

Deliberating in the presence of manipulation

Yishai Cohen

Published in the *Canadian Journal of Philosophy*

Abstract According to deliberation compatibilism, rational deliberation is compatible with the belief that one's actions are causally determined by factors beyond one's control. This paper offers a counterexample to recent accounts of rational deliberation that entail deliberation compatibilism. The counterexample involves a deliberator who believes that whichever action she performs will be the result of deterministic manipulation. It is further argued that there is no relevant difference between the purported counterexample and ordinary doxastic circumstances in which a determinist deliberates.

Keywords Deliberation · Determinism · Compatibilism · Incompatibilism · Manipulation

1. Introduction

According to causal determinism, a proposition describing the complete state of the world at any time, in conjunction with the laws of nature, entails every proposition describing the complete state of the world at every other time (van Inwagen 1983, 65). Besides the usual threats that causal determinism poses, *belief* in the truth of causal determinism has likewise posed a threat to certain things that we value, such as rational deliberation. Those who find this threat illusory are deliberation compatibilists:

Deliberation Compatibilism (DC) *S*'s deliberating and being rational is compatible with *S*'s believing that her actions are causally determined by causal antecedents beyond her control.

Their opponents are deliberation incompatibilists:

Deliberation Incompatibilism (DI) *S*'s deliberating and being rational is incompatible with *S*'s believing that her actions are causally determined by causal antecedents beyond her control.

This paper aims to show that deliberation incompatibilists have an untapped resource at their disposal. Just as cases of deterministic manipulation can support the incompatibility of moral responsibility and causal determinism (henceforth 'determinism'), cases in which an agent believes that she will be deterministically manipulated can support the incompatibility of rational deliberation with a belief in the truth of determinism.

In section 2 I introduce the concept of rational deliberation employed in this paper. In section 3 I discuss recent accounts of rational deliberation that entail *DC* (henceforth 'pro-*DC* views'), and then offer a counterexample to these accounts in section 4. The counterexample involves a deliberator who believes that whichever action she performs will be due to deterministic manipulation. In section 5 I incorporate this counterexample into a four-case

deliberation argument (FCDA) against *DC* that is structurally similar to Derk Pereboom's four-case manipulation argument (FCMA) against compatibilism, the thesis that moral responsibility and determinism are compatible. After responding to objections in section 6, I argue in section 7 that the structural similarity between the two arguments suggests an important lesson: both arguments stand or fall together because there is a certain kind of control that is relevant to both rational deliberation and moral responsibility.

2. Rational deliberation: one account of what it is, and what it isn't

Deliberation is a process or activity in which one is figuring out what to do. As Richard Taylor (1966: 168) notes, deliberation has "as its aim or goal a decision to act", as opposed to the goal of merely forming a *belief* about which action one will perform. Unlike the epistemic activity of inferring or predicting what will occur, deliberation is an activity or process that is intended to play an explanatory role with respect to what one ends up doing.

This paper marks a sharp distinction between deliberation *simpliciter* and rational deliberation, the latter of which is at issue in the *DC/DI* debate. All doxastic requirements for the former are requirements for the latter, but not vice versa. A central and unique doxastic requirement for rational deliberation is that the agent must not possess any inconsistent beliefs that are salient to her deliberation. Let us call this the 'no inconsistent beliefs' (*NIB*) thesis:

(*NIB*) In order to rationally deliberate about whether to do A1 or A2, where A1 and A2 are distinct actions, an agent must not have any inconsistent beliefs that are salient to her deliberation about whether to do A1 or A2.

To illustrate, suppose that Billy correctly believes that he cannot fly to Mars. If Billy were to deliberate about whether to fly to Mars, then his deliberation would not be an instance of rational deliberation. This is because deliberation about whether to fly to Mars presumably commits Billy to the belief that he can fly to Mars, or at least that he might be able to fly to Mars.¹ But these beliefs are inconsistent with Billy's initial belief that he cannot fly to Mars. Moreover, these inconsistent beliefs are salient to his deliberation about whether to fly to Mars. So Billy has not satisfied *NIB*'s condition in this instance of deliberation. While it may be difficult to offer precise conditions under which an agent's belief is salient to her deliberation, such precision is peripheral to the present discussion.²

I will understand the doxastic requirements for rational deliberation to imply that the kinds of beliefs one must possess in order to rationally deliberate are only beliefs to which an agent is rationally committed upon reflection, as opposed to occurrent beliefs or dispositional

¹ I share van Inwagen's (1983: 154) assumption that one's beliefs can be manifested in one's non-verbal behavior, and specifically in the process of deliberation.

² John Keller has made the plausible suggestion that *NIB* can be replaced with a principle that is restricted to an agent's lacking inconsistent beliefs about her *abilities* that are salient to her deliberation (personal correspondence).

beliefs.³ According to this approach, rational deliberation requires neither occurrent beliefs nor dispositional beliefs of a certain kind. Since many of the suggested doxastic conditions for rational deliberation require a high level of cognitive sophistication, we get the intuitively correct result that agents can rationally deliberate even if they have never read or considered the kinds of doxastic conditions for rational deliberation discussed in academia.

The somewhat idealized concept of rational deliberation that I want to explore has one important feature. According to this interpretation of *DC* and *DI*, an agent is rational insofar as she does not have a false belief about whether determinism is compatible with the ability to do otherwise. There are other perfectly good conceptions of rational deliberation that lack this feature. But I focus upon this concept of rational deliberation for two reasons. First, this concept should be of interest to anyone who wishes to deny the following claim: ‘any determinist who deliberates would come to see upon proper reflection that, in virtue of being a determinist, she is rationally committed to the possession of inconsistent beliefs that are salient to her deliberation’. Second, we don’t want inconsistent beliefs that are salient to our deliberation at least partly because we don’t want to have any false beliefs that are salient to our deliberation. So it is not much of a stretch to connect rational deliberation with the demand that one not have a false belief concerning the compatibility between determinism and the ability to do otherwise.

Given this somewhat idealized version of rationality, my argument assumes that the ability to do otherwise is in fact incompatible with determinism (Ginet 1966; van Inwagen 1983; Cohen 2015; Todd 2016). If we maintain instead that they are compatible, then it becomes extremely difficult to argue for *DI*. This is evident by the fact that many deliberation compatibilists have attempted to defend their position while granting (or at least never denying) that determinism is incompatible with the ability to do otherwise.

With this understanding of rational deliberation in place, and its accompanying assumptions, I now turn to discuss a doxastic requirement for deliberation *simpliciter*, since all such requirements are also requirements for rational deliberation (but not vice versa). According to Taylor (1966: 177), van Inwagen (1983: 155), and Coffman and Warfield (2005), deliberation requires the agent to believe that, for each action under consideration, she *can* perform that action. Coffman and Warfield (2005) refer to this as the “Belief in Ability Thesis” (BAT). Since there is always more than one action under consideration, a deliberator must thus believe that, under the current circumstances, she can do other than what she will in fact do. So if BAT is true, and if the ability to do otherwise is incompatible with determinism (as I am assuming to be the case), it follows that *DI* is true, given my aforementioned understanding of rational deliberation.

But BAT appears to be false. According to an alternative viewpoint, deliberation only requires the agent to believe that, for each action under consideration, she *might* be able to perform the action. In other words, the action under consideration only needs to be a viable epistemic possibility from the viewpoint of the deliberator. In the next section, we will review arguments for BAT by Taylor and van Inwagen, and see how a deliberation compatibilist can resist these arguments.

³ This is a suggestion made by Nelkin, which is discussed by Pereboom (2014, 119).

3. Doxastic requirements for rational deliberation

Besides satisfying *NIB*'s condition, as well as other conditions that cut across the *DC/DI* debate, recent pro-*DC* views may be divided into three groups: those that affirm an epistemic openness condition, those that affirm a deliberative efficacy condition, and those that affirm both. The epistemic openness condition may be articulated as follows:

(EO) In order to rationally deliberate about whether to do A1 or A2, where A1 and A2 are distinct actions, an agent must not believe that she will do A1, or that she will do A2. Moreover, all of the agent's beliefs must be consistent with the agent's performing A1, as well as with the agent's performing A2.

We don't need to review other versions of the epistemic openness condition since the forthcoming counterexample will focus primarily upon the deliberative efficacy condition. According to this condition, in order to rationally deliberate, one must believe that one's deliberation will, in some sense, make a difference with respect to how one will in fact act. The details of each version of the deliberative efficacy condition must be carefully inspected for our purposes, beginning with Tomis Kapitan's version.

Kapitan (1996, 436) says that rational deliberation requires a certain kind of metaphysical openness:

(PO) At t_1 , S takes his K-ing at t_2 to be open for him iff at t_1 , S believes that relative to what he himself then believes (i) he would perform K at t_2 were he to undertake K-ing at t_2 and he would refrain from K-ing at t_2 were he to undertake not to K at t_2 , and (ii) his undertaking K at t_2 is, as yet, contingent.

I interpret "the undertaking of some action" as the performance of a basic mental action, such as a decision or a formation of an intention. On this interpretation, Kapitan affirms the following thesis:

(PO*) Rational deliberation about some action K requires an agent S to take K to be open to her in the sense that S must believe that S would perform K were S to decide to K, and S must believe that S would refrain from K-ing if S were to decide to refrain from K-ing.

Randolph Clarke (1992: 103) proposes (but does not endorse) a similar thesis, but with at least one crucial addition. On a causal interpretation of Clarke's thesis,⁴ Clarke also maintains that the rational deliberator must believe that a causal connection obtains between the deliberator's action and her deliberation:

(CF') In order to [rationally] deliberate, an agent must believe that there are at least two distinct actions, *A* and *B*, such that (i) were she to choose to *A* (*B*), she can *A* (*B*) on the basis of that deliberation.

⁴ Pereboom (2014, 118) also interprets Clarke's thesis in this manner.

A further natural interpretation of Clarke’s thesis suggests that the rational deliberator must also believe that which action she performs will be causally connected to her choice or decision, which in turn will be causally connected to her deliberation.

Next, Daniel Dennett (1984, 115) considers a deterministic robot—the “Mark 1 Deterministic Deliberator”—that has various interests, and that can rationally deliberate when it has opportunities open to it. What kinds of opportunities? Here is what Dennett (1984, 118) says:

[A] real opportunity is an occasion where a self-controller “faces”—is informed about—a situation in which the outcome of its subsequent “deliberation” will be a decisive (as we say) factor. In such a situation more than one alternative is “possible” so far as the agent or self-controller is concerned; that is, the critical nexus passes through its deliberation.

In order for a situation to be one in which the outcome of some deliberation is a decisive factor, perhaps the following must be true: ‘if an agent were to believe that she has most reason, all things considered, to ϕ rather than not- ϕ , then she would in fact ϕ .’ Under this interpretation, rational deliberation is incompatible with the belief that one might act akratically, i.e. she might perform an action that she believes she has most reason, all things considered, to refrain from performing. But rational deliberation appears to be compatible with this type of belief. For example, consider a millionaire who deliberates about how much of her annual income to donate to charities, and suppose that she has the following belief: ‘I may donate a percentage of my annual income that is smaller than the obligatory percentage.’ It nevertheless seems possible for such a millionaire to rationally deliberate under these doxastic circumstances.⁵

In order to account for the compatibility of rational deliberation with a belief that one might act akratically, a weaker version of Dennett’s position would suggest that the agent must believe that she is reasons-responsive (Fischer and Ravizza 1998), and thus will at least occasionally act in accordance with what she has most reason to do, all things considered.⁶ Here is one possible formulation of the reasons-responsiveness position:

(RR) In order to rationally deliberate, an agent must believe that her action will issue from a moderately reasons-responsive mechanism as proposed by Fischer and Ravizza (1998).

Next, Dana Nelkin (2011, 142) offers a view that employs Dennett’s concept of a critical nexus:

(EN) Rational deliberators must believe, in virtue of their nature as rational deliberators, that they have multiple alternatives from which to choose, where their deliberation is the *explanatory nexus* among those alternatives.

⁵ A reviewer correctly points out that akratic action is a kind of practical irrationality, since one is going against what one believes one has most reason to do, all things considered. For the purposes of this paper, such irrationality must be sharply distinguished from one who fails to satisfy a doxastic requirement for rational deliberation.

⁶ Fischer (2006, 184–187) himself may welcome such a position.

Recall that I am assuming that the ability to do otherwise is incompatible with determinism. In that case, the notion of having multiple alternatives from which to choose does not track the alternatives associated with the ability to do otherwise. For if these alternatives were associated with the ability to do otherwise, then *EN* would be incompatible with *DC*. So, I will understand the alternatives of deliberation in the following sense:

(*) An agent *S* believes that she has multiple alternatives from which to choose if *S* believes both that she might perform some action, and that she might refrain from performing said action.

According to Nelkin (2011, 142–143), in order for one’s deliberation to be the explanatory nexus among the relevant alternatives, one’s deliberation must be a difference-maker among those alternatives by providing certain contrastive explanations. For instance, an agent’s deliberation concerning what to order while perusing the lunch menu is an explanatory nexus of what she ordered only if her deliberation can explain why she ordered, e.g., an appetizer rather than a salad. I will return to discuss a great virtue of *EN* in a moment.

In addition to endorsing an epistemic openness requirement for rational deliberation, Pereboom (2014a, 108) endorses the following deliberative efficacy requirement for rational deliberation:

(*DE*) In order to rationally deliberate about whether to do A1 or A2, where A1 and A2 are distinct actions, an agent must believe that if as a result of her deliberating about whether to do A1 or A2 she were to judge that it would be best to do A1, then, under normal conditions, she would also, on the basis of this deliberation, do A1; and similarly for A2.

In order to fully appreciate the ingenuity of both *EN* and *DE*, we need to revisit the claim that BAT, in conjunction with the assumption that determinism is incompatible with the ability to do otherwise, entails *DI*. Taylor and van Inwagen each offer a case that is supposed to motivate BAT. Here is Taylor’s (1966, 177) case:

Soldier “[C]onsider a soldier who knows that daily orders regarding the bearing of arms are enforced, and that he has no choice but to obey them. Suppose he does not know whether or not he shall be required to arm himself today, though he knows that the order has been posted.”

Here is an abbreviated version of van Inwagen’s (1983, 154) case:

Two Doors A subject *S* is in a room with two doors. *S* believes that one of the doors is unlocked, and that the other door is locked and impassable. However *S* lacks a belief about which door is locked and about which one is unlocked.

According to the argument for BAT, the soldier cannot deliberate about whether to arm himself, and *S* cannot deliberate about which door to open. The best explanation for this fact is that, in

each case, the agent lacks the belief that she can perform each action under consideration. Moreover, given what is stipulated in each case, if the agent were to *also* believe that she can perform each action under consideration, then that agent would have inconsistent beliefs, and thus rational deliberation is incompatible with the belief that one lacks the ability to do otherwise. The deliberation compatibilist can, however, offer an alternative explanation for why each agent cannot rationally deliberate.

According to Hilary Bok (1998, 111), in *Two Doors* *S*'s deliberation makes no difference to one door opening, rather than the other one opening.⁷ Similarly, in *Soldier* the soldier's deliberation makes no difference with respect to whether the soldier arms himself (assuming that the order is truly enforced). Moreover, these facts are ones to which *S* and the soldier would be rationally committed upon proper reflection. By contrast, when a determinist is deliberating under ordinary doxastic circumstances, she can, upon proper reflection, form the true belief that her deliberation makes a difference with respect to which action she performs (and which decision she makes). So there is an explanation for why the agents cannot rationally deliberate in *Soldier* and *Two Doors* that does not apply to ordinary doxastic scenarios in which a determinist deliberates. This insight is precisely what *EN* and *DE* tap into.

Since *S* believes in *Two Doors* that one of the doors is locked and impassable, *S* does not believe that *S* has multiple alternatives from which to choose, where *S*'s deliberation is the explanatory nexus among those alternatives. In other words, *S*'s deliberation does not explain why one door opened, rather than the other one opening. So *S* does not satisfy *EN*'s condition (Nelkin 2011, 140). Additionally, *S* does not believe that for each door, were *S* to judge that it would be best to open that door, then, under normal conditions (including the actual ones), *S* would in fact open that door. So *S* also does not satisfy *DE*'s condition (Pereboom 2014, 119). Moreover, the exact same lesson applies to *Soldier*. As a result, neither *Soldier* nor *Two Doors* appears to establish BAT, or establish the view that rational deliberation requires one to believe that one can do otherwise. But deliberation compatibilists aren't out of the woods yet.

4. The counterexample

Recall that according to *NIB*, rational deliberation requires the absence of inconsistent beliefs that are salient to one's deliberation. I thus understand all pro-*DC* views to maintain that their proposed requirements for rational deliberation—in conjunction with *NIB* and any other requirement that cuts across the *DC/DI* debate—are necessary and sufficient for rational deliberation (Pereboom 2014a, 108–109). To illustrate, according to a full-fledged pro-*DC* view along the lines of Kapitan's proposal, the conditions embedded in *PO** and *NIB* (along with any other requirement for rational deliberation that cuts across the *DC/DI* debate) are both necessary and sufficient for rational deliberation. Kapitan's view implies that a determinist can satisfy all of the doxastic requirements for rational deliberation, and thus *DC* is true.

⁷ One's deliberation may still explain why some door opened, rather than *that* door not opening.

A deliberation incompatibilist can accept *PO** if she also accepts a requirement for rational deliberation that is inconsistent with *DC*. So rather than offering a counterexample to a position like *PO**, this paper's aim is to offer a counterexample to all recent pro-*DC* views, including ones that affirm *PO**. The putative counterexample involves an agent who satisfies all of the requirements for rational deliberation according to deliberation compatibilists, and yet the agent apparently cannot rationally deliberate about what to do in light of her belief concerning her impending deterministic manipulation. Here are the details.

Case 1 (I)–(VIII) are true.

(I) Betty believes the following. Betty is offered a choice to press one of the two buttons in front of her. If she presses the left button (henceforth 'LEFT'), Betty will receive \$1,000.00. If she presses the right button (henceforth 'RIGHT'), then Oxfam will instead receive \$1,000.00. Betty cannot alter who receives the money once the first button is pressed. Moreover, if Betty presses both buttons simultaneously or presses no button at all, then no one receives the money.

Betty is a U.S. citizen who is financially better off than most people in the world, but is nevertheless burdened with financial debt. She has a strong desire to press LEFT in order to pay off some of her debt. However, Betty believes that donating the money to Oxfam will benefit people who are far worse off than her (Betty knows that she frequently but not exclusively undergoes rationally egoistic tendencies).

(II) Betty believes neither that she will press LEFT, nor that she will press RIGHT. Moreover, all of Betty's beliefs are consistent with the proposition that she will press LEFT, and the proposition that she will press RIGHT. **So Betty satisfies the condition in *EO*.**

(III) Betty believes that if she were to decide to press LEFT (RIGHT), she would in fact press LEFT (RIGHT), and that if she decided to refrain to press LEFT (RIGHT), she would in fact refrain from pressing LEFT (RIGHT). **So Betty satisfies the condition in *PO**.**

(IV) Betty believes that were she to choose to press LEFT (RIGHT), she can press LEFT (RIGHT) on the basis of that deliberation. **So Betty satisfies the condition in *CF'*.**

(V) Betty believes that whichever action she performs will issue from a moderately reasons-responsive mechanism. **So Betty satisfies the condition in *RR*.**

(VI) Betty believes that she has multiple alternatives from which to choose (according to (*)), and that her deliberation is the *explanatory nexus* between the alternatives of deciding to press LEFT and deciding to press RIGHT. **So Betty satisfies the condition in *EN*.**

(VII) Betty believes that if as a result of her deliberating about whether to press LEFT or RIGHT she were to judge that it would be best to press LEFT, then, under normal conditions, she would also, on the basis of this deliberation, press LEFT; and similarly for pressing RIGHT. **So Betty satisfies the condition in *DE*.**

(VIII) Betty believes the following. A team of neuroscientists has the ability to manipulate her neural states at any time by radio-like technology. Prior to Betty’s deliberation, the neuroscientists have decided arbitrarily (on the basis of a coin toss) to causally affect Betty’s imminent decision (cf. Pereboom 2014a, 76–77). As a result, the neuroscientists will manipulate Betty to press (and decide to press) one of the buttons by exerting either an egoism-enhancing or egoism-diminishing momentary influence upon Betty. If they exert a momentary egoism-enhancing influence, then Betty will press LEFT. If they exert a momentary egoism-diminishing influence, then Betty will press RIGHT.

While Betty does not know which kind of influence she will undergo, she believes that the neuroscientists only have the capability of *either* enhancing *or* diminishing Betty’s egoistic tendencies. In other words, if the neuroscientists are capable of diminishing Betty’s egoistic tendencies, then they do not have the capability of enhancing such tendencies (and vice versa).⁸

Finally, the neuroscientists will manipulate Betty’s *decision* (which results from her deliberation) to press one of the buttons. The neuroscientists do not in any way alter Betty’s ultimate *judgment* concerning what she has most reason to do, all things considered. So Betty’s decision will be manipulated by slightly altering Betty’s egoistic tendencies while Betty deliberates in order for her deliberation to generate a different “output” than it might otherwise generate in the absence of such a manipulation.

According to *Case 1*, regardless of what Betty judges, the neuroscientists have already settled what Betty will do (and decide to do). Betty will press RIGHT (donate to Oxfam) if the neuroscientists exert a momentary egoism-diminishing influence, and Betty will press LEFT (receive \$1,000.00) if the neuroscientists exert a momentary egoism-enhancing influence. These four possibilities are illustrated in the following chart:

	Betty judges that she has most reason, all things considered, to press RIGHT (donate to Oxfam)	Betty judges that she has most reason, all things considered, to press LEFT (receive \$1,000.00)
Betty’s egoistic tendencies are momentarily enhanced	Betty presses (and decides to press) LEFT (receive \$1,000.00)	Betty presses (and decides to press) LEFT (receive \$1,000.00)
Betty’s egoistic tendencies are momentarily diminished	Betty presses (and decides to press) RIGHT (donate to Oxfam)	Betty presses (and decides to press) RIGHT (donate to Oxfam)

⁸ This stipulation mirrors Pereboom’s (2014b, 220) recent modification of the FCMA in order for *Cases 1* and *4* of the FCMA to “not feature a relevant difference in causal profiles of nearby possible worlds.” Similarly, this stipulation is meant to establish that there is no relevant difference between *Case 1* and *Case 4* of the FCDA (to be discussed below) with respect to the causal profiles of nearby possible worlds, according to Betty’s beliefs about the actual world.

Notice that (according to Betty's beliefs) the neuroscientists will in fact intervene, even if Betty would perform the same action in the absence of such manipulation. If I were to stipulate instead that the neuroscientists intervene only if, in the absence of manipulation, Betty would not have pressed the neuroscientists' pre-selected button, then *Case 1* would strike a resemblance with Frankfurt-style cases, which in turn would raise numerous vexing issues that are beyond the scope of this paper. For this reason, I maintain that the intervention by the neuroscientists does not depend upon what Betty would do in the absence of manipulation. Moreover, we may also stipulate that Betty has no belief about what she would in fact do in the absence of this manipulation since, according to (I), she believes that she frequently but not exclusively undergoes rationally egoistic tendencies. With this understanding of *Case 1*, let us inspect whether it is a genuine counterexample to the aforementioned pro-DC views.

If Betty cannot rationally deliberate about which button to press, despite the fact that Betty satisfies all of the aforementioned requirements for rational deliberation, then *Case 1* is a counterexample to all recent pro-DC views. The argument may be formulated as follows:

1. In *Case 1*, Betty satisfies the conditions in *EO*, *PO**, *CF'*, *RR*, *EN*, and *DE* with respect to rationally deliberating about which button to press.
2. In *Case 1*, Betty satisfies *NIB*'s condition with respect to rationally deliberating about which button to press.
3. In *Case 1*, Betty cannot rationally deliberate about which button to press.
4. If (1)–(3) are true, then *Case 1* is a counterexample to all recent pro-DC views.
5. Therefore, *Case 1* is a counterexample to all recent pro-DC views.

Since premises (1) and (4) need no defense, let's turn to premise (3). Why think that Betty cannot rationally deliberate about which button to press? It is worth pointing out that Nelkin (2004, 223; 2011, 129)—a deliberation compatibilist—considers a similar case in which one learns that one will be manipulated, and shares the intuition that in such a case one cannot rationally deliberate.⁹ So, my intuitions are arguably not based solely upon a prior, dyed-in-the-wool commitment to *DI*. At any rate, I argue for premise (3) on the basis of the following principle:

Causal Influence Necessarily, if an agent *S* believes the following,

- Either agent *T* will ϕ or *T* will ψ (but *T* will not perform both actions).
- *S* cannot causally contribute to either *T*'s ϕ -ing or *T*'s ψ -ing.
- *T*'s ϕ -ing is (in conjunction with the laws of nature) causally sufficient for the occurrence of event *e*.

⁹ I interpret Nelkin as maintaining that Betty doesn't satisfy *EN*'s condition in *Case 1*. I will defend the contrary position shortly.

- T 's ψ -ing is (in conjunction with the laws of nature) causally sufficient for the non-occurrence of e .

then S cannot rationally deliberate about whether to permit the occurrence of e .

In order to motivate this principle, consider following case. Alex is viewing a live television broadcast of an eight-ball billiards match. The opening move is made, but no balls are pocketed. The next player will strike the cue ball towards the left or towards the right. In conjunction with the laws of nature, striking the cue ball towards the left is causally sufficient for the pocketing of the #14 striped ball, and striking the cue ball towards the right is causally sufficient for the pocketing of the #7 solid ball. Viewing this match from home, Alex cannot causally contribute to the player's next move. Moreover, Alex believes all of this. So Alex cannot rationally deliberate about whether to permit the pocketing of the #14 striped ball or the #7 solid ball. Generalizing from this case, it follows that *Causal Influence* is true.

Given the truth of *Causal Influence*, we can establish premise (3) once we recall what Betty believes according to (VIII). Either the neuroscientists will decide that Betty presses LEFT, or the neuroscientists will decide that Betty presses RIGHT. She cannot causally contribute to the neuroscientists' nefarious activity (recall that their decision is based on a coin toss). The neuroscientists' decision that Betty presses LEFT is (in conjunction with the laws of nature) causally sufficient for the occurrence of Betty's pressing LEFT (and similarly for RIGHT). It thus follows from *Causal Influence* that Betty cannot rationally deliberate about whether to permit the occurrence of Betty's pressing LEFT or Betty's pressing RIGHT. So Betty cannot rationally deliberate about which button to press.

According to premise (2), Betty satisfies *NIB*'s condition. In other words, given what Betty believes according to (I)–(VIII), Betty has no inconsistent beliefs that are salient to her deliberation about which button to press, even upon proper reflection. To begin our evaluation of this premise, we need to look closely at (VIII), and its relation to (I)–(VII).

Betty's beliefs, according to (I) and (VIII), are consistent with one another. The same holds for (II) and (VIII); even though Betty believes that the neuroscientists have already decided which button she will press, Betty has no clue as to which button has been pre-selected. So both the proposition that she will press LEFT and the proposition that she will press RIGHT are consistent with her beliefs according to (VIII). Betty's beliefs, according to (III) and (VIII), are also consistent with one another. Even though Betty believes that her action (and decision) will be manipulated, Betty also *consistently* believes that if she were to decide to press LEFT (RIGHT), she would press LEFT (RIGHT).

The case for the consistency of Betty's beliefs given (IV)–(VIII) rests primarily upon two points. First, as we saw in the discussion on Dennett's position, rational deliberation is consistent with the belief that one may act akratically. Second, rational deliberation is consistent with the belief that one's egoistic tendencies may be momentarily strengthened or reduced, just as one can perform an action for which one is morally responsible under conditions in which one's

egoistic tendencies are momentarily strengthened or reduced (Shabo 2010: 376; Pereboom 2014a, 76).

To illustrate, suppose that Brian is overseeing a children's birthday party at Chuck E. Cheese's, and Brian receives a text from a friend who is asking for help with moving furniture later in the evening. While Brian knows that he will be exhausted by the time the birthday party is over, Brian nevertheless believes that it would be best to help his friend. Additionally, Brian believes that he may undergo a momentary egoism-enhancing influence during his oversight of the birthday party (Brian is not particularly fond of his present circumstances), and thus he may act akratically at least partly due to such an influence. Despite Brian's doxastic circumstances, it seems that he can still rationally deliberate about whether to help his friend. We may now apply these results to Betty's beliefs according to (IV) and (VIII).

On a causal reading of *CF'*, Betty believes that were she to press LEFT, a causal relation would obtain between her deliberation and her decision to press LEFT, as well as between her decision to press LEFT and her overt bodily action of pressing LEFT; and similarly for RIGHT. Now, let's consider this kind of belief in a manipulation-free context, such as the Chuck E. Cheese's example. Suppose that Brian also believes that were he to help his friend, a causal relation would obtain between his deliberation and his decision to help his friend, as well as between his decision to help his friend and the overt action(s) of helping his friend. This belief is consistent both with Brian's belief that he may undergo certain egoism-enhancing influences, and his belief that he may act akratically.

If Betty's beliefs, according to (IV) and (VIII), are inconsistent with one another, then there must be some relevant difference between the doxastic circumstances of Betty and those of Brian. One notable difference concerns their beliefs about the causal source of the potential momentary egoism-enhancing (or egoism-diminishing) influence. Betty believes that the causal source involves the intentions of other agents—the neuroscientists. By contrast, Brian believes that the intentions of other agents are not involved in any direct way with the causal source of his potential momentary egoism-enhancing influence. Nevertheless, the causal source of a momentary influence upon one's egoistic tendencies apparently cannot, by itself, make a difference with respect to whether a causal relation obtains between one's deliberation and a basic mental action, or between a basic mental action and an overt bodily action. For this reason, there is good reason to think that Betty's beliefs according to (IV) and (VIII) are consistent with one another. Let's move on to (V), which concerns *RR*.

The causal source of a momentary influence upon one's egoistic tendencies apparently cannot, by itself, make a difference with respect to whether one's basic mental action issues from a moderately reasons-responsive mechanism. Just as Brian can decide to act akratically through a moderately reasons-responsive mechanism in the presence of a momentary egoism-enhancing influence (due to his oversight of the Chuck E. Cheese's birthday party), the same goes Betty. Hence, satisfying the condition in *RR* does not appear to preclude Betty's satisfying the condition in *NIB*, and thus no inconsistency appears to arise in Betty given (V) and (VIII).

In an attempt to find an inconsistency, a deliberation compatibilist may wish to supplement *RR* in the following manner: in order to rationally deliberate, an agent must believe that the moderately reasons-responsive mechanism through which one acts is “one’s own”. Fischer and Ravizza (1998, 210–214) maintain that moral responsibility requires the mechanism through which one acts to be “one’s own” in the sense that one “takes responsibility” for one’s action. Taking responsibility for one’s action consists of, among other things, seeing oneself as an agent, seeing one’s choices and actions as causally efficacious, and seeing these things on the basis of one’s evidence. Call this revised position *RR+*.

Brian can consistently believe that he will act on a mechanism that is his own and that he may act akratically at least partly due to a momentary egoism-enhancing influence. So why think that the same cannot hold for Betty? Once again, the only notable difference appears to be Betty’s belief concerning the causal source of the potential momentary egoism-enhancing (or diminishing) influence. This belief appears to be consistent with the beliefs that Betty is an agent, that her choices and actions are causally efficacious, and that these beliefs are made on the basis of the evidence available to her. So it appears that Betty’s beliefs according to (VIII) are consistent with the beliefs she would need to possess in order to satisfy the conditions in *RR+* (cf. Pereboom 2014b, 220).

Let us now turn to the relationship between (VI) and (VIII). According to the first part of *EN*, an agent must believe that she has multiple options from which to choose. As previously noted, since I am assuming that determinism is incompatible with the ability to do otherwise, the notion of having multiple options is not understood in a metaphysically robust sense. Instead, an agent takes herself to have multiple options if there is more than one action that she believes she *might* perform. Since Betty believes that she might perform either action, Betty believes that she has multiple options in the sense at issue in *EN*.

Let’s now consider the second part of *EN*. Consider Betty’s belief that her deliberation is the explanatory nexus among certain alternatives, such that Betty’s deliberation explains why she presses LEFT rather than RIGHT (or vice versa). Is this belief inconsistent with Betty’s beliefs according to (VIII)? Perhaps Betty’s deliberation *doesn’t* contrastively explain her decision since the egoism-enhancing/diminishing occurrence is not a component of Betty’s deliberation, precisely because this occurrence was a result of an intervention by other agents. But this suggestion is dubious. Whether an egoism-enhancing/diminishing occurrence is a component of someone’s deliberation cannot depend *merely* upon whether the causal source of that occurrence involves an agent’s intention.

Consider a slightly different objection. While Betty’s deliberation may provide a contrastive explanation, it doesn’t provide the *best* contrastive explanation because it is in some sense trumped by a contrastive explanation that appeals to the neuroscientists’ coin toss. There are two concerns with this objection. First, we need an explanation for why something like a coin toss trumps the contrastive explanation of one’s deliberation, but other factors beyond one’s control that causally determine one’s action don’t trump the contrastive explanation of one’s deliberation. The fact that the coin toss involves an agent does not appear to make a difference.

Instead, all that matters is which causal counterfactuals are true (according to the beliefs of the deliberator).

Second, the very notion of a *better* contrastive explanation may be intelligible given certain pragmatic considerations. For instance, in order to explain the presence of a fire, the striking of a match (rather than its absence) is a better explanation than the presence of oxygen (rather than its absence), given ordinary pragmatic considerations. But such pragmatism is wholly irrelevant to the requirements for rational deliberation. Instead, one's beliefs concerning the metaphysical facts are relevant to rational deliberation, whereby such facts rule out the very idea of a better explanation (cf. Paul & Hall 2013, 35–36). For this reason, Betty's beliefs according to (VI) and (VIII) appear to be consistent with one another.

Finally, let's turn to the relationship between (VII) and (VIII). Recall that Betty believes that if, as a result of her deliberating about which button to press, she were to judge that it would be best to press LEFT, then, under normal conditions, she would also, on the basis of this deliberation, press LEFT; and the same holds for RIGHT. By employing the phrase "under normal conditions" rather than "under the actual conditions", Pereboom (2014a, 119–120) can maintain that an agent's beliefs that satisfy *DE*'s condition are consistent with the belief that one might act akratically. So, whatever else "under normal conditions" concerns, if *S* believes that under the actual conditions *S* might act akratically, then *S* does not believe that *S* is under normal conditions.

Since Betty does believe that she might act akratically, Betty believes that she is not under normal conditions. In that case, Betty's beliefs concerning what she would do under normal conditions are consistent with her beliefs about her impending manipulation, according to (VIII). As Pereboom (2014a, 120) correctly notes, *DE* could be modified in a different manner in order to accommodate the possibility of akrasia. Pereboom's suggestion may be presented as follows:

(DE*) In order to rationally deliberate about whether to do A1 or A2, where A1 and A2 are distinct actions, an agent must believe that if as a result of her deliberating about whether to do A1 or A2 she were to judge that it would be best to do A1, then, *under the actual conditions*, she *might* also, on the basis of this deliberation, do A1; and similarly for A2.

*DE** has the advantage of sidestepping the issue of individuating normal conditions, while simultaneously granting the compatibility of rational deliberation with the belief that one might act akratically. However, since one can satisfy *DE**'s condition and consistently believe that one might act akratically, Betty can satisfy *DE**'s condition and consistently believe what is stipulated in (VIII).¹⁰ This concludes my defense of premise (2), and my articulation of *Case 1* as

¹⁰ As we have seen in our discussion of *CF'*, Betty's beliefs according to (VIII) are consistent with her belief that her decision will be based on her deliberation.

a counterexample to all recent pro-DC views. I will now employ *Case 1* for the purpose of developing the FCDA for deliberation incompatibilism.

5. Completing the four-case deliberation argument

According to *Case 2*, (I)–(VII) are true and (VIII) is false. Instead, (IX) is true:

(IX) Betty believes the following. Long ago, a team of neuroscientists decided arbitrarily (on the basis of a coin toss) which button Betty is to press (and decide to press). As a result, these neuroscientists have programmed Betty at the beginning of her life in such a manner that she will press (and decide to press) one of the buttons, though Betty has no belief about which button the neuroscientists want her to press (cf. Pereboom 2014a, 77).

According to *Case 3*, (I)–(VII) are true and (VIII) is false. Instead, (X) is true:

(X) Betty believes the following. The training practices of Betty's community (which were completed before she developed the ability to prevent or alter these practices) causally determined the nature of her deliberative reasoning processes such that, in conjunction with certain background conditions, Betty is causally determined to press (and decide to press) one of the buttons. Though, Betty has no belief about which button she will in fact press (cf. Pereboom 2014a, 78).

According to *Case 4*, (I)–(VII) are true and (VIII) is false. Instead, (XI) is true:

(XI) Betty believes the following. Everything that happens in the universe is causally determined by its past states together with the laws of nature. Betty is an ordinary human being raised in normal circumstances. As a result, Betty's deliberative reasoning processes, in conjunction with certain background conditions, will causally determine Betty to press (and decide to press) one of the buttons. Though, Betty has no belief about which button she will in fact press (cf. Pereboom 2014a, 79).

Whether Betty believes that the process of manipulation begins a few seconds prior to her decision (in *Case 1*) or at the beginning of her life (in *Case 2*) does not appear to make a difference with respect to Betty's ability to rationally deliberate about which button to press. The same is true with respect to Betty's beliefs in *Case 2* and Betty's beliefs in *Case 3* that Betty's decision to press a certain button is a result of the training practices of Betty's community; and similarly for *Cases 3* and *4*. So in light of the fact that Betty cannot rationally deliberate in *Case 1*, Betty cannot rationally deliberate in any of these four cases. The best explanation for Betty's inability to rationally deliberate in all four cases is that Betty believes that whichever action she performs (and decides to perform) will be causally determined by factors beyond her control.

We have established that rational deliberation requires an agent to lack the belief that her action will be causally determined by factors beyond her control. But there is presumably a more fundamental explanation as to why *DC* is false. Perhaps one must believe that one will be the source of one's action, such that being the source of one's action is incompatible with determinism (Kant 1785, 448; Taylor 1964, 76; Castañeda 1975, 134–135). However, instead of proposing that an agent must believe some proposition *p*, the deliberation incompatibilist can resort to the weaker claim that an agent must *lack* the belief that *not-p*. I thus propose the following requirement for rational deliberation:

Source In order to rationally deliberate about whether to do *A1* or *A2*, where *A1* and *A2* are distinct actions, an agent *S* must not believe that it is not the case that *S* will be the source of whichever action *S* performs, such that being the source of one's action is incompatible with that action being causally determined by factors beyond one's control.

Just as there are different ways to construe sourcehood with respect to the control required for moral responsibility, there are likewise different ways to construe sourcehood with respect to the control one must not believe that one lacks in order to rationally deliberate. In section 7, I suggest that the same notion of sourcehood is relevant to both moral responsibility and rational deliberation. In the next section, we will consider how a deliberation compatibilist might respond to the FCDA.

6. Objections

There are essentially two ways to respond to the argument. According to a so-called hard-line reply, we first inspect *Case 4* in which it seems that Betty can (or might be able to) rationally deliberate about which button to press. Since there is no relevant difference between *Cases 4* and *1*, we should conclude that Betty can (or might be able to) rationally deliberate about which button to press in *Case 1*, despite the truth of (VIII) (cf. McKenna 2008; 2014; Haas 2013). So premise (3) is false, or at least not obviously true. The main concern with this response is that one must deny the *Causal Influence* principle, since this principle entails premise (3).

Recall that *Causal Influence* was motivated by the eight-ball billiards match example. *Causal Influence* renders the correct verdict that Alex cannot rationally deliberate about whether to permit the pocketing of the #14 striped ball or the #7 solid ball. A hard-liner who seeks such an alternative explanation could modify the last part of *Causal Influence* in the following manner: "...then *S* cannot rationally deliberate about whether to permit the occurrence of *e*, unless *e* concerns a decision by *S*." This modified principle can account for Alex's inability to rationally deliberate, and can accommodate the position that Betty can rationally deliberate in *Case 1*. While I find this modification to *Causal Influence* to be *ad hoc*, I don't expect deliberation compatibilists to share this intuition. Breaking this stalemate in the future will require further dialectical maneuvers.

According to a so-called soft-line reply, there is a relevant difference between what Betty believes in some adjacent pair of cases, such that Betty can rationally deliberate in one of these

cases, but not in the other. Here are a few differences one might highlight. First, Betty cannot rationally deliberate in *Cases 1-3* because it is only in *Case 4* that she does not believe that the causal determination of her choice includes in some manner the intentional actions of other agents (cf. Lycan 1997, 115–119). Second, since Betty believes that she is being manipulated in a particularly invasive manner only in *Case 1* (cf. Demetriou 2010; Fischer and Tognazzini 2011, 18–25) or only in *Cases 1-2* (cf. Mele 2006, 141–144), Betty cannot rationally deliberate in these cases.¹¹

Recall that *Cases 1-4* of the FCDA were constructed in such a manner that there are no relevant differences among them with respect to the causal profiles of nearby possible worlds, according to what Betty believes to be the case. So if being manipulated in a particularly invasive manner implies that, in contrast to ordinary deterministic cases, there are relevant differences with respect to the causal profiles of nearby possible worlds (according to what the deliberator believes), then Betty is not being manipulated in a particularly invasive manner in *Case 1*, or in *Case 2*. A soft-liner must thus maintain that an agent's belief about the causal source of an egoism-enhancing/diminishing influence can determine whether she can rationally deliberate. Specifically, if the causal source involves a conscious intention, then rational deliberation is precluded. I have stressed multiple times that if this position is to be taken seriously, we will need some explanation for why a belief about the presence of a conscious intention precludes rational deliberation. For it is counterintuitive to suppose that if Betty were to believe that the neuroscientists in *Case 1* are philosophical zombies who are devoid of consciousness, then Betty *could* rationally deliberate about what to do.

7. Deliberation, responsibility, and control

These objections to the FCDA are similar to objections to Pereboom's FCMA. Just as a hard-line reply to the FCDA maintains that there is no relevant difference with respect to what Betty *believes* in all four cases, a hard-line reply to the FCMA maintains that there is no relevant difference with respect to what *obtains* in all four cases. Similarly, just as a soft-line reply to the FCDA attempts to mark a relevant difference between a pair of adjacent cases with respect to what Betty *believes*, a soft-line reply to the FCMA attempts to mark a relevant difference between a pair of adjacent cases with respect to what *obtains*.

The structural similarity of these objections to each argument is accounted for by the structural similarity of the arguments themselves. In each case of the FCDA, what Betty *believes* to obtain mirrors what *in fact* obtains in each case of the FCMA. For instance, in the FCDA's *Case 1*, Betty *believes* that a team of neuroscientists has the ability to manipulate her neural states at any time by radio-like technology, and that they will manipulate Betty's decision to press one of the buttons. Similarly, in *Case 1* of Pereboom's (2014a, 76–77) FCMA, it *is* the case

¹¹ For further soft-line replies to Pereboom's FCMA that may be appropriated to my argument, see e.g. Haji (2009, 166–168), Barnes (2015), and Schlosser (2015).

that a team of neuroscientists has the ability to manipulate Professor Plum’s neural states at any time by radio-like technology, and that they manipulate Plum to kill (and decide to kill) White.

The structural similarity between both arguments suggests that they stand or fall together; the FCMA is sound if and only if the FCDA is sound. A deliberation incompatibilist can account for this symmetry by maintaining that the same notion of sourcehood is relevant to both moral responsibility and rational deliberation. An agent must be the source of her action in order to be morally responsible for that action. Moreover, according to *Source*, in order to rationally deliberate, an agent must refrain from believing that she is not the source of whichever action she will in fact perform.

An incompatibilist who upholds the principle of alternative possibilities (a so-called “leeway incompatibilist”) can maintain that sourcehood is to be defined at least partly in terms of the possession of alternative possibilities. According to this view, Plum’s non-responsibility is due to the fact that the deterministic manipulation of his behavior removed his possession of alternative possibilities. Similarly, Betty’s inability to rationally deliberate is due to the fact that she believes that her decision will be deterministically manipulated, and such manipulation rules out alternative possibilities.

A deliberation compatibilist who denies that the same kind of control is relevant to both arguments may wish to uphold what I will call the *distinctness strategy*. According to this position, the kind of control that Suzy must not believe that she lacks in order to rationally deliberate (*RD control*) is numerically distinct from the kind of control that Suzy must in fact possess in order to be morally responsible for her actions (*MR control*). Specifically, unlike *RD control*, *MR control* is incompatible with determinism.

In order to think further about the differences between these two kinds of control, we must inspect their relationship to indeterminism. If *MR control* is compatible with indeterminism, then presumably the same holds for *RD control*. Otherwise, the presence of one kind of control implies the absence of the other kind. So let’s assume that a proponent of the *distinctness strategy* wishes to uphold the compatibility of *RD control* with indeterminism. This position is subject to a provisional dilemma.

Suppose that all indeterministic worlds in which one possesses *MR control* (at a specific time) are worlds in which one also possesses *RD control*. In that case, perhaps *RD control* amounts to all of the ingredients that make up *MR control*, minus control over *which* option is performed. An agent has control over which option is performed if and only if it is up to that agent as to which option is performed. This kind of control is sometimes referred to as the power to settle or select among multiple options (Steward 2012; Schlosser 2014). This power is also sometimes reduced to or defined in terms of agent causation (O’Connor 2000; Clarke 2003; Pereboom 2014). According to this approach, without agent causation, *even if* more than one future is nomologically possible, it is still not up to the agent as to which possibility is realized.¹²

¹² Franklin (2016) argues that this motivation for a libertarian agent-causal theory fails, and that the motivation for being an agent-causal theorist cuts across the compatibilism/incompatibilism debate. If Franklin is correct, this will arguably not provide the *distinctness strategy* with any advantage.

So, unlike *RD* control, *MR* control requires agent causation. A surprising implication for this approach is that Suzy can rationally deliberate about which action to perform while simultaneously *knowing* that it is not up to her as to which action she performs.

Suppose instead that there are indeterministic worlds in which one possesses *MR* control, but lacks *RD* control. Since possessing *MR* control implies that one's deliberation is efficacious, what is the missing ingredient that, if present, would bestow the agent with *RD* control? I grant for the sake of argument that an agent can possess *MR* control while failing to satisfy the epistemic openness requirement for rational deliberation.¹³ But the epistemic openness requirement is irrelevant to the kind of control one must not believe that one does not possess in order to rationally deliberate. By contrast, deliberative efficacy is at least one component of *RD* control, even though it also seems to be a component of *MR* control. So, on this horn of the dilemma, we still need to find the missing ingredient for *RD* control that is not also a requirement for *MR* control. If no such ingredient is to be found, we will have further reason to uphold the symmetry between the FCDA and the FCMA.

8. Conclusion

In a deterministic world, seemingly trivial alterations to our egoistic tendencies can guarantee that we will act akratically. Whether such alterations originate in some non-agential phenomenon or the intentions of another agent makes no difference to our responsibility. Similarly, whether we *believe* that such alterations to our egoistic tendencies originate in some non-agential phenomenon or the intentions of other agents makes no difference to our ability to rationally deliberate. So if we conclusively discover in the future that determinism (or something close enough) is true, we must live with the fact that we cannot rationally deliberate since we won't satisfy the requirement for rational deliberation according to *Source*.

It may also turn out that being the source of one's action is incompatible with causal indeterminism in light of considerations about luck and randomness. This opaque notion sourcehood would thus be metaphysically impossible. But there is no need for despair. Even if sourcehood is metaphysically impossible, we will continue to weigh reasons for actions that are viable epistemic possibilities from our point of view, and we will continue to perform actions that we justifiably believe will make the world a better place.¹⁴

References

Barnes, Eric C. 2015. "Freedom, Creativity, and Manipulation." *Noûs* 49 (3): 560–588.

¹³ While I grant this point for the sake of argument, I'm skeptical of the claim that the ability to do otherwise is compatible with future contingents. For this reason, I'm skeptical of the claim that one can *know* what one will do, and nevertheless have the ability to do otherwise.

¹⁴ I am grateful to Eric Barnes, Andrei Buckareff, Janice Dowell, Maddie Eller, John Fischer, Justin Fisher, André Gallois, Mark Heller, David Hershenov, John Keller, Robert Loudon, Kelly McCormick, Julien Murphy, Derk Pereboom, Jason Read, Seth Shabo, Travis Timmerman, Kathleen Winger, audiences from the 2014 Eastern APA, the 2016 *Blameless Buffalo?* conference, the University of Southern Maine, as well as two anonymous referees for this journal for helpful feedback on earlier drafts of this paper.

- Bok, Hilary. 1998. *Freedom and Responsibility*. Princeton, NJ: Princeton University Press.
- Castañeda, Héctor-Neri. 1975. *Thinking and Doing: The Philosophical Foundations of Institutions*. Dordrecht: D. Reidel Publishing.
- Clarke, Randolph. 1992. "Deliberation and Beliefs About One's Abilities." *Pacific Philosophical Quarterly* 73 (2): 101–113.
- _____. 2003. *Libertarian Accounts of Free Will*. New York: Oxford University Press.
- Coffman, E. J., and Ted A. Warfield. 2005. "Deliberation and Metaphysical Freedom." *Midwest Studies in Philosophy* 29 (1): 25–44.
- Cohen, Yishai. 2015. "The Manipulation Argument, At the Very Least, Undermines Classical Compatibilism." *Philosophia* 43 (2): 291–307.
- Demetriou, Kristin. 2010. "The Soft-Line Solution to Pereboom's Four-Case Argument." *Australasian Journal of Philosophy* 88 (4): 595–617.
- Dennett, Daniel. 1984. *Elbow Room*. Cambridge: MIT Press.
- Fischer, John Martin. 2006. *My Way: Essays on Moral Responsibility*. New York: Oxford University Press.
- Fischer, John Marin, and Mark Ravizza. 1998. *Responsibility and Control: A Theory of Moral Responsibility*. Cambridge: Cambridge University Press.
- Fischer, John Martin, and Neal A. Tognazzini. 2011. "The Physiognomy of Responsibility." *Philosophy and Phenomenological Research* 82 (2): 381–417.
- Franklin, Christopher E. 2016. "If Anyone Should Be an Agent-Causalist, then Everyone Should Be an Agent-Causalist." *Mind* 125 (500): 1101–1131.
- Ginet, Carl. 1962. 1966. "Might We Have No Choice?" in *Freedom and Determinism*, ed. Keith Lehrer. New York: Random House, 87–104.
- Haas, Daniel. 2013. "In Defense of Hard-Line Replies to the Multiple-Case Manipulation Argument." *Philosophical Studies*, 163 (3): 797–811.
- Haji, Ishtiyaque. 2009. *Incompatibilism's Allure: Principal Arguments for Incompatibilism*. Peterborough ON: Broadview Press.
- Kant, Immanuel. 1785/1981. *Grounding for the Metaphysics of Morals*. J. Ellington (tr.). Indianapolis: Hackett.
- Kapitan, Tomis. 1996. "Modal Principles in the Metaphysics of Free Will." *Philosophical Perspectives* 10: 419–445.
- Lycan, William G. 1997. *Consciousness*. Cambridge, MA: MIT Press.
- O'Connor, Timothy. 2000. *Persons and Causes: The Metaphysics of Free Will*. New York: Oxford University Press.
- McKenna, Michael. 2008. "A Hard-Line Reply to Pereboom's Four-Case Manipulation Argument." *Philosophy and Phenomenological Research* 77 (1): 142–159.
- _____. 2014. "Resisting the Manipulation Argument: A Hard-Liner Takes it On the Chin." *Philosophy and Phenomenological Research* 89 (2): 467–484.
- Mele, Alfred R. 2006. *Free Will and Luck*. New York: Oxford University Press.
- Nelkin, Dana K. 2004. "Deliberative Alternatives." *Philosophical Topics* 32 (1/2): 215–240.
- _____. 2011. *Making Sense of Freedom and Responsibility*. New York: Oxford University Press.
- Paul, L.A., and Ned Hall. 2013. *Causation: A User's Guide*. New York: Oxford University Press.
- Pereboom, Derk. 2014a. *Free Will, Agency, and Meaning in Life*. New York: Oxford University Press.

- _____. 2014b. "Response to John Martin Fischer and Dana Nelkin." *Science, Religion & Culture* 1 (3): 218–225.
- Schlosser, Markus E. 2014. "The Luck Argument Against Event-Causal Libertarianism: It Is Here to Stay." *Philosophical Studies* 167 (2): 375–385.
- _____. 2015. "Manipulation and the Zygote Argument: Another Reply." *The Journal of Ethics* 19 (1): 73–84.
- Shabo, Seth. 2010. "Uncompromising Source Incompatibilism." *Philosophy and Phenomenological Research*, 80 (2): 349–383.
- Steward, Helen. 2012. *A Metaphysics for Freedom*. New York: Oxford University Press.
- Taylor, Richard. 1964. "Deliberation and Foreknowledge." *American Philosophical Quarterly* 1 (1): 73–80.
- _____. 1966. *Action and Purpose*. Englewood Cliffs, NJ: Prentice-Hall, Inc.
- Todd, Patrick. 2016. "Manipulation Arguments and the Freedom to do Otherwise." *Philosophy and Phenomenological Research*. doi: 10.1111/phpr.12298
- van Inwagen, Peter. 1983. *An Essay on Free Will*. New York: Oxford University Press.