.
- Fischer, J.M. (2008). Freedom, foreknowledge, and Frankfurt: a reply to Vihvelin. *Canadian Journal of Philosophy* 38: 327–342.

- Fischer, J. M. (2010). The Frankfurt cases: the moral of the stories. *The Philosophical Review*, 119, 315–336.
- Fischer, J. M., & Ravizza, M. (1998). *Responsibility and control: a theory of moral responsibility*. Cambridge: Cambridge University Press.
- Fischer, J. M., & Tognazzini, N. A. (2007). Exploring evil and philosophical failure: a critical notice of Peter van Inwagen's *The Problem of Evil*. *Faith and Philosophy*, 24, 458–474.
- Frankfurt, H. (1969). Alternate possibilities and moral responsibility. *The Journal of Philosophy*, 66, 829–839.
- Frankfurt, H. (1971). Freedom of the will and the concept of a person. *Journal of Philosophy*, 68, 5–20.
- Franklin, C.E. (2011). Masks, abilities, and opportunities: why the new dispositionalism cannot succeed. *The Modern Schoolman* 88: 89–103.
- French, P., & Wettstein, H. K. (Eds.). (2006). *Midwest Studies in Philosophy: shared intentions and collective responsibility*, vol. 30.
- Ginet, C. (1966). Might we have no choice? In K. Lehrer (Ed.), *Freedom and determinism*, 87–104. New York: Random House.
- Haji, I. (1998). *Moral accountability*. New York: Oxford University Press.
- Haji, I. (2013). Historicism, non-historicism, or a mix? *Journal of Ethics*, 17, 185–204.
- Horgan, T. (1979). 'Could', possible worlds, and moral responsibility. *The Southern Journal of Philosophy*, 17, 345–358.
- Kapitan, T. (2011). A compatibilist reply to the consequence argument. In R. Kane (Ed.), *The oxford handbook of free will 2nd ed*, 131–150. New York: Oxford University Press.
- Lehrer, K. (1968). Cans without ifs. *Analysis*, 29, 29–32.
- Lehrer, K. (1976). 'Can' in theory and in practice: a possible worlds analysis. In M. Brand and D. Walton (Eds.), *Action Theory*, Dordrecht: Reidel.
- Levy, N. (2013). Manipulations for manipulation cases. Posted on Flickers of Freedom, June 13, 2013. Accessed on September 6, 2014 from: http://philosophycommons.typepad.com/flickers_of_freedom/2013/06/manipulations-for-manipulation-cases.html.
- Lewis, D. (1981). Are we free to break the laws? *Theoria*, 47, 113–121.
- Lewis, D. (1997). Finkish dispositions. *Philosophical Quarterly*, 47, 143–158.
- Lycan, W. G. (1997). *Consciousness*. Cambridge, MA: MIT Press.
- Manley, D., & Wasserman, R. (2008). On linking dispositions and conditionals. *Mind*, 117, 59–84.
- Mele, A. R. (1995). *Autonomous agents*. New York: Oxford University Press.
- Mele, A.R. (2005). A critique of Pereboom's 'four-case' argument for incompatibilism. *Analysis*, 65, 75–80.
- Mele, A. R. (2006). *Free will and luck*. New York: Oxford University Press.
- Mele, A.R. (2013). Moral responsibility, manipulation, and minutelings. *Journal of Ethics*, 17, 153–166.
- McKenna, M. (2005). Where Frankfurt and Strawson meet. *Midwest Studies in Philosophy*, 29, 163–180.
- McKenna, M. (2008). A hard-line reply to Pereboom's four-case argument. *Philosophy and Phenomenological Research*, 77, 142–159.
- McKenna, M. (2012). Moral responsibility, manipulation arguments, and history: assessing the resilience of nonhistorical compatibilism. *Journal of Ethics*, 16, 145–174.
- Murray, D., & Lombrozo, T. (Manuscript). Effects of manipulation on attributions of causation,

- free will, and moral responsibility.
- Nelkin, D. K. (2011). *Making sense of freedom and responsibility*. New York: Oxford University Press.
- Pereboom, D. (2001). *Living without free will*. Cambridge: Cambridge University Press.
- Pereboom, D. (2014). *Free will, agency, and meaning in life*. New York: Oxford University Press.
- Phillips, J., & Shaw, A. (2014). Manipulating morality: third-party intentions alter moral judgments by changing causal reasoning. *Cognitive Science*, 38, 1–28.
- Perry, J. (2008). Can't we all just be compatibilists?: a critical study of John Martin Fischer's "My Way". *The Journal of Ethics*, 12, 157–166.
- Sartorio, C. (2011). Actuality and responsibility. *Mind*, 120, 1071–1097.
- Sartorio, C. (Forthcoming). Vihvelin on Frankfurt-style cases and the actual-sequence view. *Criminal Law and Philosophy*.
- Saunders, J. T. (1968). The temptation of powerlessness. *American Philosophical Quarterly*, 5, 100–108.
- Shabo, S. (2010). Uncompromising source incompatibilism. *Philosophy and Phenomenological Research*, 80, 349–383.
- Smith, M. (2003). Rational capacities, or: how to distinguish recklessness, weakness, and compulsion. In Stroud and Tappolet (Eds.), *Weakness of Will and Practical Irrationality*, 17–38. New York: Oxford University Press.
- Strawson, G. (1994). The impossibility of moral responsibility. *Philosophical Studies*, 75, 5–24.
- Van Inwagen, P. (1983). *An essay on free will*. Oxford: Oxford University Press.
- Vihvelin, K. (2004). Free will demystified: a dispositional account. *Philosophical Topics*, 32, 427–450.
- Vihvelin, K. (2013). *Causes, laws, and free will: why determinism doesn't matter*. New York: Oxford University Press.
- Wallace, R. J. (1994). *Responsibility and the moral sentiments*. Cambridge: Harvard University Press.
- Watson, G. (1975). Free agency. *The Journal of Philosophy*, 83, 517–522.
- Wolf, S. (1990). *Freedom within reason*. New York: Oxford University Press.