Syracuse University

SURFACE

Dissertations - ALL   SURFACE

May 2016

Incompatibilist Alternative Possibilities

Yishai Cohen
Syracuse University

ABSTRACT

This dissertation has two primary aims. The first aim is to defend the position that the alternative possibilities that are required for basic desert moral responsibility are incompatible with causal determinism. The second aim is to investigate such incompatibilist alternative possibilities, specifically in relation to manipulation, deliberation, and Molinist counterfactuals of freedom.

In chapter 1, "Fischer's Deterministic Frankfurt-Style Argument", I examine John Martin Fischer's deterministic Frankfurt-style argument against the principle of alternative possibilities (PAP). Fischer attempts to show that if causal determinism rules out an agent's moral responsibility, it is not in virtue of its eliminating the agent's alternative possibilities. I contend that, once we focus upon the distinction between entailment and explanation, the incompatibilist defender of PAP can successfully rebut Fischer's argument. I argue for this claim while granting Fischer a number of assumptions that only render a defense of PAP more difficult. Additionally, I cast doubt upon David Palmer's critique of Fischer's argument, which in turn renders my defense of PAP all the more critical.

In chapter 2, "Leeway Compatibilism and Frankfurt-Style Cases", I examine the new dispositionalists' defense of the position that an agent in a deterministic Frankfurt-style case (FSC) has the ability to do otherwise, where that ability is the one at issue in PAP. Focusing specifically on Kadri Vihvelin's proposal, I argue against this position by showing that it is incompatible with the existence of structurally similar cases to FSCs in which a preemptive intervener bestows an agent with an ability.

In chapter 3, "The Manipulation Argument, At the Very Least, Undermines Classical Compatibilism", I argue precisely what the title suggests. More specifically, I argue that 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.

In chapter 4, "Manipulating Deliberators", I offer a counterexample to Derk Pereboom's account of rational deliberation, and then employ it to develop a four-case manipulated deliberation argument against deliberation compatibilism, the view that rational deliberation is compatible with the belief that one's actions are causally determined by factors beyond one's control. Given the structural similarity between my argument and Pereboom's four-case manipulation argument against compatibilism, both arguments stand or fall to together. As a result, a hard incompatibilist such as Pereboom faces a dilemma: either give up deliberation compatibilism, as well as the view that one can (as a hard incompatibilist) rationally deliberate, or give up the four-case manipulation argument.

In chapter 5, "Counterfactuals of Divine Freedom", I argue that, contrary to the commonly held position of Luis de Molina, Thomas Flint and others, counterfactuals of divine freedom (CDFs) are pre-volitional for God within the Molinist framework. That is, CDFs are not true even partly in virtue of some act of God's will. As a result, I argue that the Molinist God fails to satisfy an epistemic openness requirement for rational deliberation, and thus cannot rationally deliberate about which world to actualize.

In chapter 6, "Molinists (Still) Cannot Endorse the Consequence Argument", I defend Ken Perszyk's argument for the claim that Molinists cannot consistently endorse the consequence argument because of a structurally similar argument for the incompatibility of true Molinist counterfactuals of freedom (CCFs) and the ability to do otherwise. More specifically, Edward

Wierenga has argued that on the proper understanding of CCFs, there is a relevant difference between the consequence argument and the anti-Molinist argument. I argue that, even on Wierenga's understanding of CCFs, there is in fact no relevant difference between the two arguments. Moreover, I strengthen Perszyk's argument by highlighting further relevant similarities between CCFs and facts about the laws of nature.

INCOMPATIBILIST ALTERNATIVE POSSIBILITIES

Yishai Abraham Cohen

B.A. Southern Methodist University 2010

A Dissertation

Submitted in Partial Fulfillment of the Requirements

For the Degree of Doctor of Philosophy

Syracuse University

May 2016

ACKNOWLEDGEMENTS

I am first and foremost grateful to my advisor, Mark Heller, for numerous philosophical exchanges on just about every aspect of this dissertation, for his written comments on too many drafts of this dissertation, as well as for his constant encouragement throughout the process. I am also grateful to the other members of my dissertation committee: André Gallois, Janice Dowell, John Martin Fischer, and Derk Pereboom. All of them provided me with invaluable written comments on my dissertation, and also generated an engaging and lively exchange with multiple insightful questions during the dissertation defense.

I should also note that attending Derk's Fall 2013 free will course at Cornell and John's (and Matthew Talbert's) 2015 free will summer school at Moscow State University played a vital role in the development of my philosophical thinking about free will.

Various chapters of my dissertation have also received helpful feedback from a variety of people (in addition to those on my dissertation committee) to whom I am grateful: Kenny Boyce, Sean Clancy, Lorenza D'Angelo, Maddie Eller, Bob Hartman, Sofia Jeppsson, Simon Kittle, Benjamin Matheson, Kelly McCormick, Kris McDaniel, Dan Miller, Michael Rea, Seth Shabo Steve Steward, Matthew Talbert, Travis Timmerman, Patrick Todd, and Edward Wierenga.

Lastly, I wish to thank my parents, Susan and Akiva Cohen, my brother, Elad Cohen, and my partner, Lisa Wanek, for their unremitting love and support.

TABLE OF CONTENTS

1

Fischer's Deterministic Frankfurt-Style Argument

A Frankfurt-style case (*FSC*) is a putative counterexample to the principle of alternative possibilities (Frankfurt 1969):

> **PAP** An agent is morally responsible for what she has done only if she could have done otherwise.

In a traditional *FSC*, Jones ϕ–s on his own, but lacks the ability to avoid ϕ–ing. This inability is at least partly due to a preemptive intervener, Black, who is lurking by, ready to directly alter or manipulate Jones' mental processing if, absent Black's intervention, Jones were about to avoid ϕ–ing. So, according to the Frankfurtian, although Jones cannot avoid ϕ–ing, Jones is nevertheless morally responsible for ϕ–ing since Black in no way causally contributes to Jones' ϕ–ing.

Incompatibilist defenders of *PAP* offer the Dilemma Defense in response (Kane 1985, 51; 1996, 142–145; Widerker 1995a, 1995b; Ginet 1996; Wyma 1997): if causal determinism is true, then although Jones cannot avoid ϕ–ing, it is question-begging to assume that Jones is morally responsible for ϕ–ing. On the other hand, if causal indeterminism (of the right sort) is true, there cannot be a prior sign that infallibly indicates to Black whether Jones will (in the absence of Black's intervention) in fact ϕ. So, although Jones may be morally responsible for ϕ–ing, Jones

could have avoided ϕ–ing. Hence, once *FSC*s are fleshed out in further detail, one can see that they fail to undermine *PAP*.

There have been many ingenious attempts to circumvent the Dilemma Defense.[1] In addressing the Dilemma Defense's deterministic horn, John Martin Fischer (2006, ch. 10) agreed that it would be question-begging against the incompatibilist defender of *PAP* to assume that Jones cannot do otherwise in light of causal determinism, but is nevertheless morally responsible. Instead, Fischer attempted to argue that Black's presence, device, and dispositions in themselves and apart from causal determinism render Jones unable to avoid ϕ–ing. If this claim were correct, then *PAP* would appear to be false. For, Black's presence, device, and dispositions in no way causally contribute to Jones' ϕ–ing, and thus appear to be irrelevant to Jones' moral responsibility for ϕ–ing.

However, Stewart Goetz (2005, 85) has highlighted the fact that unless Jones' ϕ–ing is *deterministically* caused by factors beyond his control, there cannot be a prior sign that infallibly indicates to Black whether Jones will (in the absence of Black's intervention) in fact ϕ. So it is false that Black's presence, device, and dispositions *in themselves and apart from causal determinism* render Jones unable to avoid ϕ–ing. Fischer (2010, 323–325) now concedes Goetz's point, and thus attempts to construct an improved deterministic Frankfurt-style argument against *PAP* (henceforth '*Fischer's Improved Argument*') that also doesn't assume that Black's presence, device, and dispositions *in themselves and apart from causal determinism* render Jones unable to avoid ϕ–ing. More specifically, Fischer attempts to refute *PAP* by showing that if

---

[1] Frankfurt-style arguments that take on the indeterministic horn of the Dilemma Defense include (but are not limited to) Stump (1996; 2003), Haji (1998), Hunt (2000; 2003), McKenna (2003), Mele and Robb (1998; 2003), and Pereboom (2003; 2014, ch. 1). Like Fischer, Funkhouser (2009) appeals to a deterministic *FSC* against *PAP*. I, of course, can only address Fischer's argument in this paper. Hence, my aim is not to show that *PAP* has not been refuted. Rather, I only aim to defend the incompatibilist defender of *PAP* against Fischer's argument.

causal determinism rules out an agent's moral responsibility, it is not in virtue of its eliminating the agent's alternative possibilities.

I contend that, once we focus upon the distinction between entailment and explanation, the incompatibilist defender of *PAP* can successfully rebut *Fischer's Improved Argument*. I argue for this claim while granting Fischer a number of assumptions that only render a defense of *PAP* more difficult. Additionally, in the appendix I cast doubt upon David Palmer's (2014) critique of *Fischer's Improved Argument*, which in turn renders my defense of *PAP* all the more critical.[2] Let's now turn to *Fischer's Improved Argument*.

1.1   Fischer's Improved Argument

Consider Fischer's deterministic *FSC*, which I'll refer to as ***The Voting Case***:

> [Black] has secretly inserted a chip in Jones's brain which enables Black to monitor and control Jones's activities. Black can exercise this control through a sophisticated computer that he has programmed so that, among other things, it monitors Jones's voting behaviour. If Jones were to show any inclination to vote for McCain (or, let us say, anyone other than Obama), then the computer through the chip in Jones's brain, would intervene to assure that he actually decides to vote for Obama and does so vote. But if Jones decides on his own to vote for Obama (as Black, the old progressive, would prefer), the computer does nothing but continue to monitor—without affecting—the goings-on in Jones's head (Fischer 2010: 316).

---

[2] With the exception of footnote 7, I will not discuss Widerker and Goetz's (2013) critique of *Fischer's Improved Argument*, to which Fischer (2013) has replied.

Jones decides "on his own" at t2 to vote for Obama [at t3]; and imagine that Jones involuntarily exhibits some sign, such as (appropriately enough) a furrowed left brow, at an earlier time t1 (Fischer 2010, 319).

With *The Voting Case* in mind, **Fischer's Improved Argument** against *PAP* goes as follows:

1.   Assume that causal determinism obtains and that the Frankfurt case of Jones and Black unfolds as above.

2.   At this point in the argument, causal determinism is not assumed in itself to rule out access to alternative possibilities. (Neither is it to be supposed here that Black's presence, device, and dispositions in themselves rule out such access.)

3.   Causal determinism plus Black's presence, device, and dispositions rule out Jones's freedom at t2 to choose otherwise.

4.   If Jones is not morally responsible for choosing at t2 to vote for Obama at t3, it is not in virtue of the *mere* fact that he was not free at t2 to choose otherwise.

Thus:

5.   If causal determinism rules out Jones's moral responsibility for his choice at t2, it is *not* in virtue of its eliminating alternative possibilities (if in fact it does eliminate alternative possibilities) (Fischer 2010, 328).

Fischer (2010, 335) thinks that if (5) is true, then *PAP* appears to be false. The idea here is that if *PAP* is true, and *if* determinism eliminates Jones' alternative possibilities, then determinism *would* rule out Jones' moral responsibility in virtue of its eliminating such alternatives. For this

reason, if (5) is true, then *PAP* is false. I grant Fischer this point for the sake of argument.[3] I also grant the inference from (4) to (5). After all, if Jones' lacking some feature *f* is irrelevant to Jones' moral responsibility for his voting decision, then the mere fact that determinism eliminates *f* surely cannot imply that determinism renders Jones blameless for his voting decision.

The real question, then, is why Fischer accepts (4). The short answer is that (4) may be inferred from (3). My criticism of *Fischer's Improved Argument* consists of a detailed assessment of this inference. To give a preview of what is to come, I will offer two interpretations of (3). One interpretation invokes the notion of entailment, while the other invokes the notion of explanation. Only the truth of (3) under the latter interpretation warrants an inference to (4). However, within the context of arguing against an incompatibilist defender of *PAP*, Fischer is not rationally permitted to affirm the truth of (3) under the latter interpretation, or so I will argue.

Now, I said that I am granting Fischer a number of assumptions that only render a defense of *PAP* more difficult. I now turn to the first assumption. According to Widerker (1995b, 248) the opponent of *PAP* must demonstrate the truth of the following thesis:

---

[3] Since (5) employs an 'in virtue of' locution, and *PAP* does not (at least as both Fischer (2010) and I articulate the principle), one might think that the truth of (5) *only* undermines the following strong reading of *PAP* as extensively discussed by Leon and Tognazzini (2010, 558):

> (*PAP$_{cc}$*) A conceptually necessary constituent of being morally responsible is having the ability to do otherwise.

I am, however, willing to concede for the sake of argument that if *PAP$_{cc}$* is false, then so is *PAP* as both Fischer (2010) and I characterize the principle. For further discussion of these different readings of *PAP*, see also Fischer (1994, ch. 7), Della Rocca (1998), and Funkhouser (2009).

IRR There may be circumstances in which a person performs some action which although they make it impossible for him to avoid performing that action, they in no way bring it about that he performs it.

Fischer (2010, 333–335) thinks that even if X causally contributes to (or brings about) Jones' voting decision, so long as X is irrelevant to Jones' moral responsibility for his voting decision *insofar as X plays some role in ruling out Jones's ability to do otherwise*, then *PAP* is false. So, contrary to Widerker, the opponent of *PAP* need not demonstrate the truth of *IRR*. This is the first assumption I grant Fischer. In order to be perfectly clear about this, throughout the paper I will invoke the term 'irrelevance*' which is to be understood as follows:

X is **irrelevant*** to an agent S's moral responsibility for ϕ–ing *iff* X is irrelevant to S's moral responsibility for ϕ–ing *insofar as X plays some role in ruling out S's ability to avoid ϕ–ing.*[4]

The other assumption I grant Fischer is that an overdetermination of explanation for some proposition is possible. But in order to see *why* this assumption renders a defense of *PAP* more difficult, we must wait until section 6.

In the next section, I will offer two interpretations of premise (3). But before proceeding, it is important to remember that I am only concerned with rebutting *Fischer's Improved Argument* on behalf of the incompatibilist defender of *PAP*. For, as will become evident later on, my critique

---

[4] The careful reader will see that the term 'X' is being used loosely in this discussion to refer to either a proposition or an event. If you think propositions cannot "play a role in ruling out an agent's ability", I invite the reader to think instead of a proposition's playing some explanatory role with respect to an agent's lacking an ability.

of *Fischer's Improved Argument* arguably won't help a compatibilist defender of *PAP*. Let's proceed now to the two possible interpretations of premise (3).

1.2   Two Avenues to Premise (4)

I will employ abbreviations for the following propositions (a sentence enclosed in brackets is a proposition):

[Jones cannot do otherwise] $=_{df}$ [Jones is unable at t2 to choose otherwise]

[Determinism] $=_{df}$ [Deterministic laws L obtain, and Jones' choice at t2 is causally determined by factors beyond his control]

[Black is present] $=_{df}$ [Black is present, Black's device is present, and Black has certain dispositions]

Here, then, are the two possible interpretations of (3) that I offer:

3i.   [[Determinism] and [Black is present]] *entails* [Jones cannot do otherwise].

3ii.   [[Determinism] and [Black is present]] *explains* [Jones cannot do otherwise].

Before going into the details behind Fischer's move from (3) to (4), some clarifying remarks on the distinction between (3i) and (3ii) are in order. I am going to maintain the virtually unanimous position that even if *P* entails *Q*, it doesn't necessarily follow that *P* even partly explains *Q*. To demonstrate this point with just one example, every single contingent fact about

The Beatles entails, but does not even partly explain, [2+2=4]. As a result, we should not immediately infer (3ii) from (3i) without further argument.

Additionally, I will not adopt any particular account of explanation for the purposes of this paper. This is a virtue of my critique of *Fischer's Improved Argument*, rather than a defect. For, I will be arguing that, in order to establish (4), Fischer must show us that some proposition at least partly explains [Jones cannot do otherwise], while simultaneously being irrelevant* to Jones' moral responsibility for his voting decision. So, *whatever* account of explanation one thinks Fischer must employ in order to establish (4) in the above manner, I contend that my objection to Fischer's argument for (4) will go through.

Let us now look into the details of Fischer's inference from (3) to (4). Fischer (2010: 330) claims to move from (3) to (4) by holding the following crucial claim:

C. The fact that Black's device (and dispositions) in a causally deterministic context rule out Jones's freedom to choose and do otherwise is irrelevant to Jones's moral responsibility.

According to (C), (3) is true, and the truth of (3) is irrelevant to Jones' moral responsibility for his voting decision. Although it is not made explicit in (C), as previously noted, Fischer (2010, 333–335) thinks that in order to move from (3) to (4), Fischer must only show that the truth of (3) is *irrelevant** to Jones' moral responsibility. I will thus understand (C) in that manner.

Notice that Black (and his device and dispositions) in a causally deterministic context is supposed to be irrelevant* to Jones' moral responsibility since Black (and his device and dispositions) in no way causally contributes to Jones' decision at t2 to vote for Obama. This line

of reasoning appears to rely upon the following principle: if a fact is irrelevant to the causal explanation of a person's action, then that fact is irrelevant to *whether or not* that person is morally responsible for her action. I grant the truth of this principle.[5]

Now, since there are two possible interpretations of (3), and since (C) presupposes the truth of (3), there are likewise two possible interpretations of (C):

Ci. The truth of (3i) is irrelevant* to Jones' moral responsibility for his voting decision.

Cii. The truth of (3ii) is irrelevant* to Jones' moral responsibility for his voting decision.

We thus have two potential avenues to the truth of (4):

The First Avenue

3i. [[Determinism] and [Black is present]] *entails* [Jones cannot do otherwise].

Ci. The truth of (3i) is irrelevant* to Jones' moral responsibility for his voting decision.

Thus:

4. If Jones is not morally responsible for choosing at t2 to vote for Obama at t3, it is not in virtue of the *mere* fact that he was not free at t2 to choose otherwise.

---

[5] In the appendix, I defend this principle against an alleged counterexample by Palmer (2014). Notice, moreover, that if principle is true, then a restricted form of this principle that employs the notion of irrelevance* rather than irrelevance is also true.

The Second Avenue

3ii.   [[Determinism] and [Black is present]] *explains* [Jones cannot do otherwise].

Cii.   The truth of (3ii) is irrelevant* to Jones' moral responsibility for his voting

decision.

Thus:

4.   If Jones is not morally responsible for choosing at t2 to vote for Obama at t3, it is

not in virtue of the *mere* fact that he was not free at t2 to choose otherwise.

Let us now turn our attention to *The First Avenue*.

1.3   The First Avenue

Let's begin by assessing the truth of each premise of *The First Avenue*. I grant that there is good

reason to accept the truth of (3i), as I will now explain. Recall that, according to (2), we are to

remain agnostic about the truth of the following proposition:

(^) Causal determinism rules out access to alternative possibilities.

So, given (2), Fischer is rationally permitted to affirm the truth of (3i) only if its truth does not

depend upon the truth of (^). Fortunately for Fischer, (3i) appears to be true *irrespective* of

whether (^) is true, as I will now illustrate.

If (^) is true, then [Jones cannot do otherwise] is entailed by [Determinism] in conjunction

with *any* other true proposition, including [Black is present]. So if (^) is true, so is (3i). On the

other hand, if (^) is false, then [Determinism] does not by itself entail [Jones cannot do

otherwise]. However, I agree with Fischer that [[Determinism] and [Black is present]] entails [Jones cannot do otherwise]. One way to motivate this claim is by conceiving of an agent's abilities in terms of her access to possible worlds, as I will now explain.

If an agent *S* who ϕ–s had the ability to avoid ϕ–ing, then there is a sufficiently nearby world *w* in which *S* does not ϕ, such that *w* was accessible to *S* (Lehrer 1976, 253–254). Accordingly, if (^) is false, then in a deterministic world *S* has access to worlds with a different past or different laws of nature (or both), at least if such worlds are sufficiently close to the world at which *S* is located. However, given [Determinism] *and* [Black is present], the closest world in which Jones does otherwise is, we may suppose, a world in which Black is absent (or Black's device is absent or defective, or Black has different dispositions). But *that* world is presumably sufficiently far such that it is *not* accessible to Jones. So Jones cannot do otherwise, *even if* (^) is false.

To be clear, my intention is not to commit myself to an 'accessibility to worlds' account of abilities. Rather, I am only suggesting one natural way to assess the position that (3i) is true even if (^) is false. Moreover, I acknowledge that this position is contested by the so-called new dispositionalists (Smith 2003; Fara 2005; 2008; Vihvelin 2000; 2004; 2008).[6] Notice, however, that by siding with Fischer (2008) on this position, I am by no means making matters easier for the incompatibilist defender of *PAP*. So, assuming for the sake of argument that Jones cannot do otherwise, *even if* (^) is false, we may conclude that (3i) is true *irrespective* of whether (^) is true. For this reason, Fischer is rationally permitted to affirm the truth of (3i), even given (2).[7]

Let's turn now to (Ci). I will assume that if (3i) true, then the truth of (3i) is irrelevant* to Jones' moral responsibility for his voting decision if the conjunctive proposition [[Determinism]

---

[6] For a response to the new dispositionalists on this issue, see Clarke (2009), Whittle (2010), and Franklin (2011a). For a response specifically to Vihvelin, see Fischer (2008) and Kittle (2014).

[7] Widerker and Goetz (2013) may be interpreted as rejecting (3i) given (2). For a response to Widerker and Goetz, see Fischer (2013). To repeat, I am willing to side with Fischer on this point for the sake of argument.

and [Black is present]] is irrelevant* to Jones' moral responsibility for his voting decision. But what does it take for a conjunctive proposition to be irrelevant* to Jones' moral responsibility? More specifically, is [[Determinism] and [Black is present]] irrelevant* to Jones' moral responsibility if at least one of its conjuncts is irrelevant* to Jones' moral responsibility? Well, if *The First Avenue* has any hope of being non-question-begging against the incompatibilist defender of *PAP*, it would appear so. For, if *both* conjuncts need to be irrelevant* to Jones' moral responsibility for his voting decision in order for (Ci) to be true, then, *a fortiori*, [Determinism] needs to be irrelevant* to Jones' moral responsibility. But it is question-begging against the incompatibilist defender of *PAP* to assume that [Determinism] is irrelevant* to Jones' moral responsibility. After all, if this assumption were not question-begging, then *FSC*s would be utterly superfluous to refuting *PAP*; Fischer could simply establish (5) by pointing out that [Determinism] at least partly explains Jones' inability to do otherwise while simultaneously being irrelevant* to Jones' moral responsibility.

So, in order for *The First Avenue* to be non-question-begging against the incompatibilist defender of *PAP*, we must assume that [[Determinism] and [Black is present]] is irrelevant* to Jones' moral responsibility for his voting decision if at least one of its conjuncts is irrelevant* to Jones' moral responsibility. In other words, we should understand the truth conditions for (Ci) according to the following principle:

> ***The Irrelevance*-Responsibility Principle* (IRP)** If at least one conjunct of a conjunctive proposition P is irrelevant* to an agent S's moral responsibility for ϕ–ing, then P is irrelevant* to S's moral responsibility for ϕ–ing.

Fortunately, [Black is present] *is* plausibly irrelevant* to Jones' moral responsibility precisely because Black in no way causally contributes to Jones' voting decision. So, given the truth of (3i), and given *IRP,* there is good reason to accept (Ci).[8]

So, since there is good reason to affirm both premises of *The First Avenue* (at least given *IRP*), should we thereby affirm (4) on the basis of these premises? No. For, as I will now explain, since *The First Avenue* does not invoke the notion of explanation, there is good reason to think that it is invalid.

Consider once again *The Voting Case,* except that Black and his device and dispositions are deleted. Moreover, suppose that on the other side of town, Lisa is giving a lecture to her class. Call this **The Lecture Case**. Now, like *The Voting Case*, we are supposing that causal determinism obtains in *The Lecture Case*. Let us also suppose that (^) is true.[9] Consequently, [Determinism] entails [Jones cannot do otherwise]. *A fortiori*, [Jones cannot do otherwise] is entailed by the following conjunctive proposition: [[Determinism] and [Lisa is lecturing]]. Moreover, since [Lisa is lecturing] is clearly irrelevant* to Jones' moral responsibility for his voting decision, it follows from *IRP* that [[Determinism] and [Lisa is lecturing]] is irrelevant* to Jones' moral responsibility. So, we have an argument for (4) that is structurally similar to *The First Avenue*:

---

[8] One may have reason to doubt the truth of *IRP*. However, as I have just explained, in order for *The First Avenue* to be non-question-begging against the incompatibilist defender of *PAP*, *The First Avenue* requires the truth of *IRP*. I am thus willing to grant the truth of *IRP* for the sake of argument, regardless of what doubts one may have about its truth.

[9] Fischer initially does not presuppose the truth of (^) in *The Voting Case*. It does not follow, however, that it is problematic for me to assume the truth of (^) in *The Lecture Case* for the purpose of showing that the notion of entailment (rather than the notion of explanation) is not the right tool for establishing (4). This issue will be further discussed at the end of this section. Moreover, as we will see in section 6, Fischer himself thinks (and I agree) that it *is* dialectically appropriate to assume the truth of (^) in *The Voting Case*.

The Bad Argument

3i*.   [[Determinism] and [Lisa is lecturing]] *entails* [Jones cannot do otherwise].

Ci*.   The truth of (3i*) is irrelevant* to Jones' moral responsibility for his voting

decision.

Thus:

4.   If Jones is not morally responsible for choosing at t2 to vote for Obama at t3, it is

not in virtue of the *mere* fact that he was not free at t2 to choose otherwise.

Clearly, *The Bad Argument* is invalid. The best explanation for this is that, while [Lisa is

lecturing] is no doubt irrelevant* to Jones' moral responsibility for his voting decision, [Lisa is

lecturing] does not even partly *explain* [Jones cannot do otherwise]. What the opponent of *PAP*

needs to do, then, is replace [Lisa is lecturing] with a proposition that at least partly *explains*

[Jones cannot do otherwise] while simultaneously being irrelevant* to Jones' moral

responsibility for his voting decision. More specifically, in order to establish (4), Fischer must

establish the truth of the following thesis:

> ***The Thesis*** There is some proposition *P* that at least partly explains [Jones cannot do
>
> otherwise] while simultaneously being irrelevant* to Jones' moral responsibility for his
>
> voting decision in *The Voting Case*.

In other words, *The Thesis* should be included as a premise of *The First Avenue*, whereby (4)

follows from *The Thesis*, and *The Thesis* follows from (3i) and (Ci). The same holds, *mutatis*

mutandis, for *The Second Avenue*.

So, even though (3i) and (Ci) appear to be true (given *IRP*), these premises don't establish *The Thesis*, just as (3i*) and (Ci*) don't establish *The Thesis*. Hence, *The First Avenue* doesn't establish the truth of (4). I will now consider an important objection to the conclusions that have been reached so far.

Recall that I assumed the truth of (^) in *The Lecture Case*. Hence, one might think that since we are not assuming the truth of (^) in *The Voting Case,* one can consistently maintain that [Jones cannot do otherwise] is not even partly explained by [Lisa is lecturing] in *The Lecture Case*, but *is* at least partly explained by [Black is present] in *The Voting Case*. The thought here is that since [Jones cannot do otherwise] is entailed by [Determinism] in *The Lecture Case* but not in *The Voting Case*, [Black is present] must be doing some work in explaining [Jones cannot do otherwise], while [Lisa is lecturing] is doing no such work. So, (3i) and (Ci) do establish *The Thesis* after all.

The critical flaw in the above line of reasoning is the assumption that [Determinism] doesn't entail [Jones cannot do otherwise] in *The Voting Case*. We are simply not entitled to this assumption. For, recall that in *The Voting Case* we are to remain *agnostic* about the truth of (^), rather than *deny* its truth. Consequently, we must remain *agnostic* about whether or not [Determinism] entails [Jones cannot do otherwise] in *The Voting Case*. Fischer could of course replace (2) with the falsity of (^) in an attempt to establish *The Thesis*. I grant that such a move may succeed against the *compatibilist* defender of *PAP*. But such a move is dialectically inappropriate within the context of arguing against the *incompatibilist* defender of *PAP*. This concludes my assessment of *The First Avenue*. Let's now turn to the second one.

1.4   The Second Avenue

I have argued that in order to arrive at (4), Fischer must establish *The Thesis*. So, *The Thesis* should be viewed as a premise in *The First Avenue*, which is to be inferred from (3i) and (Ci). Similarly, *The Thesis* should be viewed as a premise in *The Second Avenue*, which is to be inferred from (3ii) and (Cii).

Since, (3i) only invokes the notion of entailment, *The Thesis* cannot be inferred from (3i) and (Ci). By contrast, since premise (3ii) of *The Second Avenue* invokes the notion of explanation, perhaps (3ii) and (Cii) can establish *The Thesis*, which in turn would establish the truth of (4). Let's begin, then, by assessing the truth of (Cii).

Similar to my remarks on *The First Avenue*, I will assume that if (3ii) true, then (3ii) is irrelevant* to Jones' moral responsibility for his voting decision if the conjunctive proposition [[Determinism] and [Black is present]] is irrelevant* to Jones' moral responsibility for his voting decision. So, similar to my remarks on (Ci), we must understand the truth conditions for (Cii) according to *IRP*: if (3ii) is true, and if at least one of the conjuncts of [[Determinism] and [Black is present]] is irrelevant* to Jones' moral responsibility for his voting decision, then (Cii) is true. Moreover, I have conceded that [Black is present] is irrelevant* to Jones' moral responsibility. So, I grant that (Cii) is true *if* (3ii) is true. So, let's turn now to (3ii).

Given that we accept the truth conditions for (Cii) according to *IRP*, a constraint must be placed on the truth conditions for (3ii) according to the following principle:

> **The Non-Monotonicity-Explanation Principle (NMEP)** In order for some conjunctive proposition *P* to even partly explain some further proposition *Q*, each conjunct of *P* must at least partly explain *Q*.

Besides its intuitive plausibility (Cf. Rosen 2010, 116–117), why must we accept the truth conditions for (3ii) according to *NMEP*? Well, suppose we instead held that (3ii) is true if at least one of the conjuncts of [[Determinism] and [Black is present]] at least partly explains [Jones cannot do otherwise]. We know that, irrespective of whether (^) is true, [Determinism] at least partly explains [Jones cannot do otherwise]. So, (3ii) would undoubtedly be true if we rejected *NMEP*. So, if we rejected *NMEP*, then both premises of *The Second Avenue* would arguably be true. But notice that the incompatibilist defender of *PAP* would not thereby have any reason to accept *The Thesis*, and thus would not have any reason to accept (4), as I will now explain.

Recall once more that it is question-begging against the incompatibilist defender of *PAP* to assume that [Determinism] is irrelevant* to Jones' moral responsibility for his voting decision, but it is *not* question-begging to assume that [Black is present] is irrelevant* to Jones' moral responsibility.[10] In that case, in order to establish *The Thesis*, Fischer must establish that [Black is present] at least partly explains [Jones cannot do otherwise]. But if we reject *NMEP*, then (3ii) can be (and arguably is) true without [Black is present] even partly explaining [Jones cannot do otherwise]. For this reason, in order for *The Second Avenue* to be non-question-begging against the incompatibilist defender of *PAP*, we must accept *NMEP*. So, given *NMEP*, the truth of (3ii)—and ultimately the success of *Fischer's Improved Argument*—hinges upon whether [Black is present] at least partly explains [Jones cannot do otherwise].

I won't argue that [Black is present] does not even partly explain [Jones cannot do otherwise] in *The Voting Case*. Instead, I will argue that the proponent of *Fischer's Improved Argument* is not rationally permitted to hold that [Black is present] even partly explains [Jones cannot do otherwise] in *The Voting Case*. If I am correct, then Fischer is not rationally permitted to affirm

---

[10] This is because Black in no way causally contributes to Jones' voting decision.

the truth of (3ii), and thus is not rationally permitted to affirm the truth of either *The Thesis* or (4). In that case, the incompatibilist defender of *PAP* has been given no reason to give up her position. My argument is as follows:

6. If (^) is true, then [Black is present] does not fully explain [Jones cannot do otherwise] in *The Voting Case*.

7. If (^) is true, then [Black is present] does not partly explain [Jones cannot do otherwise] in *The Voting Case*.

8. If (^) is true, then [Black is present] does not even partly explain [Jones cannot do otherwise] in *The Voting Case*. [(6), (7)]

9. If both (2) and (8) are true, then the proponent of *Fischer's Improved Argument* is not rationally permitted to hold that [Black is present] even partly explains [Jones cannot do otherwise] in *The Voting Case*.

Thus:

10. The proponent of *Fischer's Improved Argument* is not rationally permitted to hold that [Black is present] even partly explains [Jones cannot do otherwise] in *The Voting Case*. [(2), (8), (9)]

Working backwards, let's consider premise (9). Suppose I know the following: if *P* is true then *Q* is false, and I am (and should be) agnostic about the truth of *P*. Consequently, I am not rationally permitted to affirm the truth of *Q*. Similarly, if I am to remain agnostic about the truth of (^), and I know that *if* (^) is true, then [Black is present] does not even partly explain [Jones cannot do otherwise] in *The Voting Case*, then I am not rationally permitted to hold that [Black is

present] even partly explains [Jones cannot do otherwise] in *The Voting Case*. So premise (9) is very plausible.[11] Next, since (8) follows from (6) and (7), one must reject either (6) or (7) in order to reject (10).

No one should reject premise (6). Remember that Fischer is not arguing for the claim that [Black is present] *alone* rules out Jones' ability to do otherwise. For, as noted in the first section, Fischer (2010, 323–325) concedes Goetz's (2005, 85) point that unless Jones' ϕ–ing is *deterministically* caused by factors beyond his control, there cannot be a prior sign that infallibly indicates to Black whether Jones will (in the absence of Black's intervention) in fact ϕ. In that case, it is safe to assume that, irrespective of the truth of (^), [Black is present] does not fully explain [Jones cannot do otherwise]. *A fortiori*, (6) is true.

The crux of my critique of *Fischer's Improved Argument*, then, hinges upon the truth of (7). As a first step towards establishing its truth, I contend that we should accept the following principle:

> **The Entailment-Explanation Principle (EEP)** If proposition *P* partly explains proposition *Z*, and the following conditions obtain,
>
> **i.**   *P* and *Z* are both contingently true.
>
> **ii.**   There is no overdetermination of explanation of *Z*.
>
> **iii.**   *P* does not entail *Z*.
>
> then the further proposition *Q* that, in conjunction with *P*, fully explains *Z*, is such that *Q* does not entail *Z*.

---

[11] While denying some epistemic closure principle (or employing some other related epistemic move) may be appropriate when it comes to rejecting external world skepticism, I assume that no such move is appropriate for the purposes of rejecting (9).

The irony is not lost upon me that I am appealing to *EEP* while simultaneously emphasizing the importance of distinguishing between entailment and explanation. However, distinguishing between entailment and explanation is perfectly consistent with acknowledging certain connections between them, including the rather weak connection embodied in *EEP*. Now, I offer two cases in support of *EEP*.

First, let us consider whether *EEP* renders the correct verdict that [Rafael is reading] does not partly explain [The singleton set {Edna} exists]. Suppose there is no overdetermination of explanation of [The singleton set {Edna} exists]. In that case, condition (ii) of *EEP* is satisfied. It is safe to assume that conditions (i) and (iii) are likewise satisfied.[12] Moreover, let us rightly suppose that [The singleton set {Edna} exists] is at least partly explained by [Edna exists]. Now, according to *EEP*, if [Rafael is reading] partly explains [The singleton set {Edna} exists], then [Edna exists] must not entail [The singleton set {Edna} exists]. But it does. So, *EEP* renders the correct verdict that [Rafael is reading] does not partly explain [The singleton set {Edna} exists].

Second, let us consider whether *EEP* renders the correct verdict that [Lisa is lecturing] does not partly explain [Jones cannot do otherwise] in *The Lecture Case*. So, let us once again assume the truth of (^), and also explicitly stipulate that there is no overdetermination of explanation of [Jones cannot do otherwise]. In that case, it is safe to assume that conditions (i)–(iii) of *EEP* are satisfied. Now, according to *EEP*, if [Lisa is lecturing] partly explains [Jones cannot do otherwise], then whatever further proposition that, in conjunction with [Lisa is lecturing], fully explains [Jones cannot do otherwise], that proposition must not entail [Jones cannot do otherwise]. Plausibly, the proposition that, in conjunction with [Lisa is lecturing], fully explains [Jones cannot do otherwise] is [Determinism]. But since [Determinism] *does* entail [Jones cannot

---

[12] If one accepts Williamson's (2002) position that everything necessarily exists, I ask the reader to consider whether *EEP* renders the correct verdict in this case under the assumption that Williamson's position is false.

do otherwise], *EEP* thus renders the correct verdict that [Lisa is lecturing] does not partly explain [Jones cannot do otherwise].

I have offered two cases in support of *EEP*. Many more could be given. I thus hope to have shown that, in the absence of counterevidence, we should accept this principle. Now, the truth of *EEP* is significant since, as I will now show, it establishes the truth of (7).[13]

Suppose that (^) is true in *The Voting Case*.[14] Now, with respect to the question of whether [Black is present] partly explains [Jones cannot do otherwise], it is safe to assume that conditions (i) and (iii) of *EEP* are satisfied. Moreover, given (6), we may also explicitly stipulate that there is no overdetermination of explanation of [Jones cannot do otherwise].[15] So condition (ii) of *EEP* is also satisfied. Now, according to *EEP*, if [Black is present] partly explains [Jones cannot do otherwise], then whatever further proposition that, in conjunction with [Black is present], fully explains [Jones cannot do otherwise], that proposition must not entail [Jones cannot do otherwise]. Plausibly, the proposition that, in conjunction with [Black is present], fully explains [Jones cannot do otherwise], is [Determinism]. But since [Determinism] *does* entail [Jones cannot do otherwise],[16] *EEP* thus renders the verdict that [Black is present] does not partly explain [Jones cannot do otherwise]. So, (7) is true.

---

[13] Even if *EEP* turned out to be false, there would presumably be some *other* principle that renders the correct verdicts in the two aforementioned cases. We would thus need to see if *that* principle similarly establishes the truth of (7). If, however, one is skeptical of the existence of general principles that are supposed to explain a variety of philosophical phenomena, then I may put the point this way: even if *EEP* turned out to be false, the burden is arguably upon the opponent of *PAP* to show us, under the assumption of (^), what the relevant difference is between *The Voting Case* and, e.g., *The Lecture Case*, such that [Black is present] partly explains [Jones cannot do otherwise] in *The Voting Case*, but [Lisa is lecturing] does not partly explain [Jones cannot do otherwise] in *The Lecture Case*.

[14] The fact that Fischer does not initially assume the truth of (^) in *The Voting Case* is irrelevant to my present argument for the truth of (7).

[15] Since *The Voting Case* is supposed to involve an ordinary case of causal determinism (plus Black's lurking nearby), and since Fischer (rightly) grant's Goetz's point that Black (with his device and dispositions) does not *all by himself* render Jones unable to do otherwise, it should not be controversial to explicitly stipulate that there is no overdetermination of explanation of [Jones cannot do otherwise]. However, if one still objects to this stipulation, see my discussion of this issue in the next section.

[16] Recall that we are presently assuming the truth of (^).

One might be left with a suspicion that (7) is false since there appears to be some important difference between [Black is present] in *The Voting Case* and, e.g., [Lisa is lecturing] in *The Lecture Case*. I will now explain the source of this understandable suspicion. I concede that *if* (^) is false in both cases, *then* [Jones cannot do otherwise] is partly explained by [Black is present] in *The Voting Case*, but is not even partly explained by [Lisa is lecturing] in *The Lecture Case*. I, however, have not been concerned here with the implications of the falsity of (^). Rather, my aim has been to defend (7), which is a claim about the implications of the *truth* of (^). So, the above difference between [Black is present] and [Lisa is lecturing] in no way undermines (7).

This completes my argument for (10), which says that the proponent of *Fischer's Improved Argument* is not rationally permitted to hold that [Black is present] even partly explains [Jones cannot do otherwise] in *The Voting Case*. Thus, the proponent of *Fischer's Improved Argument* is not rationally permitted to affirm the truth of (3ii). So, she is likewise not rationally permitted to affirm the truth of either *The Thesis* or (4).

So far, I have attempted to show that in order to arrive at (4), Fischer must establish the truth of *The Thesis*. However, within the context of arguing against the incompatibilist defender of *PAP*, (3i) and (Ci) don't establish *The Thesis*. Moreover, while *The Thesis* may be inferred from (3ii) and (Cii) (given *IRP* and *NMEP*), as I have just shown, Fischer is not rationally permitted to affirm the truth of (3ii). I now turn to why it is important to my critique of *Fischer's Improved Argument* that I grant Fischer the assumption that an overdetermination of explanation for some proposition is possible.

1.5   On the Overdetermination of Explanation

Fischer (2010, 330) emphasizes that "*exactly the same core argument*" can be made against *PAP*

given the truth of (^). In other words, Fischer thinks that a commitment to (2) is not essential to

offering a non-question begging case against the incompatibilist defender of *PAP*. While I agree

with Fischer on this point, it is worth exploring why Fischer thinks this, and how this bears upon

my critique of *Fischer's Improved Argument*. Consider the following remarks:[17]

> [I]t is not at all clear that the mere fact that a prior state of the universe (together with the
>
> laws of nature) explains why Jones cannot at t2 choose otherwise entails that no other
>
> fact can play this sort of explanatory role. Why does the explanation in terms of causal
>
> determination "crowd out" all other explanations, including the explanation in which
>
> Black plays a crucial role (Fischer 2010, 332)?

Notice that I have *not* defended (and need not defend) the position that an overdetermination of

explanation of [Jones cannot do otherwise] is impossible. Indeed, as noted at the outset of this

paper, I grant Fischer the assumption that an overdetermination of explanation for some

proposition is possible. As an incompatibilist defender of *PAP*, I consider the following scenario

to be one in which there is an overdetermination of explanation of [Jones cannot do otherwise]:

> **The Overdetermination Case** Everything that obtains in *The Voting Case* obtains in this
>
> case, except that Black, and his device and dispositions, are deleted. Moreover, Jones'
>
> choice at t2 is causally overdetermined: first, Jones' choice at t2 is deterministically

---

[17] Fischer's employment of the term 'explains' rather than 'entails' in the passage below lends further confirmation
to (3ii) being the correct and charitable interpretation of (3).

caused in exactly the same way it is caused in *The Voting Case*. Let us suppose that both in this case and in *The Voting Case* Jones' choice is deterministically caused by some belief-desire pair (henceforth '*BD*') that belongs to Jones. Second (and in addition), Jones' choice at t2 is deterministically caused in this case by an electromagnetic field, whereby this electromagnetic field and *BD* are mereologically and causally independent of one another.

Here are abbreviations for the following two propositions:

[BD] $=_{df}$ [*BD* deterministically causes Jones' choice at t2]

[Electromagnetic field] $=_{df}$ [The electromagnetic field deterministically causes Jones' choice at t2]

Given the truth of (^), [*BD*] fully explains [Jones cannot do otherwise]. But this does not crowd out the possibility of other such explanations. Indeed, [Jones cannot do otherwise] is *also* fully explained by [Electromagnetic field]. Moreover, notice that this verdict is perfectly consistent with *EEP*, which is only concerned with scenarios in which there is *no* overdetermination of explanation for some proposition.

Can Fischer show that there *is* an overdetermination of explanation of [Jones cannot do otherwise] in *The Voting Case* in order for *EEP* not to apply to *The Voting Case*? Well, Fischer would presumably have to show that [Black is present] fully explains [Jones cannot do otherwise], and that some further proposition also fully explains [Jones cannot do otherwise]. But, as previously noted in support of (6), Fischer (2010, 323–325) concedes Goetz's (2005, 85)

point that Black's presence, device, and dispositions *in themselves and apart from causal determinism* do *not* render Jones unable to avoid ϕ–ing.

Of course, *The Voting Case* could be modified in such a way that Jones' voting decision is *causally* overdetermined, thus guaranteeing that there *is* an overdetermination of explanation for Jones' voting decision. One may simply consider a case just like *The Voting Case,* except that (like *The Overdetermination Case*) Jones' voting decision is deterministically caused both by a belief-desire pair *and* an electromagnetic field. Call this case **The Voting Case***.[18] If Fischer were to replace *The Voting Case* with *The Voting Case**, then *EEP* would no longer apply, since this principle only concerns cases in which there is no overdetermination of explanation of [Jones cannot do otherwise]. However, replacing *The Voting Case* with *The Voting Case** is a false hope. For, if [Black is present] does not even partly explain [Jones cannot do otherwise] in *The Voting Case*, then it is surely safe to assume that the same holds for *The Voting Case**. Surely, adding in an electromagnetic field that deterministically causes Jones' voting decision cannot suddenly make it the case that [Black is present] even partly explains [Jones cannot do otherwise].

We have thus seen that if Fischer were to run exactly the same core argument with (^) rather than (2), I would not object to such an argument on the grounds that an overdetermination of explanation of [Jones cannot do otherwise] is impossible. Rather, I would argue, as I already have, that premise (8) is true: if (^) is true, then [Black is present] does not even partly explain [Jones cannot do otherwise] in *The Voting Case*. So, if the proponent of *Fischer's Improved Argument* replaces (2) with (^), then (3ii) is false. Moreover, these results are perfectly consistent with the assumption that an overdetermination of explanation of [Jones cannot do otherwise] is possible.

---

[18] Unlike *The Overdetermination Case,* in *The Voting Case** it is true that [Black is present].

1.6   Conclusion

Within the context of *Fischer's Improved Argument* being directed at an incompatibilist defender of *PAP*, I have argued for the following: in order to establish (4), one must establish *The Thesis*. But premises (3i) and (Ci) of *The First Avenue* don't establish *The Thesis*. Next, while premises (3ii) and (Cii) of *The Second Avenue* do establish *The Thesis* (under the constraints of *IRP* and *NMEP*), the proponent of *Fischer's Improved Argument* who accepts (2) is not rationally permitted to affirm the truth of (3ii). Moreover, if (2) is replaced with (^), then (3ii) is false. I have argued for these claims while granting Fischer two assumptions that only render a defense of *PAP* more difficult. The first assumption was that the opponent of *PAP* need not establish the truth of *IRR*, and the second one was that an overdetermination of explanation for some proposition is possible.

Now, in fairness to Fischer, I wish to emphasize that if the proponent of *Fischer's Improved Argument* were to replace (2) with the assumption that (^) is *false*, then *PAP* has arguably been refuted. This is due to the fact that, if (^) is false in *The Voting Case*, then [Jones cannot do otherwise] still appears to be true, and [Black is present] appears to partly explain the truth of [Jones cannot do otherwise], despite being irrelevant* to Jones' moral responsibility for his voting decision. So, if Fischer replaces (2) with the *falsity* of (^), *Fischer's Improved Argument* is arguably successful against a *compatibilist* defender of *PAP*. However, I have only been concerned with showing that the *incompatibilist* defender of *PAP* can successfully rebut *Fischer's Improved Argument*.

1.7 Appendix: Palmer's Objection

Palmer (2014) thinks that the inference from (3) to (4) is unwarranted because he rejects (C). Palmer argues against (C) by entertaining and then in turn undermining a variety of principles that, if true, would seem to support (C). I am not convinced that Palmer has refuted all such principles. More specifically, I do not think Palmer has refuted the following principle that, if true, would seem to support (C) (given the truth of (3)):

> **(IP-W)** If a fact is irrelevant to the causal explanation of a person's action, then that fact is irrelevant to *whether or not* that person is morally responsible for her action (Palmer 2014, 3856).

Palmer thinks that (*IP-W*) is false because it is subject to the following alleged counterexample:

> ***Modified-promise-breaker*** Suppose that Jones finds himself with an irresistible desire to decide to break his promise…[and] Jones is a moral person. So, upon finding himself with this desire, Jones tries his very best to resist it. However, despite trying his very best to resist the desire, he eventually succumbs to it and decides to break his promise (Palmer 2014, 3857).

The idea is that Jones' trying his very best to resist his desire to decide to break his promise is irrelevant to the causal explanation of Jones' decision to break his promise. However, Jones' trying his very best to resist his desire to decide to break his promise *is* intuitively relevant to whether or not Jones is morally responsible for his decision to break his promise. More

specifically, Jones is intuitively *not* morally responsible for his decision to break his promise at least partly *because* he tried his very best to resist the desire to decide to break his promise. So, according to Palmer, (*IP-W*) is false.

Before addressing *Modified-promise-breaker*, I will precisify (*IP-W*) in two ways in order to stay clear of potential ambiguities. First, decisions are plausibly basic mental actions; they are things we do not in virtue of doing anything else. Second, the kind of basic desert moral responsibility with which *PAP* is concerned is *non-derivative* (or direct) as opposed to derivative (or indirect). With these two considerations in mind, let's reformulate (*IP-W*) in the following manner:

(**IP-W***) If a fact is irrelevant to the causal explanation of a person's basic mental action, then that fact is irrelevant to *whether or not* that person is non-derivatively morally responsible for her basic mental action.

There is nothing objectionable about precisifying (*IP-W*) in the above manner given that (*IP-W**) likewise supports (C)—a proposition which itself concerns non-derivative moral responsibility for basic mental actions. For, recall that *FSC*s are meant to show that Jones is non-derivatively morally responsible for performing a basic mental action that he could not have avoided.

Now, once we focus on (*IP-W**) which explicitly concerns basic mental actions, it is initially puzzling as to how *Modified-promise-breaker* is supposed to falsify (*IP-W**). After all, *Modified-promise-breaker* involves Jones' *doing* something (viz. trying to resist his desire) before Jones decides to break his promise, which, *ex hypothesi*, is not something Jones does in virtue of doing

anything else. So, how could Jones' trying to resist his desire be relevant to whether or not Jones is non-derivatively morally responsible for his decision to break his promise?

Perhaps Jones' trying to resist his desire is relevant to whether or not Jones is non-derivatively morally responsible for deciding to break his promise *at a particular time*. However, Jones' trying to resist his desire *is* relevant to the causal explanation of the *time* at which Jones decides to break his promise. So, irrespective of whether we focus on Jones' decision to break his promise *simpliciter* or whether we focus on the *time* at which Jones decides to break his promise, *Modified-promise-breaker* does not appear to be a counterexample to (*IP-W**). In order to see this more clearly, I suggest fleshing out *Modified-promise-breaker* in the following manner:

> **Modified-promise-breaker*** Although it is up to Jones whether to refrain from deciding to break his promise from t1–t4, if Jones does not decide to break his promise at t4 or any earlier time, Jones will be causally determined by factors beyond his control to decide to break his promise at t5. Now, it so happens that from t1–t4 Jones tries to resist his desire by refraining from deciding to break his promise from t1–t4. As a result, at t5, Jones decides to break his promise, and Jones' decision is causally determined by factors beyond his control.

Two preliminary questions need to be asked in order to see whether *Modified-promise-breaker** falsifies (*IP-W**).

First question: Is Jones' refraining from deciding to break his promise from t1–t4 irrelevant to the causal explanation of Jones' decision to break his promise at t5? No. This fact *is* relevant.

For, the fact that Jones refrained from deciding to break his promise from t1–t4 explains why, at t5 rather than some earlier time, Jones decided to break his promise.

Second question: Is Jones' refraining from deciding to break his promise from t1–t4 irrelevant to the causal explanation of Jones' decision to break his promise *simpliciter*? Yes. For, Jones' refraining from deciding to break his promise from t1–t4 makes no difference to whether Jones decides to break his promise *simpliciter*.

Now, *if* Jones' refraining from deciding to break his promise from t1–t4 is relevant to whether or not Jones is non-derivatively morally responsible for deciding to break his promise *simpliciter*, then (*IP-W**) is false. However, by the lights of the incompatibilist defender of *PAP*, this is *not* relevant to whether Jones is non-derivatively morally responsible for deciding to break his promise *simpliciter*. After all, it is not up to Jones whether to decide to break his promise *simpliciter*. So, by the lights of the incompatibilist defender of *PAP*—whom Palmer is defending—no matter at which time Jones decides to break his promise, Jones is not non-derivatively morally responsible for deciding to break his promise *simpliciter* in *Modified-promise-breaker**.

To be clear, the incompatibilist defender of *PAP* can still maintain that if Jones were to have instead decided at t4 to break his promise, then Jones would have been non-derivatively morally responsible for deciding *at t4* to break his promise. Indeed, along with Ginet (1996) and Franklin (2011b), Palmer (2011; 2013) himself defends this position.[19] However, this position is consistent with (*IP-W**).[20]

---

[19] For a critique of this approach to moral responsibility that focuses on the time at which an action is performed, see Hunt and Shabo (2013).

[20] Alternatively, the incompatibilist defender of *PAP* may plausibly maintain that if Jones had decided at t4 to break his promise, then Jones would have been *derivatively* morally responsible for deciding to break his promise *simpliciter* in virtue of being non-derivatively morally responsible for deciding to break his promise *simpliciter* on his own (Robinson 2012). This position is also consistent with (*IP-W**).

To conclude, I offered a reason to doubt that *Modified-promise-breaker* undermines (*IP-W**). Moreover, I have tried to flesh out the details of *Modified-promise-breaker* in order to further strengthen such doubt. Hence, unless we are given some new reason to reject (*IP-W**), the incompatibilist defender of *PAP* has good reason to accept (*IP-W**) since it is *prima facie* plausible. So, it appears that *Fischer's Improved Argument* can successfully resist Palmer's objection.

2

Leeway Compatibilism and Frankfurt-Style Cases

According to a traditional Frankfurt-style case (*FSC*) Jones *X*-s on his own, but lacks the ability to avoid *X*-ing. This inability is at least partly due to a preemptive intervener, Black, who is lurking by, ready to directly alter or manipulate Jones' mental processing if, absent Black's intervention, Jones were about to avoid *X*-ing. So, according to the Frankfurtian, although Jones cannot avoid *X*-ing, Jones is nevertheless morally responsible for *X*-ing since Black in no way causally contributes to Jones' *X*-ing. We should thus reject the principle of alternative possibilities (*PAP*), according to which an agent is morally responsible for what she has done only if she could have done otherwise (Frankfurt 1969).

Incompatibilist defenders of *PAP* offer the Dilemma Defense in response (Kane 1985, 51; 1996, 142–145; Widerker 1995; Ginet 1996; Wyma 1997): if causal determinism is true, then although Jones cannot avoid *X*-ing, it is question-begging to assume that Jones is morally responsible for *X*-ing. On the other hand, if causal indeterminism (of the right sort) is true, then there cannot be a prior sign that infallibly indicates to Black as to whether Jones will in fact *X*. So, although Jones may be morally responsible for *X*-ing, Jones *could have* avoided *X*-ing. Hence, according to the Dilemma Defense, *FSC*s fail to undermine *PAP*, at least when *FSC*s are directed at an incompatibilist defender of *PAP*.

Leeway compatibilists—proponents of both *PAP* and the compatibility of moral responsibility and causal determinism—obviously cannot employ the Dilemma Defense in response to *deterministic FSC*s since it is *not* question-begging against *them* to assume that Jones

is morally responsible for *X*-ing, even if Jones' *X*-ing was causally determined by factors beyond his control. As a result, leeway compatibilists must employ a different strategy for defending *PAP* against deterministic *FSC*s.

Kadri Vihvelin (2000; 2004; 2008; 2013), a champion of the so-called new dispositionalism, offers us such a strategy. She argues that under a certain kind of deterministic *FSC* (to be discussed below) Jones *does* have the ability to do otherwise, where that ability is the one at issue in *PAP*.[1] I argue against this position by showing that it is incompatible with the existence of structurally similar cases to *FSC*s in which a preemptive intervener bestows an agent with an ability.[2] Before turning to my argument, let us review Vihvelin's position of *FSC*s.[3]

---

[1] Other new dispositionalists include Smith (2003) and Fara (2008). According to the new dispositionalism, abilities are, or behave in a similar manner to, dispositions, whereby non-manifested dispositions *are* compatible with causal determinism (see Clarke's (2009) thorough discussion of the new dispositionalism). Nelkin (2011) is a close cousin of the new dispositionalists in light of her position that an 'interference-free capacity' sense of an ability to do otherwise is compatible with causal determinism. Like Vihvelin, these three authors claim that an agent in a deterministic *FSC* retains the ability to do otherwise, whereby that ability is the one at issue in *PAP* (Nelkin actually holds that only blameworthiness requires the ability to do otherwise). I will rest content with arguing against one of the more developed leeway compatibilist responses to deterministic *FSC*s. For, it should be fairly straightforward as to how my criticism of Vihvelin's position applies to the other three authors' positions.

[2] Vihvelin's position has also been contested, among others, by Fischer (2008), Whittle (2010), Franklin (2011a), and Kittle (2014). My argument is independent of these criticisms. Moreover, unlike previous criticisms, I will be engaging with Vihvelin's recent (2013) work seems to depart in subtle ways from her (2000; 2004; 2008) work in regards to deterministic *FSC*s (see footnote 4).

[3] After writing this paper, it has come to my attention that, in developing a distinct argument from my own, Levy (2008; 2012; 2014) likewise employs what he calls a 'Frankfurt-enabler case'. So, lest I be perceived as rehashing the same argument, it is crucial that I highlight that we are arguing for strikingly different conclusions.

By employing a Frankfurt-enabler case, Levy cleverly attempts to show that "[t]he mere presence of a counterfactual intervener…can strip an agent of a responsibility underwriting capacity" (2012, 607). So, according to Levy, while an agent in an *FSC* may indeed lack the ability to do otherwise, we should not accept that the agent is responsible for what they have done. Hence, FCSs do not succeed in motivating Frankfurt-style compatibilism, according to which an agent's moral responsibility for her action only depends upon her actual sequence properties (For replies to Levy, see Haji and McKenna (2011) and Cova (2014)).

Notice, then, that Levy implicitly rejects what I will call *The Principle*: if an agent *S* ϕ–s, and something, *Z*, in no way causally contributes to *S*'s ϕ–ing, then *Z* is irrelevant to whether or not *S* is directly/non-derivatively morally responsible for ϕ–ing. My argument, by contrast, is consistent with *The Principle*, as well as with the success of *FSC*s in motivating Frankfurt-style compatibilism. For, I will only be defending the claim that, even if the ability to do otherwise is compatible with causal determinism, an agent in a deterministic *FSC* lacks the ability to do otherwise, where that ability is the one at issue in *PAP*. This is a conclusion that the Frankfurt-style compatibilist can (and is likely to) wholeheartedly accept.

2.1 Vihvelin on Frankfurt-Style Cases

Vihvelin (2013) offers a helpful distinction between two methods that a preemptive intervener *F* can employ in an *FSC* in order to ensure that subject *S* does what *F* wants *S* to do. One is the **Bodyguard** method in which *F*'s intervention is "causally triggered by the *beginnings of any action* (overt or mental) contrary to the intervener's plan" (Vihvelin 2013, 97). Here is an example:

> ***Bodyguard FSC*** Unbeknownst to Jones, Black implants a device in Jones' brain. Black wants Jones to decide to vote republican. Jones decides to vote republican on his own. However, if Black detects that Jones is beginning to decide to vote democrat, Black is committed to intervening by pressing a button that, via the implanted device, causes Jones to decide to vote republican instead.

I grant for the sake of argument that "no story about a Bodyguard [FSC] can, *in principle*, provide a genuine counterexample to PAP" (Vihvelin 2013, 99). What I disagree with Vihvelin about concerns the other method *F* can employ. According to the **Preemptor** method, *F*'s intervention is *not* causally triggered by the beginning of any action. Instead, *F*'s intervention is causally triggered "by some *earlier event* that is a reliable indicator of the fact that the subject will, in the absence of intervention, perform some action (overt or mental) contrary to the Preemptor's wishes" (Vihvelin 2013, 98). Here is an example:

> ***Preemptor FSC*** Unbeknownst to Jones, Black implants a device in Jones' brain. Black wants Jones to decide to vote republican. Jones decides to vote republican

on his own. Moreover, Jones would have (in the absence of Black's intervention) decided to vote democrat on his own if an involuntary prior sign *s* had occurred to Jones. However, if Black detects *s*, Black is committed to intervening by pressing a button that, via the implanted device, causes Jones to decide to vote republican instead.

Vihvelin says that the "presence of a Preemptive intervener of the sort Frankfurt described does not rob a free agent of the ability to choose otherwise" (Vihvelin 2013, 105). Now, the kind of ability to which Vihvelin is referring must be the one at issue in *PAP*. Otherwise, *PAP* will not have been rescued from deterministic *FSC*s. While Vihvelin does distinguish between two kinds of abilities (to be discussed momentarily), she does not explicitly say which ability Jones retains in *Preemptor FSC*. Let us, then, consider both abilities in order to offer the most charitable interpretation of Vihvelin's position.

S has the **narrow ability** to *X iff S* "has what it takes" to *X* (Vihvelin 2013, 11), i.e. *S* has the salient capacities and skills to *X* which are constituted by *S*'s intrinsic properties. For instance, Joan Mitchell has the narrow ability to paint an abstract masterpiece even if she is presently flying on an airplane in which there is no brush, paint, or canvas. Next, *S* has the **wide ability** to *X iff S* has the narrow ability to *X*, and *S* has "the means and the opportunity and nothing external stands in her way" (Vihvelin 2013, 11). So, if Mitchell is presently flying on a plane in which there is no brush, paint, or canvas, she has the narrow ability to paint, but lacks the wide ability to paint. By contrast, if Mitchell is in her studio with a brush, paint, and canvas, then she has the wide ability (and, *a fortiori*, the narrow ability) to paint an abstract masterpiece.

With this basic understanding of these abilities in place, we may thus offer two interpretations of *PAP*:

Narrow-*PAP* An agent is morally responsible for what she has done only if she had the narrow ability to do otherwise.

Wide-*PAP* An agent is morally responsible for what she has done only if she had the wide ability to do otherwise.

Since Wide-*PAP* entails Narrow-*PAP*, the question at issue is not whether Vihvelin must defend Narrow-*PAP*. Rather, the question is whether Vihvelin must *also* defend Wide-*PAP*. I will now argue that Vihvelin must in fact defend Wide-*PAP*.

Consider the following two cases:

Grand Canyon Michael Phelps is on vacation, and is walking around a section of the Grand Canyon in which there is no large reserve of water nearby (and he knows this). An armed and irrational Phelps-fan tells Phelps that unless he presently completes a 200-meter butterfly stroke, the irrational fan will kill an innocent bystander.

Pool Michael Phelps is on vacation, and is swimming in an Olympic-size pool. An armed and slightly less irrational Phelps-fan tells Phelps that unless he

presently completes a 200-meter butterfly stroke, the irrational fan will kill an

innocent bystander.

Whereas in *Grand Canyon* Phelps only has the narrow ability to complete a 200-meter butterfly

stroke, in *Pool* Phelps has the wide ability to complete a 200-meter butterfly stroke. Now, in

contrast to *Pool*, Phelps clearly cannot be morally responsible for not completing a 200-meter

butterfly stroke in *Grand Canyon*. Hence, *if* moral responsibility requires the ability to do

otherwise, the best explanation for why Phelps cannot be morally responsible for not completing

a 200-meter butterfly stroke in *Grand Canyon* is that moral responsibility requires the *wide*

ability (the narrow ability plus the *opportunity*) to do otherwise. For this reason, Vihvelin must

uphold Wide-*PAP*, and thus must maintain that Jones retains the wide ability to avoid deciding to

vote republican in *Preemptor FSC*.[4] However, I will now argue against the position that Jones

retains the wide ability to avoid deciding to vote republican in *Preemptor FSC* by showing that

this position is incompatible with the existence of structurally similar cases to *FSC*s in which a

preemptive intervener bestows an agent with an ability.

2.2 The Argument

Consider the following admittedly bizarre case:

---

[4] In her earlier (2004) work, Vihvelin seemed to only commit to Narrow-*PAP* (see especially Vihvelin 2004, 447). For this reason, Whittle (2010, 12–15) and Franklin (2011a, 97–99) contest that Vihvelin's position on deterministic *FSC*s is unsuccessful since moral responsibility for one's action arguably requires both the (narrow) ability and the *opportunity* to do otherwise (similar remarks apply, *mutatis mutandis*, to Fara's (2008) defense of *PAP*; see Whittle (2010, 8–12) and Franklin (2011a, 99–102)). Hence, by committing to Wide-*PAP*, Vihvelin can accommodate Whittle's and Franklin's point that responsibility requires the opportunity to do otherwise. Of course, whether possessing the opportunity to do otherwise is compatible with causal determinism is a contested issue that I will not take up here (cf. Franklin 2011c).

Bridge Julie is standing on the edge of cliff *A* which is adjacent to cliff *B*. There is

an injured camper at cliff *B* who needs help getting back to camp. Julie cannot

jump from *A* to *B* because the gap is far too wide. However, Julie knows the

following: Black is hundreds of miles away, and thus cannot help the camper.

However, Black has implanted a device in Julie's brain, and thus can detect

whether Julie is about to decide to take the first step off of cliff *A* towards cliff *B*.

More specifically, since causal determinism is true, Black can infallibly detect

this by observing an involuntary prior sign *s*, whereby Julie decides to take the

first step off of cliff *A* towards cliff *B* if *s* had just occurred. Moreover, if Black

detects that Julie is about to decide to take that step, Black plans on pressing a

button that results in a bridge immediately manifesting, and this bridge would in

turn allow Julie to walk with ease from *A* to *B*. Now, despite Julie's knowledge of

all of this (though she doesn't know whether she will in fact decide to help the

camper), Julie will be late for a once-in-a-lifetime concert if she helps the camper.

So Julie decides to walk away from the edge of cliff *A*. As a result, the camper

waits for hours at cliff *B* until someone else helps the camper get back to camp.

Under the assumption that the wide ability to do otherwise is compatible with causal

determinism—an assumption I make throughout the paper for the purpose of the present

dialectic—the following is obviously true:

1. Julie has the wide ability to help the camper in *Bridge*.[5]

---

[5] An anonymous referee for *Thought* has suggested to me that perhaps the plausibility of (1) derives from the
apparent fact that Julie is morally responsible for not helping the camper. Moreover, perhaps one can deny (1) and

What best explains the fact that Julie has wide ability to help the camper? I submit that the best explanation for why Julie has this ability includes, among other things, the truth of the following claim:

> **(*)** In *Bridge*, if Julie's involuntary prior sign *s* were to occur, then Black would (rather than standing idly by) press the button that results in the bridge immediately manifesting.

It is extremely difficult to see how Julie has the wide ability to help the camper if the bridge would *not* manifest if Julie's involuntary prior sign were to occur. After all, helping the camper requires walking from *A* to *B*, and walking from *A* to *B* requires the manifestation of the bridge. So, since (*) is part of the best explanation for why Julie has the wide ability to help the camper, we should conclude that the following is true:

> **2.** If Julie has the wide ability to help the camper in *Bridge*, then (*) is true.

It therefore follows from (1) and (2) that:

> **3.** (*) is true.

---

maintain instead that Julie can do something (viz. decide to take the first step) that would *result* in her having the wide ability to help the camper, because deciding to take the first step would result in the manifestation of the bridge. However, as the referee rightly notes, Vihvelin could only adopt this move at the cost of denying *PAP*, the very principle Vihvelin wishes to defend.

Now, consider the following claim about *Preemptor FSC* that is structurally similar to (*):

> (**) In *Preemptor FSC*, if Jones' involuntary prior sign *s* were to occur, then
> Black would (rather than standing idly by) press the button that results in Jones
> deciding to vote republican (via the implanted device).

Given the structural similarity between (*) and (**), I suggest that we have good reason to accept the following:

> 4. If there is no relevant difference between the relationship between Black and
>   Julie's involuntary prior sign in *Bridge*, and the relationship between Black
>   and Jones' involuntary prior sign in *Preemptor FSC*, then (*) is true *iff* (**) is
>   true.

Now, consider the following claim, which is the antecedent of (4):

> 5. There is no relevant difference between the relationship between Black and Julie's
>   involuntary prior sign in *Bridge*, and the relationship between Black and Jones'
>   involuntary prior sign in *Preemptor FSC*.

Notice that, as is the case in *Preemptor FSC*, in *Bridge* Black is only employing the Preemptor method. In other words, in *Bridge* Black's intervention is only causally triggered by an involuntary prior sign, as opposed to the beginning of an action (overt or mental). I thus

suggest that, in the absence of finding some relevant difference between *Bridge* and *Preemptor FSC*, we should accept (5). I will now consider two objections to (5), each of which attempt to posit a relevant difference.

First objection: Perhaps Vihvelin would want to maintain that the presence of idle interveners can bestow us with new wide abilities, but never remove our wide abilities.[6] But what accounts for this asymmetry? Unless we are provided with some explanation, this move seems entirely *ad hoc*. Moreover, as I will now argue, this position seems false.

Recall that, in accordance with (2), if we wish to maintain that Julie has the wide ability to help the camper, we must *not* hold fixed the fact that Black is standing idly by in the counterfactual scenario in which Julie's involuntary prior sign occurs. Now, consider a case like *Bridge*, except that the bridge between *A* and *B* is already present, but Black plans on *removing* the bridge if Julie's involuntary prior sign occurs (and Julie knows this). Call this case *Bridge 2*. If we do *not* hold fixed the fact that Black is standing idly by in the counterfactual scenario of *Bridge*, then surely the same goes for the counterfactual scenario of *Bridge 2*. But if we do not hold fixed the fact that Black is standing idly by in the counterfactual scenario of *Bridge 2*, then it follows that Black's presence, along with his device and intentions, remove Julie's wide ability to help the camper in *Bridge 2*. Hence, it is untenable to suppose that the presence of idle interveners can bestow us with new wide abilities, but never remove our wide abilities. I suggest, then, that the only potentially relevant difference between *Bridge* and *Preemptor FSC* concerns what each agent knows.[7] Let us now explore whether this difference undermines (5).

Second objection: In *Bridge*, Julie knows that Black is present with certain intentions. By contrast, in *Preemptor FSC* Jones knows neither that Black is present, nor that Black has certain

---

[6] I am grateful to an anonymous referee for *Thought* for bringing this consideration to my attention.

[7] One is at liberty to replace "knows" with "believes", "justifiably believes", "is cognitively sensitive to", etc. This detail makes no difference to my argument.

intentions. So, an objection to (5) might go as follows: Jones' lack of knowledge of Black's presence and intentions implies that when we assess what would happen if Jones' involuntary prior sign were to occur, we hold fixed the fact that Black is standing idly by. By contrast, Julie's knowledge of Black's presence and intentions implies that when we assess what would happen if Julie's involuntary prior sign were to occur, we do *not* hold fixed the fact that Black is standing idly by. Call this the *knowledge-difference* objection to (5).

It is certainly not obvious as to why an agent's knowledge makes a difference to whether we should hold fixed the fact that Black is standing idly by in the counterfactual scenario in which the involuntary prior sign occurs. We may suppose that Black does not know, or care about, what Julie and Jones know. Instead, Black only cares about pressing the button if the involuntary prior sign occurs. We may even suppose that Black does not know, or care about, what happens if he presses the button. How, then, could knowledge of Black's presence and intentions make a difference to what would happen if the involuntary prior sign were to occur? To borrow a phrase from Fischer (1994, 141) on *FSC*s made in a different context, this seems to involve a kind of alchemy.[8]

To be clear, the possession of certain pieces of knowledge can sometimes *add to* one's abilities (Shabo 2014). Indeed, I am willing to concede that if Julie did *not* have knowledge of Black's presence and intentions, Julie would lack the wide ability to help the camper. But notice that we are concerned here with the possession of certain pieces of knowledge that in turn *remove* one's wide abilities.

---

[8] Perhaps *de se* foreknowledge—foreknowledge of one's own actions—results in one losing the wide ability to do otherwise. Such a proposal seems to go against the spirit of leeway compatibilism. But never mind that. The more important point for our purposes here is that, precisely in order to avoid this type of concern, I have stipulated in *Bridge* that Julie does *not* possess *de se* foreknowledge.

Perhaps losing one's wide abilities is possible when one attains knowledge about the *consequences* of certain actions. For example, if I learn that pressing a certain elevator button will result in an explosion that will kill me, it is not implausible to suppose that I had, but now lack, the wide ability to press the elevator button given my subjective motivational set. However, Julie's knowledge in *Bridge* does not fit this description. Her knowledge of Black's presence and intentions has nothing to do with the *consequences* of helping the camper. It appears, then, that the *knowledge-difference* objection to (5) fails. Hence, since there does not appear to be any further relevant difference between the relationship between Black and Julie's involuntary prior sign in *Bridge*, and the relationship between Black and Jones' involuntary prior sign in *Preemptor FSC*, I conclude that (5) is true. Now, it follows from (4) and (5) that:

6. (*) is true *iff* (**) is true.

Moreover, it follows from (3) and (6) that:

7. (**) is true.

Lastly, I contend that we have excellent reason to accept the following claim:

8. If Jones has the wide ability to avoid deciding to vote republican in *Preemptor FSC*, then (**) is false.

What (8) says is that in order for Jones to have the wide ability to avoid deciding to vote

republican, it must be the case that if Jones' involuntary prior sign were to occur, Black would

stand idly by, rather than press the button. Besides (8)'s *prima facie* plausibility, Vihvelin seems

to (rightly, in my opinion) endorse this claim. For instance, consider these remarks:

> Insofar as Black is a Preemptor he has the *power* to make a difference to the facts
>
> about what Jones is able to choose. But since Black does not exercise his power,
>
> these facts remain unchanged and Jones remains as free to choose as he was
>
> before Black came on the scene (Vihvelin 2013, 106).

In a case in which Black is absent, and Jones decides to vote republican on his own, if Jones

retains the wide ability to do otherwise, then it must be the case that if Jones' involuntary prior

sign were to occur, then Jones would avoid deciding to vote republican.[9] So, according to

Vihvelin, since these facts remain unchanged once we add in Black's presence (and his device

and intentions), then it must be true in *Preemptor FSC* that if Jones' involuntary prior sign were

to occur, then Jones would avoid deciding to vote republican, and thus (**) is false. So, (8) is

true.

To be clear, I do not think that we should hold fixed the fact that Black is standing idly by if

Jones' involuntary prior sign were to occur (if we are assessing what occurs in the counterfactual

scenario). But this is because I do not think that Jones has the wide ability to do otherwise. Thus,

in accordance with (8), I *agree* with Vihvelin that *if* Jones has the wide ability to do otherwise in

Preemptor FSC, then we must hold fixed the fact that Black is standing idly in the counterfactual

---

[9] Recall that, within the present dialectic, we are assuming the truth of causal determinism, as well as the
compatibility of causal determinism and the wide ability to do otherwise.

scenario in which Jones' involuntary prior sign occurs. In other words, I agree that *if* Jones has

the wide ability to do otherwise, then (**) is false. So, everyone in the debate should accept (8)

without hesitation.

However, as I have shown, (*) is true (i.e. premise (7) is true). Thus, it follows from (7) and

(8) that:

9. Jones lacks the wide ability to avoid deciding to vote republican in *Preemptor*

   FSC.

We have thus seen that Vihvelin's position on *FSC*s (one of the more developed positions among

the new dispositionalists) is untenable given its incompatibility with the existence of structurally

similar cases to *FSC*s in which a preemptive intervener bestows an agent with an ability (such as

Bridge). Moreover, since leeway compatibilists cannot employ the Dilemma Defense in response

to deterministic *FSC*s, leeway compatibilism's prospects are uninviting.

3

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

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 determinism is incompatible with the ability to do otherwise (Ginet 1966; van Inwagen 1983). The consequence

---

[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.

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 3.1, I discuss Derk Pereboom's (2014) four-case manipulation argument (FCMA) which will take the center stage of this paper. In section 3.2, 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 3.3 I consider the counterintuitive implications of the conjunction of classical compatibilism and a so-called soft-line reply to the FCMA. In section 3.4 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]

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 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: $Plum_1$ (of *Case 1*) is not blameworthy for deciding to kill White. So satisfying the *common compatibilist conditions* for responsibility is not sufficient for

---

[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.

[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).

being responsible for one's action.[5] Moreover, there is no relevant difference among *Cases 1–4*

with respect to Plum's responsibility. So Plum$_4$ (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. Plum$_1$ 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 Plum$_1$'s neural states being induced from

moment to moment by neuroscientists. This gave rise to the concern that Plum$_1$ 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

Plum$_1$. Surely agency can be preserved in the face of momentary external influences (Shabo

2010, 75–77). So the compatibilist must, it seems, concede that Plum$_1$ is an agent in the new

Case 1.[6] If, however, Plum$_1$ satisfies this pre-condition for satisfying the *common compatibilist*

---

[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.

conditions for responsibility, then it should be safe to stipulate that $Plum_1$ *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 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 $Plum_1$'s egoistic tendencies just prior to $Plum_1$'s decision is part of $Plum_1$'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 $Plum_1$ is an agent, Plum's reasoning process and his decision are intrinsically indistinguishable in *Cases 1–4*, and given that is uncontroversial to stipulate that $Plum_4$ satisfies the *common compatibilist conditions* for responsibility, it should be uncontroversial to stipulate that $Plum_1$ 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 $Plum_4$, $Plum_1$ is not blameworthy for deciding to kill White because $Plum_1$'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 $Plum_1$ and $Plum_4$ 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 $Plum_1$'s decision to kill White, whereby $Plum_1$'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 $Plum_1$ lacks, i.e. a condition that involves factors that are external to the $Plum_1$'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, $Plum_1$ *is* blameworthy for his decision to kill White. So, a hard-liner does *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*.

---

[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 $Plum_1$ and $Plum_4$ 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.

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 $Plum_1$ 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.

3.2 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].

---

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

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:

> **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,

---

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

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 2011a). 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.

3.3 A Dilemma for Soft-line Classical Compatibilism

A classical compatibilist who adopts a soft-line reply maintains that, unlike Plum$_4$, Plum$_1$ 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 Plum$_1$ fails to satisfy. So, before presenting the dilemma, I want to defend this assumption against an important objection.

---

[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.

The SLCC theorist who accepts Vihvelin's framework might mistakenly object to the stipulation that there is no relevant intrinsic difference between $Plum_1$ and $Plum_4$ as follows. According to LCA-PROP-ABILITY, in order to determine $Plum_1$'s intrinsic properties, we need to hold fixed the fact that $Plum_1$ is being manipulated when considering a suitable proportion of hypothetical cases. For, being manipulated is an "actual sequence property" of $Plum_1$. However, holding fixed the fact that $Plum_1$ is being manipulated, $Plum_1$ does *not* refrain from deciding to kill White in a suitable proportion of cases. So, LCA-PROP-ABILITY renders the verdict that $Plum_1$ in fact *lacks* the narrow ability to refrain from deciding to kill White. In other words, it's not the case that $Plum_1$'s narrow ability to refrain from deciding to kill White is simply finkish or masked.[12] By contrast, when we hold fixed $Plum_4$'s intrinsic properties, in a suitable proportion of cases $Plum_4$ does, we may suppose, refrain from deciding to kill White. So $Plum_1$ and $Plum_4$ *cannot* be intrinsic duplicates of one another in all relevant respects after all. More specifically, *unlike* $Plum_1$, $Plum_4$ has the narrow ability to refrain from deciding to kill White.[13] For this reason, one cannot consistently maintain that $Plum_1$ and $Plum_4$ are intrinsic duplicates of one another in all relevant respects, and also that only $Plum_1$ is being manipulated by neuroscientists.

This objection has one critical flaw. There is no inconsistency in supposing that $Plum_1$ and $Plum_4$ are intrinsic duplicates of one another in all relevant respects, and that only $Plum_1$ is being manipulated by neuroscientists. For, in order to determine $Plum_1$'s narrow abilities according to LCA-PROP-ABILITY, we do *not* hold fixed the fact that $Plum_1$ is being manipulated when considering a suitable proportion of hypothetical cases. Let me explain.

---

[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).

The manipulation of $Plum_1$ is nothing over and above the neuroscientists exerting a momentary egoism enhancing tendency in $Plum_1$. So, the crucial difference between $Plum_1$ and $Plum_4$ is, as previously stressed, the causal *origin* of the exact strength of the egoistic tendencies in $Plum_1$ and $Plum_4$, respectively. In other words, the difference consists in factors that are *external* to $Plum_1$ and $Plum_4$, respectively. So, being manipulated is no doubt an "actual sequence property" of $Plum_1$. Nevertheless, the neuroscientists' intentional action is not intrinsic to $Plum_1$'s reasoning process or $Plum_1$'s decision. In that case, the neuroscientists' intentional action *at best* robs $Plum_1$ of the *opportunity* to decide to refrain from killing White. However, according to LCA-PROP-ABILITY, in order to determine $Plum_1$'s narrow abilities we need to consider hypothetical cases in which $Plum_1$ *has* the opportunity to decide to kill White. For this reason, in order to determine $Plum_1$'s narrow abilities, we do *not* hold fixed the fact that $Plum_1$ is being manipulated when considering a suitable proportion of hypothetical cases. So, no inconsistency results from stipulating that $Plum_1$ and $Plum_4$ are intrinsic duplicates of one another in all relevant respects, and that only $Plum_1$ is being manipulated by neuroscientists.

So, *if* there is a morally relevant difference between the actions of $Plum_1$ and $Plum_4$ it may due to the fact only $Plum_4$ 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 $Plum_1$ fails to satisfy.[14] Moreover, this position does not escape my dilemma which I will now present.

---

[14] Perhaps a compatibilist can construct an alternative account of abilities whereby it *is* objectionable to stipulate that $Plum_1$ and $Plum_4$ are intrinsic duplicates of one another in all relevant respects and yet only $Plum_1$ 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.

The SLCC theorist must hold that there is an external condition for responsibility that $Plum_1$ 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 $Plum_1$ 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 $Plum_1$ fails to satisfy) is *not* necessary for S to be able to do other than φ.

The conjunction of SLCC and (i) entails that $Plum_1$ is not blameworthy for his decision to kill White, and $Plum_1$ was not able to refrain from deciding to kill White.[15] Moreover, $Plum_1$'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 $Plum_1$'s control that causally determined $Plum_1$'s decision to kill White.

To be clear, there may be a *deeper* explanation as to why $Plum_1$ 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 $Plum_1$'s inability to do otherwise, irrespective of whether $Plum_1$ lacks the ability to do otherwise *because* of the activities of the neuroscientists. So, the conjunction of SLCC and (i) entails the following:

---

[15] The SLCC theorist who accepts (i) and *PAP* would presumably say that $Plum_1$ is not blameworthy for deciding to kill White at least partly *because* $Plum_1$ was not able to do other than decide to kill White.

(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 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 Plum_1's control are relevant to Plum_1'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 Plum_1 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 $Plum_1$'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 $Plum_1$'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 $Plum_1$'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 $Plum_1$'s decision to kill White was causally determined by factors beyond $Plum_1$'s control, and, crucially, such factors include the intentional actions of other agents, $Plum_1$ fails to satisfy some external condition for responsibility. So $Plum_1$ 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 $Plum_1$ retains the ability to refrain from deciding to kill White. Additionally, as

indicated in section 3.1, 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 $Plum_1$ is blameworthy for his decision to kill White *if* we assume that $Plum_1$ retains the ability to refrain from deciding to kill White and $Plum_1$ satisfies the *common compatibilist conditions* for responsibility. For, under these assumptions, $Plum_1$ 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 $Plum_1$ retains the wide ability to refrain from deciding to kill White, and $Plum_1$ satisfies the *common compatibilist conditions* for responsibility, $Plum_1$ is nevertheless blameless for deciding to kill White. But what could possibly be missing here that renders $Plum_1$ blameless? Barring the impossible demand for being *causa sui,* it seems that there is nothing further that *could* be demanded in order to render $Plum_1$ blameworthy.

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

---

[16] Perhaps we lack ultimate 'heaven and hell' responsibility for our actions given that we are not *causa sui*. As Clarke (2005, 20–23) notes, however, it does not follow that we lack a weaker form of basic desert responsibility.

manager's insult was influenced in the weak form whereas $Plum_1$'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 $Plum_1$ has not been given an irresistible urge to decide to kill White. In that case, it is dificult 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 $Plum_1$ 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.

3.4 A Dilemma for Hard-line Classical Compatibilism

The classical compatibilist who adopts a hard-line reply maintains that $Plum_1$ 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:

---

[17] One difference between $Plum_1$ and the manager is that $Plum_1$ 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.

(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 Plum$_1$ is blameworthy for deciding to kill White, Plum$_1$ must have the ability to refrain 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 direct[18] 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

---

[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.

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 $Plum_1$ 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 $Plum_1$ will do precisely what the neuroscientists want $Plum_1$ to do. So, if you agree that $Plum_1$ cannot do otherwise in the old *Case 1* just because the neuroscientists act on what they know will nomologically necessitate $Plum_1$'s decision to kill White, then you should agree that $Plum_1$ 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 *FSC*s 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 ~φ (in the absence of Black's intervention). So, the fact that Black would

---

[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 *FSC*s.

have intervened if Jones were about to ~φ (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) ~φ 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 ~φ (barring an intervention by Black) (Vihvelin 2013, 97–99). Vihvelin's position that Jones can do otherwise in certain deterministic *FSC*s 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; Whittle 2010; 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

---

[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.

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 ~φ. Now, according to Vihvelin, if Black is only employing the *Bodyguard* method in a deterministic *FSC* (a case in which Black does not actually intervene), then Jones lacks the wide ability to ~φ, although Jones retains the wide ability to *choose* to ~φ (or at least try or begin to choose to ~φ) (Vihvelin 2013, 100).[21] In that case, Vihvelin would surely agree

---

[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 ~φ but only retains the wide ability to, say, try to ~φ. 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

that in *Actual Intervention* according to which Black *actually* intervenes by employing the *Bodyguard* method, Jones *still* lacks the wide ability to $\sim\varphi$ (even if Jones retains the wide ability to, e.g., try to $\sim\varphi$). 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 *FSC*s 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 1*. More specifically, both *Case Actual Intervention* and *Case 1* involve an agent who is manipulated to perform some action while simultaneously retaining the *common compatibilist*

---

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 Plum$_1$'s decision, we can simply shift our attention to Plum$_1$'s *attempt* to make his decision. If Plum$_1$ is responsible for his attempt to make his decision, and *PAP* is true, then Plum$_1$ could have refrained from trying to make his decision, *despite* the fact that Plum$_1$'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 3.2, Jacqueline du Pré *can* be (non-derivatively) morally responsible for not playing the cello.

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 Plum$_1$ 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

Plum$_1$ 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

---

[22] Among the very few philosophers that have explicitly considered the relationship between classical compatibilism and the manipulation argument, Berofsky (2006, 435–439) 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).

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).

3.5 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:

(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.

4

Manipulating Deliberators

Derk Pereboom (2014, ch. 5) offers an account of rational deliberation in defense of the following thesis:

> **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).

I offer a counterexample to Pereboom's account of rational deliberation. I then employ this counterexample to develop a four-case manipulated deliberation argument for the conclusion that (DC) is false and that the following thesis is true:

> **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).

My argument also shows that rational deliberation is incompatible with believing that one's actions are not up to oneself. I concede that there are viable replies to my argument against (DC).

However, given the structural similarity between my argument and Pereboom's (2014, ch. 4) four-case manipulation argument (henceforth 'FCMA') against compatibilism, both arguments stand or fall together. As a result, a hard incompatibilist[1] such as Pereboom who defends the FCMA faces a dilemma: either give up (DC) as well as the view that one can (as a hard incompatibilist) rationally deliberate, or give up the FCMA. Before turning to Pereboom's defense of (DC), here are four clarificatory remarks on the notion of rational deliberation that bear upon the discussion at hand.

First, when I talk throughout this paper about a certain belief an agent must have or lack in order to rationally deliberate, I will not further clarify as to whether what is at issue is an occurent belief, a dispositional belief, or a belief to which only a reflective agent is rationally committed (Pereboom 2014, 119). These details do not affect the plausibility of my argument against (DC).

Second, deliberation partly consists in considering two or more incompatible courses of action. Now, suppose that the two incompatible courses of action under an agent S's consideration are φ-ing and ψ-ing. One might think that in order for S to rationally deliberate about whether to φ or to ψ, S must believe both that S can φ, and that S can ψ. For the purposes of this paper, however, I will only make the weaker assumption that in order for S to rationally deliberate about whether to φ or to ψ, S must only believe that S *might* be able to φ, and that S *might* be able to ψ.[2] This weaker assumption is consistent with (DC).

---

[1] A hard incompatibilist rejects both compatibilism and libertarianism. Pereboom (2001; 2014) does not claim, however, that free will is impossible. Rather, Pereboom only claims that we do not *in fact* have free will, contra the libertarian's position.

[2] This position is referred to as the No Belief in Ability Thesis by Coffman and Warfield (2005, 37). For a defense of this thesis, see Taylor (1966, ch. 12); Waller (1985, 49); Kapitan (1986, 235–241); Pettit (1989, 43); Clarke (1992, 110); Bok (1998, 110); and Searle (2001). This thesis is, however, disputed by Coffman and Warfield (2005, 37–38).

Third, an agent *S rationally* deliberates about some matter only if *S* deliberates and *S* does not have any inconsistent beliefs that are salient to *S*'s deliberation. I will refer to this necessary condition for rational deliberation as the 'no inconsistent beliefs' (NIB) condition:

> **(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 I rightly believed that I cannot fly to Mars. Now, if I were to deliberate about whether to fly to Mars, then my deliberation would not be an instance of *rational* deliberation. Why? The reason is that deliberating about whether to fly to Mars presumably commits me to the belief that I at least might be able to fly to Mars. But this belief is inconsistent with my initial belief that I cannot fly to Mars. Moreover, these inconsistent beliefs are salient to my deliberation about whether to fly to Mars. So, I fail to satisfy (NIB) in this instance of deliberation.[3]

Suppose by contrast that I deliberate about whether to purchase vanilla or chocolate ice cream. In this instance of deliberation, I can satisfy (NIB) *even if* I believe both that I cannot fly to Mars and that I might be able to fly to Mars. This is because these inconsistent beliefs are not *salient* to my deliberation about whether to purchase vanilla or chocolate ice cream. Obviously, it may be difficult to give a precise account of which beliefs count as salient. However, the details of such an account are peripheral to the discussion at hand.

---

[3] Throughout this paper, I will continue to speak loosely by saying, e.g., that I satisfy (NIB), rather than making the more precise claim that I satisfy the condition for rational deliberation according to (NIB).

Fourth, Coffman and Warfield (2005) argue in favor of the possibility of "double-minded deliberation" whereby *S* deliberates about whether to *φ*, and *S* believes both that *S* can *φ* and that *S* cannot *φ*. Moreover, Coffman and Warfield also seem to think that double-minded *rational* deliberation is possible, although they do not say so explicitly; Pereboom (2014, 125) interprets Coffman and Warfield in this manner. Contra what Coffman and Warfield at any rate seem to suggest, double-minded deliberation appears to be a paradigmatic case of irrational deliberation. Fortunately, Pereboom (2014, 125) himself agrees that satisfying (NIB) is a necessary condition for rational deliberation. Moreover, this is how things should be; it would, after all, be a hollow victory for the defender of (DC) if a non-libertarian could "rationally deliberate" only if she had inconsistent beliefs. Let us now turn to Pereboom's defense of (DC).

4.1 Pereboom on Rational Deliberation

Oversimplifying matters for a moment, Pereboom maintains that there are two conditions that are each necessary and jointly sufficient for rational deliberation. What Pereboom (2014, 108) really means, however, is that these two conditions, *together with other necessary conditions that cut across the debate over* (DC), are jointly sufficient for rational deliberation. It is thus crucial to my argument that I construe Pereboom's view as these two conditions, *together with* (NIB), being each necessary and jointly sufficient for rational deliberation. I will now introduce these two conditions. First, Pereboom (2014, 113) affirms the following epistemic-openness condition:

> **(S)** In order to deliberate rationally among distinct actions A1… An, for each Ai, S cannot be certain of the proposition that she will do Ai nor of the proposition that she will not do Ai; and either (a) the proposition that she will do Ai is consistent with every

proposition that, in the present context, is settled for her, or (b) if it is inconsistent with some such proposition, she cannot believe that it is.

Pereboom (2014, 113) is employing 'settled' in (S) as a technical term to be understood in the following manner:

> **(Settled)** A proposition is settled for an agent just in case she believes it and disregards any doubt she has that it is true, e.g., for the purpose of deliberation.[4]

Second, Pereboom (2014, 108) affirms the following deliberative efficacy condition:

> **(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.

To appreciate the ingenuity of (DE), notice that it is not only consistent with (DC). Rather, it also renders the intuitively correct verdict in a case first proposed by Peter van Inwagen (1983, 154) in an attempt to undermine (DC), which I summarize as follows:

---

[4] Some proponents of (DC) such as Nelkin (2011, 137) reject an epistemic openness requirement for rational deliberation. Nothing I wish to argue for below hinges on whether Nelkin is correct. Rather, my primary focus is on the inadequacy of (DE) for a full-fledged account of rational deliberation.

Two Doors *S* is in a room with two doors. *S* believes that one of the doors is unlocked,

and the other is locked and impassable, though *S* lacks a belief about which door is

locked and which door is unlocked.

Intuitively, *S* cannot rationally deliberate about which door to open in *Two Doors*. According to

van Inwagen (1983, 143–161), the reason for this is that *S* must believe that *S* can perform each

action (i.e. open each door), and thus *S* must believe that *S* can do otherwise. So, under the

assumption that causal determinism is incompatible with the ability to do otherwise (Ginet 1966;

van Inwagen 1983), it follows that (DC) is false.

   Notice, however, that since *S* believes that one of the doors is locked and impassable, *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, *S* would in fact open that door. So, *S* does not satisfy (DE) in *Two Doors*. As

a result, *Two Doors* does not refute (DC).

4.2 The Counterexample

I now proceed to offer my counterexample to Pereboom's account of rational deliberation

(henceforth '*Pereboom's Account*'), according to which (NIB), (S), and (DE) are each necessary

and jointly sufficient for rational deliberation:

   Case 1 (i)–(iv) 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 button on the left, Betty will receive one thousand

   dollars. Call this button 'LEFT'. If she presses the button on the right, then Oxfam will

instead receive one thousand dollars. Call this button 'RIGHT'. Whichever button Betty

presses first will dictate where the money goes. So Betty cannot alter who will receive 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. So Betty

cannot distribute one thousand dollars to both parties. Betty believes that giving the

money to Oxfam will in turn contribute to benefiting people who are far worse off than

Betty, a middle-class U.S. citizen. However, Betty also has a strong desire to take the

money in order to purchase a professional espresso machine that she has desired for some

time (Betty frequently but not exclusively undergoes rationally egoistic tendencies).

(ii) Betty is not certain of the proposition that she will press LEFT nor of the proposition

that she will press RIGHT; and both the proposition that she will press LEFT and the

proposition that she will press RIGHT are each consistent with every proposition that, in

the present context, is settled for Betty. **So Betty satisfies (S).**

(iii) 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 (DE).**

(iv) 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 (say, on the basis of a coin toss) which button

Betty is to press (and decide to press). As a result, the neuroscientists will manipulate

Betty to decide to press one of the buttons by exerting either an egoism-enhancing or

egoism-diminishing momentary influence upon Betty. Though, Betty has no belief about

which button the neuroscientists in fact want her to decide to press (cf. Pereboom 2014: 76–77).[5]

If Betty cannot rationally deliberate about which button to press, despite the fact that Betty satisfies (NIB), (S), and (DE), then *Case 1* is a counterexample to *Pereboom's Account*. Now, I contend that, given the truth of (iv), Betty cannot rationally deliberate about which button to press. To be clear, perhaps Betty cannot help but act *as if* she can rationally deliberate in *Case 1*, just as perhaps one cannot help but act *as if* one is free in a case in which one believes that one is not free (Korsgaard 1996, 162–163). Even so, this in no way undermines my claim.

Now, I lack a positive argument for the claim that Betty cannot rationally deliberate about which button to press given (iv), beyond simply highlighting my intuition about the matter. Notice, however, that Pereboom's own *Case 1* of the FCMA is likewise supposed to simply garner the intuition that Plum is not responsible for his decision to kill White. Additionally, it is worth pointing out that even Nelkin (2004, 223; 2011, 129)—a proponent of (DC)—considers a similar case in which one learns that one will be manipulated by a neuroscientist to vote for a certain candidate, and likewise concludes that in such a case one presumably cannot rationally deliberate about whom to vote for. So, my intuitions are arguably not based solely upon a prior, dyed-in-the-wool commitment to (DI).[6]

---

[5] I stipulate that the neuroscientists will in fact intervene, regardless of whether Betty would decide to press the neuroscientists' pre-selected button in the absence of an intervention? If I stipulate instead that the neuroscientists intervene only if, in the absence of an intervention, Betty would not have pressed the neuroscientists' pre-selected button, then *Case 1* would more closely resemble a Frankfurt-style case which would in turn raise many vexing issues that are simply beyond the scope of this paper.

[6] As I argue in the appendix, in order for Nelkin's account of rational deliberation to render the intuitively correct verdict in *Case 1*, a historical condition must be explicitly built into Nelkin's account of rational deliberation. Moreover, as I will argue in section 4.4, Pereboom cannot consistently defend the FCMA and accept a historical condition for rational deliberation.

Before proceeding, I wish to preempt a potential objection. Levy (2006, 456–459) offers an interesting case in support of (DC) that is like van Inwagen's *Two Doors* case, except that which door is unlocked is based upon a predictor's ability to predict that Sally, the deliberator, will choose to open that door (and Sally believes this). Let us call Levy's case *Newcomb's Doors*. One might infer from certain similarities between *Newcomb's Doors* and *Case 1* that if Sally can rationally deliberate about which door to open, then Betty can also rationally deliberate about which button to press (cf. Henden 2010). While I cannot adequately address whether Levy's case supports (DC) here, there is, I think, a crucial difference between *Newcomb's Door* and *Case 1*.

Betty's decision to press one of the buttons is dependent (in some sense) upon the neuroscientists' decision concerning which button Betty will press, and not vice versa. By contrast, Sally's decision to open one of the doors is *not* dependent upon the mental states of the predictor. In fact, in *Newcomb's Doors* the dependence goes the other way around; which door the predictor unlocks is dependent upon *Sally's* mental states. This critical difference between *Newcomb's Doors* and *Case 1* undermines the claim that if Sally can rationally deliberate about which door to open, then Betty can rationally deliberate about which button to press. So, we should in fact maintain that, given Betty's beliefs according to (iv), Betty cannot rationally deliberate about which button to press. I now turn to the more difficult task of showing that Betty satisfies (S), (DE), *and* (NIB).

Betty satisfies (S) given (ii). Betty also clearly satisfies (DE) since (iii) is a word-for-word reproduction of the (DE) condition. The *real* question is whether Betty has any inconsistent beliefs (that are salient to Betty's deliberation) given (i)–(iv). If Betty *does* have any inconsistent beliefs given (i)–(iv), then Betty fails to satisfy (NIB), and thus *Case 1* is not a counterexample to *Pereboom's Account* after all. Notice that Betty has no inconsistent beliefs given (i)–(iii). So,

in order to see whether Betty satisfies (NIB) in *Case 1*, we need to look closely at (iv), and its relation to (i)–(iii).

For clarificatory purposes, I wish to emphasize that according to (iv) *what* the neuroscientists are manipulating is Betty's *decision* to press one of the buttons (which results in her pressing the relevant button), rather than Betty's *judgment* of what would be best to do. Formulating (iv) in this manner results in my *Case 1* more closely resembling *Case 1* of Pereboom's FCMA, which will be important for my purposes later on. Moreover, if I were to *instead* stipulate that (what Betty believes is that) the neuroscientists manipulate Betty's *judgement* about what would be best to do, it would be less evident that *Case 1* is a counterexample to *Pereboom's Account*, or that *Case 1* poses a threat to (DC). After all, (DC) concerns a belief about one's *action* being causally determined by factors beyond one's control, *not* one's *judgment* about what would be best to do. I now proceed to show that Betty's beliefs according to (i)–(iii) are consistent with Betty's beliefs according to (iv).

The beliefs that Betty has according to (i) and (iv) are clearly consistent with one another. Moreover, the beliefs that Betty has according to (ii) and (iv) are consistent with one another; even though Betty believes that the neuroscientists have already decided which button she is to press, Betty has no clue as to which button the neuroscientists in fact want her to press. So, both the proposition that she will press LEFT and the proposition that she will press RIGHT are consistent with every proposition that, in the present context, is settled for Betty according to (iv). I now attempt to defend the less evident claim that the beliefs that Betty has according to (iii) and (iv) are consistent with one another.

According to (iii), 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. Why does Pereboom employ the phrase "under normal conditions" rather than "under the actual conditions" in (DE)? Because, as Pereboom (2014, 119–120) notes, *Pereboom's Account* must be consistent with the possibility of *S* rationally deliberating in the face of the belief that *S* might act akratically, i.e. act contrary to what *S* believes it would be best to do. 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.

Notice, then, that since Betty *does* believe that she might act akratically since she doesn't know which button the neuroscientists in fact want her to press, it follows that Betty believes that she is *not* under normal conditions. In that case, Betty's beliefs according to (iii) are consistent with Betty *lacking* the further belief 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 the actual conditions*, she would also, on the basis of this deliberation, press LEFT; and similarly for pressing RIGHT. So Betty has no inconsistent beliefs given (iii) and (iv).

As Pereboom (2014, 120) himself notes, (DE) could be modified in a different manner in order to account for the compossibility of rational deliberation and the belief that one might act akratically. We can present Pereboom's suggestion in the following manner:

> **(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 questions about what normal conditions amount to, while at the same time being consistent with the compossibility of rational deliberation and the belief that one might act akratically. Notice, however, that since Betty can satisfy (DE*) and consistently believe that she might act akratically, it seems that Betty can also satisfy (DE*) and consistently believe what is stipulated in (iv).

Perhaps I have failed, however, to pay sufficiently close attention to the phrase "on the basis of this deliberation". More precisely, if Betty satisfies (DE*), then perhaps Betty's belief about the neuroscientists manipulating her decision is inconsistent with Betty's believing that whichever decision she makes under the actual conditions might be *based on her deliberation*. While this is the most promising route for blocking *Case 1* from being a counterexample to *Pereboom's Account*, I do not think that it ultimately succeeds.

Recall that the deliberative efficacy condition should be consistent with the possibility of *S* rationally deliberating in the face of the belief that *S* might act akratically. Hence, any plausible account of rational deliberation that is consistent with (DC) must surely grant the possibility of *S* rationally deliberating in the face of *S*'s belief that antecedent factors beyond *S*'s control momentarily enhance or diminish, e.g., *S*'s egoistic tendencies, whereby this momentary influence can be a difference-maker with respect to which action *S* performs. But notice that (according to what Betty believes) this is precisely how the neuroscientists are manipulating Betty's decision, viz. by exerting an egoism-enhancing or egoism-diminishing momentary influence upon Betty. So even if Pereboom were to replace (DE) with (DE*), there is no good

reason to think that Betty cannot satisfy both (NIB) and (DE*), even given Betty's beliefs according to (iv).[7] Let us now take stock.

I at any rate suggested that Betty cannot rationally deliberate about which button to press given (iv). I then showed that Betty satisfies (S) and (DE). And, I have now just argued that Betty satisfies (NIB), *and* that Betty can *still* satisfy (NIB) if she were to satisfy (DE*) instead of (DE). So (NIB), (S), and (DE) (or instead (DE*)) are not jointly sufficient for rational deliberation. So *Case 1* is a counterexample to *Pereboom's Account*.

One final move I have not yet considered is incorporating a *historical condition* into (DE), whereby, e.g., *S* fails to satisfy (DE) if *S* believes that *S*'s action will be causally determined by factors beyond *S*'s control, and those factors include in some manner the intentional actions of other agents (cf. Lycan 1997, 115–119). I find such a position to be implausible (cf. Pereboom 2001, 125; Mele 2006, 141–142). But more importantly, I will argue in section 4.4 that Pereboom cannot build a historical condition into (DE) and consistently uphold the FCMA. Right now, however, I wish to employ *Case 1* to develop my four-case manipulated deliberation argument against (DC).

4.3 Completing the Four-Case Manipulated Deliberation Argument

According to *Case 2*, (i)–(iii) are true and (iv) is false. Instead, (v) is true:

> **(v)** Betty believes the following: long ago, a team of neuroscientists have decided
>
> arbitrarily (say, on the basis of a coin toss) which button Betty is to press (and decide to

---

[7] My remarks here are similar to Pereboom's (2014, 76) claim that agency and responsibility may be preserved in the face of egoism-enhancing (and presumably also egoism-diminishing) momentary influences. Cf. Shabo (2010, 75–77).

press). As a result, these neuroscientists have programmed Betty at the beginning of her life in such a manner that she will decide to press one of the buttons, though Betty has no belief about which button the neuroscientists want her to decide to press (cf. Pereboom 2014, 77).

Whether Betty believes that the process of manipulation of her decision begins a few seconds prior to her decision or at the beginning of Betty's life surely cannot make a difference with respect to Betty's ability to rationally deliberate about which button to press. Moreover, if Betty's beliefs according to (i)–(iii) are consistent with Betty's beliefs according to (iv), the same goes for (v). Thus, just as Pereboom's account of rational deliberation yields the intuitively incorrect verdict in *Case 1*, it likewise renders the intuitively incorrect verdict in *Case 2*.

According to *Case 3*, (i)–(iii) are true and (v) is false. Instead, (vi) is true:

> **(vi)** 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 decide to press one of the buttons. Though, Betty has no belief about which button she will in fact decide to press (cf. Pereboom 2014, 78).

Whether Betty believes that Betty is causally determined to press a certain button as a result of the programming of neuroscientists or as a result of the training practices of Betty's community does not appear to make a difference with respect to Betty's ability to rationally deliberate.

Moreover, if Betty's beliefs according to (i)–(iii) are consistent with Betty's beliefs according to (v), the same goes for (vi).

Finally, according to **Case 4**, (i)–(iii) are true and (vi) is false. Instead, (vii) is true:

> **(vii)** 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 circumstance. As a result, Betty's deliberative reasoning processes, in conjunction with certain background conditions, will causally determine Betty to decide to press one of the buttons. Though, Betty has no belief about which button she will in fact decide to press (cf. Pereboom 2014, 79).

Once again, there is no relevant difference between (vi) and (vii). So since Betty cannot rationally deliberate about which button to press in *Case 3*, the same holds for *Case 4*.

What explains the fact that Betty cannot rationally deliberate about which button to press in *Cases 1–4*? The best explanation is that Betty believes that her decision to press one of the buttons is (or will be) causally determined by factors beyond her control. So (DC) is false. However, I suggest that there is yet a more fundamental explanation as to why Betty cannot rationally deliberate.

Consider *Cases 1–4*, except that in each case (Betty believes that) Betty's decision to press one of the buttons is *indeterministically* caused by factors beyond her control, such that it is not up to her whether one decision rather than another is indeterministically caused by factors beyond her control (cf. Pereboom 2001, 41–54; 2014, 83–86). Call these cases *Cases 1*–4**. If you agree that Betty cannot rationally deliberate in *Case 1*, you are likely to (or at any rate

should) agree that the same goes for *Case 1**. Moreover, since there is arguably no relevant difference among *Cases 1*–4**, we need an explanation for why Betty cannot rationally deliberate in *Cases 1*–4** that does not appeal to Betty's belief that her decision is causally determined by factors beyond her control since Betty *lacks* this belief in *Cases 1*–4**. What, then, is the missing ingredient that explains why Betty cannot rationally deliberate either in *Cases 1–4* or in *Cases 1*–4** (assuming there is one single missing ingredient)?

According to Richard Taylor (1964, 76), rational deliberation requires that an agent believe that it is up to oneself whether to do it or not. I think that Taylor is correct, or at least very close to the truth. That is, we can relax Taylor's condition if we like to the claim in order to rationally deliberate about whether to φ, one must at least *lack* the belief that it is *not* up to oneself whether to φ. I thus propose the following requirement for rational deliberation:

> **(UT)** 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 up to S whether to do A1 or A2, whereby S believes that it is not up to S in the aforementioned sense if S believes that causal determinism is true, or if S believes that performing A1 or A2 is (or will be) indetermistically caused by factors beyond S's control, such that it is not up to S whether A1 or A2 is indeterministically caused by factors beyond S's control.

On my view, then, (NIB), (S), (DE), and (UT) are each necessary and jointly sufficient for rational deliberation.[8] So, in order to rationally deliberate about which button to press in *Cases*

---

[8] Notice, then, that on my view *S* cannot rationally deliberate about which door to open in *Two Doors* because *S* fails to satisfy (DE) *and* (UT).

1–4 and *1*–4**, Betty must lack the belief that it is not up to her, in the libertarian sense, to choose which button to press. But lacking this belief is inconsistent with what Betty believes in *1–4* and *1*–4**. So, in *1–4* and *1*–4**, Betty either fails to satisfy (UT) or Betty fails to satisfy (NIB). Betty cannot satisfy both. So my account renders the intuitively correct verdict in all of these cases. Notice that satisfying both (UT) and (NIB) is not only incompatible with believing that causal determinism is true. Rather, satisfying both (UT) and (NIB) is also incompatible with a commitment to hard incompatibilism. So, on my view, irrespective of whether Pereboom believes that determinism is true, as a hard incompatibilist, Pereboom cannot rationally deliberate.

4.4 Objections

There are many ways to respond to my argument against (DC) (and ultimately in favor of (UT)). We can begin by distinguishing between two types of replies—the hard-line reply and the soft-line reply. According to a hard-line reply, we start with *Case 4* in which it seems that Betty can (or might be able to) rationally deliberate about which button to press. And, since there is no relevant difference between *Case 4* and *Case 1*, we can conclude that Betty can in fact (or might be able to) rationally deliberate about which button to press in *Case 1*, despite the truth of (iv). So *Case 1* is not a counterexample to *Pereboom's Account* after all (cf. McKenna 2008; 2014; Haas 2013).

According to a 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 just two potential differences one might note. First, Betty cannot rationally deliberate only 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 thus cannot rationally deliberate.

I grant that at least some the above replies to my argument are viable. That is, I do *not* take

myself to have conclusively shown that (DC) is false (or that (UT) is true). However, the

important thing to note is that these replies—*including any soft-line reply that explicitly*

incorporates a historical condition for rational deliberation—are structurally similar to certain

replies to Pereboom's FCMA. As a result of this isomorphism, any reply (and counter-reply) to

one argument can be appropriately tailored to be directed towards the other argument, as I will

now illustrate.

A hard-line reply to Pereboom's FCMA simply needs to be tailored in such a way that it is

claimed that there is no relevant difference with respect to what Betty *believes* in all four cases.

Similarly, for any soft-line reply to the FCMA that attempts to mark a relevant difference

between any pair of adjacent cases, a similar move can be made in response to my argument by

marking a relevant difference between a pair of adjacent cases with respect to what Betty

believes. Conversely, any hard-line or soft-line reply to my argument can be tailored to apply to

the FCMA by focusing on what obtains in any of the four cases of the FCMA, rather than

focusing on what Betty *believes* in any of the four cases of my argument. So, if Pereboom were

to endorse a soft-line reply to my argument by adopting a historical condition for rational

deliberation, it is difficult to see how Pereboom could consistently object to a soft-line reply to

the FCMA that similarly adopts a historical condition for responsibility. It is for this reason that I

claimed in section 4.2 that if Pereboom were to incorporate a historical condition into (DE) in

order to render the intuitively correct verdict in *Case 1*, Pereboom could not consistently uphold the FCMA.

Why do the replies to each argument share a structural similarity? Because the arguments themselves share a structural similarity. In each case of my argument, part of what Betty *believes* to obtain mirrors what in fact obtains in Pereboom's corresponding case. For instance, in my *Case 1* Betty *believes* that a team of neuroscientists have the ability to manipulate her neural states at any time by radio-like technology, and that they will in fact manipulate Betty's decision to press one of the buttons. Similarly, in *Case 1* of Pereboom's (2014, 76–77) FCMA, it just is the case that a team of neuroscientists have the ability to manipulate Plum's neural states at any time by radio-like technology, and that they manipulate Plum to kill (and decide to kill) White. Due to this kind of structural similarity between my argument and the FCMA, both arguments seem to stand or fall together. As a result, Pereboom faces the following dilemma: either give up (DC) as well as the view that one can (as a hard incompatibilist) rationally deliberate, or give up the FCMA.

Perhaps there is a way for Pereboom to resist the claim that both arguments stand or fall together. After all, my argument is concerned with rational deliberation, and Pereboom's FCMA is concerned with basic desert moral responsibility. So, perhaps Pereboom could claim that the kind of control one must believe one possesses (or at least must not believe that one does not possess) in order to rationally deliberate (henceforth '*RD control*') is *distinct* from the kind of control one must in fact possess for moral responsibility (henceforth '*MR control*'). As a result, perhaps Pereboom can consistently accept a soft-line or hard-line reply to my argument without thereby undermining the FCMA. Let us call this the *distinctness objection*.

The *distinctness objection* seems to imply that there is a possible world *w* in which *S* lacks *MR control*, and yet *S* believes *truly* that *S* possess *RD control*.[9] What, then, is missing in *w* in order for *S* to possess *MR control*? Since *S* believes *truly* in *w* that *S* possess *RD control*, the missing ingredient cannot be that *S*'s actions are causally determined by factors beyond *S*'s control, or that *S*'s actions are indeterministically caused by factors beyond *S*'s control, such that it is not up to *S* whether one action rather than another is indeterministically caused by factors beyond *S*'s control. Pereboom (2014, 51–54) does maintain that possessing *MR control* requires being (fundamentally) an agent-cause of one's action. So perhaps believing truly that one possesses *RD control* does not require being (fundamentally) an agent-cause of one's actions.

This move is *ad hoc*. If, as an agent-causal theorist, one holds that the mere truth of causal indeterminism does not confer any more control upon an agent in the sense required for responsibility, why think that the mere truth of causal indeterminism *does* confer any more control upon an agent in the sense required for believing truly that one possesses *RD control*? I see no satisfying answer to this question. At any rate, unless we are provided with some positive motivation for the *distinctness objection*, it should be relegated to the status of an *ad hoc* maneuver, devised solely for the purpose of escaping my proposed dilemma.

4.5 Conclusion

I offered a counterexample to Pereboom's account of rational deliberation, and then employed it to develop a four-case manipulated deliberation argument against (DC), and ultimately for the position that rational deliberation requires an agent to *lack* the belief that it is *not* up to her, in the

---

[9] If the *distinctness objection* does not have this implication, it would then follow that *RD control* and *MR control* are necessarily co-extensive which in turn would strongly suggest that these two forms of control are in fact identical.

libertarian sense, to choose what to do. I then noted a number of replies to my argument, and conceded that at least some of them are viable. However, given the structural similarity between my argument and Pereboom's FCMA, both arguments stand or fall together. So, while my argument against (DC) (and ultimately for (UT)) may not convince a non-libertarian, I claim to have at least shown that a hard incompatibilist such as Pereboom who defends the FCMA faces a vexing dilemma: either give up (DC) as well as the view that one can (as a hard incompatibilist) rationally deliberate, or give up the FCMA.

4.6 Appendix: Nelkin's Account of Rational Deliberation

As noted in section 4.2, Nelkin—a defender of (DC)—apparently would hold that Betty cannot rationally deliberate about which button to press in *Case 1*. One might therefore think that Pereboom could thus adopt Nelkin's account of rational deliberation in order to escape my dilemma. This is, however, a false hope. For, I will now argue that in order to render the intuitively correct verdict in *Case 1* that Betty cannot rationally deliberate about which button to press, Nelkin's account must explicitly incorporate a historical condition for rational deliberation. If I am correct, then, as explained in section 4.4, Pereboom cannot consistently adopt Nelkin's account and uphold the FCMA.

According to *Nelkin's Account*, the following condition, in conjunction with (NIB), is both necessary and sufficient condition for rational deliberation (Nelkin 2011, 142):

> **(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.

For one's deliberation to be the explanatory nexus among the relevant alternatives is roughly for one's deliberation to be a difference-maker among those alternatives, whereby such deliberation can provide certain *contrastive* explanations. For instance, my deliberation concerning what to order while perusing the lunch menu is an explanatory nexus with respect to what I order only if my deliberation can explain why I ordered an appetizer *rather than*, say, a salad. Now, does Betty satisfy (NIB) and (EN) in *Case 1* given her beliefs according to (iv)? I think that, unless a historical condition is explicitly built into (EN), the answer is 'yes'. Let me explain.

As noted in section 4.2, any plausible account of rational deliberation that is consistent with (DC) must grant the possibility of *S* rationally deliberating in the face of *S*'s belief that antecedent factors beyond *S*'s control momentarily enhance or diminish, e.g., *S*'s egoistic tendencies, whereby this momentary influence can be a difference-maker with respect to which action *S* performs. However, (according to what Betty believes) this is precisely how the neuroscientists are manipulating Betty's decision, viz. by exerting an egoism-enhancing or egoism-diminishing momentary influence upon Betty. So, in order for *Nelkin's Account* to be plausible, it must be possible for *S* to satisfy (EN) and (NIB) in the face of *S*'s belief that antecedent factors beyond *S*'s control enhances or diminishes, e.g., *S*'s egoistic tendencies, whereby this momentary influence can be a difference-maker with respect to which action *S* performs. But, again, (according to what Betty believes in *Case 1*) this is precisely how the neuroscientists are manipulating Betty's decision, viz. by exerting an egoism-enhancing or egoism-diminishing momentary influence upon Betty. So, it appears that in *Case 1* Betty satisfies (NIB) and (EN) after all, and thus that *Nelkin's Account* renders the intuitively incorrect verdict (even by Nelkin's lights) in *Case 1*. So, in order for *Nelkin's Account* to render the intuitively correct verdict in *Case 1*, a historical condition must be explicitly built into (EN)

which entails, *inter alia*, that if S believes that the factors beyond S's control that causally influence S's decision include the intentional actions of other agents, then S fails to satisfy both (NIB) and (EN).

So, if Pereboom were to adopt (EN) in an attempt to arrive at the intuitively correct verdict in *Case 1*, Pereboom would have to adopt a historical condition for rational deliberation along the lines suggested above. However, as explained in section 4.4, Pereboom could not then consistently object to a soft-line reply to the FCMA that endorses a similar historical condition for responsibility. So, even if Pereboom were to adopt a version of (DE) or (EN) that explicitly includes some historical condition for rational deliberation in order to render the intuitively correct verdict in *Case 1*, Pereboom would not thereby escape my dilemma.

5

Counterfactuals of Divine Freedom

Molinism purports to reconcile a robust account of divine providence with creaturely libertarian freedom.[1] The central Molinist thesis is that God has middle knowledge. This knowledge consists (at least) of counterfactuals of creaturely freedom (*CCF*s)—counterfactuals concerning what some creature would freely do if they were placed in certain circumstances. A distinctive feature of *CCF*s is that they have their truth-value both contingently and logically prior to any act of God's will.

While Molinism is perceived to have a number of fruitful applications in the philosophy or religion and theology,[2] it has been simultaneously plagued with a number of philosophical difficulties due to its commitment to true *CCF*s. Specifically, anti-Molinists argue that true *CCF*s do not exist because there is nothing to ground such truths (Adams 1977; Hasker 1989, 29–52; Cowan 2003),[3] or because they are ruled out by Lewis-Stalnaker possible world semantics (van Inwagen 1997).[4] Anti-Molinists have also argued that true *CCF*s are incompatible with libertarian freedom (Hasker 1986).[5] True *CCF*s are perceived to be a threat not only to free

---

[1] For the purpose of this paper I will be assuming that Molinism is committed to libertarianism. Though, see Perszyk's (2000) attempt to reconcile Molinism with compatibilism. Moreover, I will also be focusing exclusively on actions that agents can avoid performing. While not all Molinists may affirm the Principle of Alternative Possibilities (Frankfurt 1969, 829), it is a standard assumption within the Molinist framework that the kinds of free actions in the consequent of *CCF*s are ones in which an agent does in fact have the ability to do otherwise.

[2] For a summary of such applications, see Perszyk (2013, 764–765). Molinism has also been applied to areas outside of philosophy of religion, such as in Stump (2003, 148–150) and Hartman (ms).

[3] For a reply to this argument, see Flint (1998, 121–137), Wierenga (2001), Craig (2001), and (Merricks 2007, 146–155).

[4] For a reply to this argument, see Plantinga (1974, 178; 1985, 377–378).

[5] For a reply to this argument, see Flint (1999).

creatures, but also to certain features that *God* is supposed have. Specifically, true *CCF*s pose a threat to God's sovereignty or omnipotence since their truth-value is independent of any act of God's will (Mackie 1982, 174; Rogers 2007). Additionally, true *CCF*s pose a threat to God's omniscience given that God may have no means of knowing which *CCF*s are true (O'Connor 1992). In contrast to *CCF*s, very little attention has been given to counterfactuals of *divine* freedom (*CDF*s), and to what sorts of problems they may raise for Molinism. My aim is to investigate this rather unexplored territory.

I will argue that, contrary to the position of Luis de Molina, Thomas Flint, and others, *CDF*s are pre-volitional for God within the Molinist framework. That is, *CDF*s are not true even partly in virtue of some act of God's will. As a result, I argue that the Molinist God alone fails to satisfy an epistemic openness requirement for rational deliberation, and thus she cannot rationally deliberate about which world to actualize. Before turning to my argument, I must highlight certain aspects of Molinism that are crucial to our discussion of *CDF*s.

5.1 Molinism

According to Molinism, God's knowledge is divided into three types: natural, middle, and free. These types of knowledge stand in a non-temporal, logical or explanatory relation to one another with respect to God's decision to actualize a world. Moreover, these types of knowledge are individuated on the basis of both the modal status of their truth-value as well as the dependence status of their truth-value on acts of God's will. Thomas Flint (2003, 93) calls a proposition whose truth-value does not obtain in virtue of some act of a person *S*'s will a pre-volitional truth for *S*. Although it is not explicitly stated, I take Flint to mean that such a truth-value does not

obtain *even partly* in virtue of some act of *S*'s will. So, the notion of pre-volitionality I will be working with goes as follows:

> A truth *TR* is pre-volitional for a person *S* if and only if *TR* is not true even partly in virtue of some act of *S*'s will.[6]

By contrast, a truth that *does* obtain at least partly in virtue of some act of *S*'s will is post-volitional for *S*. With these two notions in hand, I now turn to the three types of divine knowledge within the Molinist framework.

First, God's *natural* knowledge consists of knowledge of necessary truths. These truths are pre-volitional for God. Second, God's *middle* knowledge consists of propositions that are, like God's natural knowledge, pre-volitional for God. But unlike the propositions of God's natural knowledge, the truth-value of these propositions is a contingent matter. Third, God's *free* knowledge consists of propositions whose truth-value is a contingent matter and which are post-volitional for God.

As noted above, the central propositions that are a part of God's middle knowledge are *CCF*s that have the following structure:

> If creaturely person *S* were in circumstance *C*, *S* would freely ϕ.

In order for these counterfactuals to be relevant to God's decision concerning which world to actualize, '*C*' must refer to a complete or maximally detailed description of the relevant circumstances since ordinary counterfactuals do not permit strengthening the antecedent; adding

---

[6] I employ '*TR*' to refer to a truth in order to reserve the term '*T*' to refer to a creaturely world-type.

more to the antecedent can change the truth-value of the counterfactual (Lewis 1973b, 31–32).

Hence, as Dean Zimmerman (2009, 56–59) puts it, it is the maximally described *ultima facie*

counterfactuals, rather than the non-maximally described *prima facie* counterfactuals, that are

useful to God.

　　Next, a creaturely world-type is, roughly, any complete set of *CCF*s that are compossibly

true. There are a number of ways to formulate a precise account of a creaturely world-type (Flint

1998, 46–50). The following relatively simple formulation of a creaturely world-type should

suffice for my purposes here:

> *T* is a creaturely world-type *iff* for any counterfactual of creaturely freedom ($C \,\square\!\!\rightarrow A$),
>
> either ($C \,\square\!\!\rightarrow A$) or ($C \,\square\!\!\rightarrow \sim\!A$) is a member of *T*, and all of the *CCF*s that are members
>
> of *T* are compossibly true.[7]

A creaturely world-type *T* is true *iff* all *CCF*s that are members of *T* are true. Now, the Molinist

should be open to the possibility of other propositions in addition to *CCF*s having their truth-

value contingently and being pre-volitional for God. For example, some Molinists entertain the

existence of contingently true, pre-volitional (for God) counterfactuals concerning

indeterministic scenarios at the microphysical level (Flint 1998, 42–43, 193–196; Craig 2007,

63; Perszyk 2011, 18). In that case, if we wish to equate God's middle knowledge with

knowledge of the true creaturely world-type, we must be open to the possibility of incorporating

---

[7] This is a modification of one of Flint's (1998, 49) formulations of a creaturely world-type. I add the compossibility component in order to rule out logically inconsistent creaturely world-types. Note that I will assume that the law of conditional excluded middle is true in order not to unnecessarily complicate things further. To be clear, how exactly one defines a creaturely world-type does not affect the core of my argument.

propositions other than *CCF*s into a creaturely world-type.[8] However, I argue later on that even

if *CDF*s are both contingently true and pre-volitional for God, God's middle knowledge should

not be equated with God's knowledge of the true creaturely world-type. The reason for this is

complex, and has nothing to do with the fact that the set in question is labeled a *creaturely*

world-type.

Next, a world *W* is feasible *iff* the creaturely world-type that is in fact true is true in *W* (Flint

1998, 51–54). It is in God's power to actualize a world only if that world is feasible. Since it is

not up to God which creaturely world-type is in fact true, God is able to choose which world to

actualize among the set of feasible worlds without in any way constraining creaturely libertarian

freedom. It is in this way that God's middle knowledge plays a central role in reconciling a

robust account of divine providence with the existence of creaturely libertarian freedom.

The final, less familiar notion that is crucial to this paper is resilience, a notion that Flint

(2003, 93) defines as follows:

> [A] truth *T*[*R*] is resilient for a person *S* if and only if *S* lacks counterfactual power over
>
> *T*[*R*]—i.e., if and only if it's not the case that *S* has or had the power to act in such a way
>
> that *T*[*R*] would not have been true.

The concept of resilience has been discussed primarily by Flint (2003) and Kvanvig (2002; 2011,

ch. 6) in regards to the coherence and plausibility of 'Maverick Molinism'. According to this

view, although *CCF*s are pre-volitional for God, they are non-resilient for God. In other words,

God could have acted in such a way such that, had she acted that way, certain *CCF*s would (or

---

[8] Flint (1998, 47) himself understands a counterfactual of creaturely freedom to include propositions in which a
being is not properly speaking free. So Flint in fact already grants that the objects of God's middle knowledge can
include propositions other than the ones I am labeling as a *CCF*.

might) have had a different truth-value from that which they actually have. I take no stand on the coherence of Maverick Molinism here. Instead, I will explore the relationship between *CCF*s and free creatures, as well as the relationship between *CDF*s and God with regards to the notion of resilience. So, with this brief sketch of Molinism in hand, I now turn to the finer details of the relationship between *CCF*s and free creatures. This will in turn help us understand the relationship between *CDF*s and God.

5.2 *CCF*s and Free Creatures

In this section I argue that the Molinist must accept two propositions concerning the relationship between *CCF*s and free creatures. Here is the first proposition:

(A) For any creaturely person *S*, every *CCF* about *S* is pre-volitional for *S*.

The Molinist must accept (A) for two reasons. Here is the first reason. According to the grounding objection, in order for *CCF*s to be true, they must be true in virtue of some categorical aspect of the world. But, so the objection goes, there is no such aspect to ground their truth (Adams 1977; Hasker 1989, 29–52; Cowan 2003). The objection relies on a grounding principle that excludes the grounding of *CCF*s. Here is an example of such a principle:

(GP) Any true contingent proposition is true in virtue of the existence or non-existence of some concrete state or event (Hasker 2004, 195; 2011, 27).

There are roughly two responses the Molinist can make to the grounding objection, and some

Molinists wish to remain neutral between which response ultimately ought to be endorsed. Either

response, however, commits the Molinist to accepting (A). First, she could maintain that *CCF*s

are grounded by replacing (GP) with the following principle:

> **(GP+)** Any true contingent proposition is true in virtue of some concrete state of affairs
>
> that does exist, or has existed, or will exist, or *would* exist (under specified conditions
>
> (Flint 2009, 281; cf. Wierenga 2001).

If the Molinist were to endorse (GP+), she would be committed to the claim that what (wholly)

grounds the truth of *CCF*s is the 'subjunctive aspect' of the world (whatever exactly that is). In

that case, *CCF*s are not true even partly in virtue of some act of a free creature's will, and thus

(A) is true. The second response to the grounding objection simply denies that all contingently

true propositions are true in virtue of some aspect of the world, and thus that *CCF*s are among

such brute or ungrounded truths (Plantinga 1985; Craig 2001; Merricks 2007, 146–155). *A

fortiori*, *CCF*s are not true even partly in virtue of some act of a free creature's will, and thus (A)

is once again true. So, no matter how the Molinist answers the grounding objection, the Molinist

is committed to (A).

I now turn to the second reason for why the Molinist must accept (A) which will only come

to light after a good deal of setup. According to Molinism, God decides which world to actualize

by considering all of the true *CCF*s. However, if any *CCF* were true even partly in virtue of

some act of a free creature's will, then that *CCF* would be true *too late*, so to speak, for the

purpose of deciding which world to actualize. But at least certain *CCF*s *are* true too late (Adams

1977, 113–114; Kenny 1979, 69–71), viz. those that are entailed by the truth of the antecedent and consequent of said *CCF*s. This has been coined by Wierenga (1989, 148) as the 'not true soon enough' objection. Let us flesh out this objection with an example. Suppose that Martha is in (maximally specified) circumstances *C* such that she has the opportunity to either go hiking or go fishing, and she in fact freely decides to go hiking. In that case, the following *CCF* is true:

> ***Hiking*** If Martha were in *C*, she would freely decide to go hiking.

The Stalnaker-Lewis possible worlds semantics, a formal theory, is often accompanied by a metaphysical theory concerning the truth-makers for counterfactuals (cf. Mares and Perszyk 2011, 97–101). Hence, it is widely assumed that the truth of *Hiking* depends upon the truth of the antecedent and consequent of *Hiking*. But then *Hiking*—and every additional *CCF* that is entailed by the truth of its antecedent and consequent—apparently cannot be true logically prior to creation, and thus cannot play a role in God's decision to actualize a world.[9] The argument may be formulated as follows:

1. The truth of *Hiking* depends at least partly upon the truth of the antecedent of *Hiking*.

2. The truth of the antecedent of *Hiking* depends at least partly upon which world God actualizes.

3. Therefore, the truth of *Hiking* depends at least partly upon which world God actualizes. [(1) & (2)]

4. If (3) is true, then truth of *Hiking* cannot be of use to God.

---

[9]Despite the somewhat misleading name, *CCFs* are really subjunctive conditionals that can have antecedents that are not contrary to fact (Lewis 1973b: 3–4).

5. Therefore, the truth of *Hiking* cannot be of use to God. [(3) & (4)]

Plantinga's (1985) reply to the 'not true soon enough' objection reveals that he thinks the inference from (1) and (2) to (3) is invalid because the notion of dependence at issue here is not transitive. Plantinga (1985, 376) offers the following counterexample to the transitivity of dependence:

1*. The truth of *The Allies won the Second World War* depends on which world is actual.

2*. Which world is actual depends on whether I mow my lawn this afternoon.

Therefore,

3*. The truth of *The Allies won the Second World War* depends on whether I mow my lawn this afternoon.

Since (1*) and (2*) are true but (3*) is clearly false, Plantinga (1985, 376) concludes that "the relation expressed by the relevant sense of 'depends' isn't transitive."[10]

Plantinga's alleged counterexample is unconvincing. The dependence relation, like explanation, is widely understood to be non-monotonic (Rosen 2010, 116–117; Raven 2012, 689). That is, roughly, if something *A* depends upon set *s* whose sole members are *W*, *X*, and *Y*, then each member of *s* must play a role in *A*'s dependence upon *s*. Hence, if *Z* plays no role in what *A* depends upon, then *A* does not depend upon set *s** whose sole members are *W, X, Y,* and *Z*. The upshot is that we should reject (1*); since Plantinga's mowing his lawn in the afternoon plays no role in what grounds the truth of *The Allies won the Second World War*, the truth of *The Allies won the Second World War* does not depend upon which world is actual. Instead, it only

---

[10] Plantinga's reply to the 'not true soon enough' objection is also endorsed by Wierenga (1989, 148–150).

depends upon a *proper part* of the actual world. So, the Molinist should not reply to the 'not true soon enough' objection by rejecting (3). Instead, she can and should reject premise (1). Let me explain.

The Molinist is committed to the truth of *CCF*s whose antecedents are not true. At least *these* counterfactuals are not true in virtue of any act of a free creature's will, as we have seen in the previous discussion of the grounding objection. Hence, in response to the 'not true soon enough' objection, the Molinist should adopt the exact same position in regards to *CCF*s whose antecedents are true. In other words, although the truth of *Hiking* is *entailed* by the truth of the antecedent and consequent of *Hiking*, the Molinist should deny that the truth of *Hiking* even partly *depends upon* the truth of the antecedent and consequent of *Hiking*. This way, the Molinist can treat all *CCF*s in the same manner *and* provide a principled reason for rejecting (1). In that case, since no *CCF* depends upon the truth of either its antecedent or consequent, no *CCF* depends upon an act of a free creature's will. So no *CCF* about a creaturely person *S* is post-volitional for *S*. This concludes my second reason for thinking that the Molinist must accept (A).

The other proposition I think that the Molinist must accept is the following:

(B) For any creaturely person *S*, if *S* freely ϕ-s in circumstance *C*, then the following *CCF* is non-resilient for *S*: 'if *S* were in circumstance *C*, *S* would freely ϕ'.

To see why the Molinist must accept (B), consider the following two *CCF*s (note that 'the restaurant' is shorthand for a maximally specified description of certain circumstances):

Pizza If Roy were in the restaurant, he would freely order pizza (rather than spaghetti).

Spaghetti If Roy were in the restaurant, he would freely order spaghetti (rather than pizza).

Suppose that *Pizza* is true, *Spaghetti* is false, and that Roy does in fact freely order pizza in the restaurant. Since Roy could have ordered spaghetti instead, and since Roy's ordering spaghetti in the restaurant entails that *Spaghetti* is true, it follows that Roy has some sort of counterfactual power over the truth-value of *Pizza* and *Spaghetti*. In other words, there is something Roy could have done (viz., order spaghetti rather than pizza), such that if he were to have done it, *Pizza* would have been false and *Spaghetti* would have been true (Flint 1999, 303). In that case, *Pizza* is non-resilient for Roy. Generalizing from this case, it follows that (B) is true. So I conclude in this section that, within the Molinist framework, (A) and (B) are true. I now turn to discuss the relationship between *CDF*s and God.

5.3 *CDF*s and God

Recall that a *CCF* has the following structure:

If a creaturely person *S* were in circumstance *C*, *S* would freely ϕ.

Similarly, the relevant kind of *CDF* with which I will be concerned has the same structure:

If God were in circumstance *C*, she would freely ϕ.

Now, Flint (1998, 55) and others conceive of "God's decision as to what he will do [to be] a single, all-encompassing one which follows upon his middle knowledge and precedes his free knowledge." I will thus understand 'ϕ' to stand for actualizing some world and 'C' to stand for the logical moment that immediately precedes God's act of will.[11],[12] Thus, C is be individuated, at least in part, on the basis of which creaturely world-type happens to be true. However, as I noted earlier, there may be other objects of God's middle knowledge besides *CCF*s. In that case, it would seem at first glance that C should be individuated on the basis of God's exhaustive middle knowledge. But this isn't exactly right either. For, *if CDF*s are pre-volitional for God (and contingently true), C should be individuated on the basis of God's exhaustive middle knowledge, *with the exception of God's knowledge of the true CDFs.*[13]

To see why the true *CDF*s should be excluded from C, it will be instructive to return for a moment to *CCF*s. Suppose once more that *Pizza* is true and that Roy freely orders pizza (rather than spaghetti) in the restaurant. Now, as I noted earlier, the salient *CCF*s for Molinism are the *ultima facie* counterfactuals, rather than the non-maximally described *prima facie* counterfactuals. This is why I noted earlier that 'the restaurant' is shorthand for a maximally specified description of certain circumstances. Now, does 'the restaurant' include the truth of the very counterfactual under discussion, viz. *Pizza*? No. For, if 'the restaurant' *did* include the truth

---

[11] For simplicity's sake, I bypass Plantinga's (1974, 169–174) weak/strong actualization distinction that is no doubt crucial for a full-fledged account of Molinism.

[12] Notice that Flint would presumably still want to grant the existence of what I'll call *local CDF*s—*CDF*s that are more narrow in scope such as the following:

> If Suzie were to freely request God to cure her of some illness, God would freely grant Suzie's request.

If one thought of God's interaction with the world as one single, comprehensive act of will, but one also wanted to affirm the existence of such local *CDF*s, one may want to posit a kind of priority of the true global *CDF* over the true local *CDF*s. For reasons I will not discuss here, I think that the true global *CDF* is not in fact prior to the true local *CDF*s within the Molinist framework (Otte 2009; Plantinga 2009; Wierenga 2011, 130). Note, however, that I will be focusing on global *CDF*s rather than local *CDF*s *merely for simplicity's sake*, as nothing I say below hinges on which kinds of *CDF*s are logically prior.

[13] I am indebted to Ed Wierenga and Kenny Boyce for helpful discussions on this issue.

of *Pizza*, then 'the restaurant' would simply *entail* that Roy orders pizza. However, within the Molinist framework, if Roy *freely* orders Pizza in the restaurant, then Roy could have refrained from ordering pizza in the restaurant. So, 'the restaurant' cannot entail the fact that Roy orders pizza. The more general point here is that *the circumstances specified in the antecedent of a CCF do not include the truth of that very CCF* (Wierenga 2011, 127–130).

In light of these considerations, we should similarly hold that *the circumstances specified in the antecedent of a CDF do not include the truth of that very CDF*. So, even if *CDF*s are pre-volitional for God (and contingently true), circumstance *C* in the antecedent of a *CDF* should be individuated on the basis of God's exhaustive middle knowledge, *with the exception of God's knowledge of the true CDFs*. So, the creaturely world-type should not include *CDF*s, even if it includes other propositions besides *CCF*s that are both contingently true and pre-volitional for God. With this understanding of *CDF*s in hand, I now turn to the more detailed kind of *CDF* I will be discussing in this paper:

If creaturely world-type *T* were true, God would freely actualize world *W*.

Now, unless there is some relevant difference between God and free creatures, just as the Molinist must affirm that *CCF*s are both pre-volitional and non-resilient (in certain circumstances) for free creatures, the Molinist likewise should maintain that *CDF*s are both pre-volitional and non-resilient (at least in certain circumstances) for God. In other words, barring some relevant difference between God and free creatures, just as the Molinist is committed to (A) and (B), the Molinist also ought to accept the following two propositions:

(A*) Every *CDF* is pre-volitional for God.

(B*) If God freely ϕ-s in circumstance *C*, then the following *CDF* is non-resilient for

God: 'if God were in circumstance *C*, God would freely ϕ'.

No Molinist should have a problem accepting (B*). To illustrate, suppose that creaturely world-

type T_1 is true, and that God actualizes world W_1. In that case, the following *CDF* is true:

(i) If T_1 were true, God would freely actualize world W_1.

Since God freely actualizes W_1, God could have actualized some other (feasible) world instead,

say, W_2. If God did that, then (i) would be false, and the following *CDF* that is in fact false

would have been true instead:

(ii) If T_1 were true, God would freely actualize world W_2.

So God has a kind of counterfactual power over the truth of (i). So (i) is non-resilient for God.

Generalizing from this case, it follows that (B*) is true. Now, while (B*) is not controversial,

many Molinists seem to assume with Molina (1998, 173/*Concordia* 4.52.11.) and Flint (1998,

56–57) that (A*) is false. Such Molinists maintain instead that *CDF*s are post-volitional for God.

The only Molinist I know of who accepts (A*) is one of Molina's contemporaries, Francisco

Suarez (1741) (cf. Craig 1988, 225–226).

Now, as previously stated, unless some relevant difference can be found between God and

free creatures, the Molinist must accept (A*) on pain of arbitrariness. Is there a relevant

difference between God and free creatures that might block the inference from (A) to (A*)? I fail to see what that difference might be.[14] At any rate, the burden seems to be squarely on the Molinist to provide some difference between God and free creatures that in turn casts doubt on the inference from (A) to (A*). As Freddoso (1988, 53) notes, if Molinists reject (A*), they "are burdened with the task of finding some way to explain the fact that God has prevolitional cognition of the free actions of creatures but lacks prevolitional cognition of His own free actions." What I will do in the next section, then, is assess the reason Molina and Flint offer for rejecting (A*), the only reason offered in the literature. I will argue that their case against (A*) is unconvincing precisely because they do not mark a relevant difference between free creatures and God. If I am correct, then the Molinist must indeed accept (A*) on pain of arbitrariness.

5.4 Are *CDF*s Post-Volitional For God?

For ease of exposition, I will refer to a *CDF* that is pre-volitional for God as a 'Pre-*CDF*', and refer to a *CDF* that is post-volitional for God as a 'Post-*CDF*'. Now, Molina (1988, 171–175/*Concordia* 4.52.11–13) thought that Pre-*CDF*s are incompatible with God's freedom, despite also affirming that pre-volitional *CCF*s for free creatures are compatible with creaturely freedom. I take Molina's argument against Pre-*CDF*s to go roughly as follows. Assume for the sake of argument that Pre-*CDF*s exist. In that case, logically prior to God's decision of what to do, God knows both that creaturely world-type *T* is in fact true, and also that if *T* were true God would actualize world *W*. So God knows logically prior to God's decision of what to do that God will (or does) actualize *W*. But this is incompatible with God's freely actualizing *W*: "For if such

---

[14] Thanks to Mike Rea for the suggestion that a proponent of divine timelessness might note as a relevant difference that the circumstances in the antecedent of a *CDF* are non-temporal, whereas the circumstances in the antecedent of a *CCF* are temporal. While the temporal/atemporal distinction certainly marks *a* difference between *CCF*s and *CDF*s (given divine timelessness), it is difficult to see how this difference is *relevant* to the issue of truth-making. And, until some reason is given for thinking it is, this difference cannot block the inference from (A) to (A*).

knowledge [of Pre-*CDF*s] existed, then He would in no way be able to choose the [alternative action]" (Molina 1998, 171). The argument may be constructed as follows:

The Argument Against Pre-CDFs

6. If Pre-*CDF*s exist, then God knows that God ϕ-s logically prior to God's ϕ-ing.

7. If God knows that God ϕ-s logically prior to God's ϕ-ing, then God does not ϕ freely.

8. Therefore, if Pre-*CDF*s exist, then God does not ϕ freely. [(6) & (7)]

9. God does ϕ freely.

10. Therefore, it's not the case that Pre-*CDF*s exist. [(8) & (9)]

Flint agrees with Molina that (A*) is false within the Molinist framework, but does not offer much of an argument to work with. Rather, Flint (1998, 57, fn 27) simply asserts that "the conditional which indicates what God would do given the *true* creaturely world-type *cannot* be true independent of God's free act of will" (emphasis in the original). I presume that Flint wishes to endorse something like *The Argument Against Pre-CDFs*. At any rate, my hope is that assessing the above argument will explain why Flint's commitment to Post-*CDF*s is unfounded. Let us then turn to that argument.

While open theists may find (7) attractive, it is extremely difficult to see how the Molinist can accept (7). This is because Molinists are committed to the compatibility of creaturely freedom and divine knowledge of the true *CCF*s logically prior to God's creative act of will. In that case, unless some relevant difference can be found between God and free creatures, the Molinist should presumably also commit to the compatibility of divine freedom and divine

knowledge of the true *CDF*s logically prior to God's creative act of will. I will now further elaborate on this point.

Is the truth of (A) incompatible with creaturely libertarian freedom? The Molinist is committed to saying no since she is committed to (B), the claim that, roughly, *CCF*s are non-resilient for free creatures. That is, the Molinist is committed to the claim that free creatures could have done otherwise, and had they done otherwise, certain *CCF*s that are in fact true would have been false, *despite* the fact that such *CCF*s are true logically prior to God's creative act of will.

Now, in regards to the question of whether (A*) is incompatible with divine freedom, the Molinist can and should provide a similar answer: given the truth of (B*), God could have actualized a different feasible world, and had God done so, at least one *CDF* that is in fact true would have been false, *despite* the fact that that *CDF* is true logically prior to God's creative act of will. We can thus see that if Molina and Flint wish to assert that (A*) and (B*) cannot both be true, they must explain how (A) and (B), by contrast, *can* both be true. And, as I have stressed in the previous section, in order to accomplish this task, they must point to some relevant difference between God and free creatures. But no such difference is provided within *The Argument Against Pre-CDFs*. So we have not been given any good reason to accept (7)—and ultimately to accept the existence of Post-*CDF*s—within the Molinist framework.

Perhaps the Molinist can establish (7), not by marking some relevant difference between God and free creatures, but rather by showing that there is a relevant difference between an agent *S* foreknowing what *S herself* will do, rather than foreknowing what some other agent *S** will do. In other words, perhaps *de se* foreknowledge alone is incompatible with freedom.

Notice that if the Molinist were to defend (7) in the above manner, she would presumably have to admit that if God provided each of us with knowledge of *our* future actions—perhaps by providing each of us with a book like '*The Life of Osmo*' (Taylor 1991, 58–63)—then we would not be able to do otherwise. Such a method for defending (7) thus comes at a serious cost, and also seems to go against the spirit of Molinism.

More importantly, however, I do not see what is unique to *de se* foreknowledge such that it precludes the possibility of a free action. This is because it is difficult to see how *gaining* knowledge can *reduce* one's abilities.[15] Perhaps losing one's abilities is possible when one attains knowledge about the *consequences* of certain actions. For example, if I come to learn that pressing a certain elevator button will result in the detonation of a nuclear bomb that will end millions of lives, it is not implausible to suppose that I had but now lack the ability to press that button given my subjective motivational set. However, God's knowledge does not fit this description; God's *de se* foreknowledge has nothing to do with the *consequences* of acting in some manner. I thus conclude that *The Argument Against Pre-CDFs* fails because premise (7) has not been established in a way that does not undermine the compatibility of (A) and (B).

To be clear, I think there *is* a case to be made for the incompatibility of (A) and (B), as well as for the incompatibility of (A*) and (B*). Elsewhere I (Cohen 2015, 243) have suggested that, roughly, in order for it to be up to an agent whether a proposition *p* is true, *p* must be true at least partly in virtue of something the agent does (or refrains from doing), or in virtue of something that was caused by something the agent did (or refrained from doing). But the important point to grasp for our purposes here is that the Molinist's commitment to both (A) and (B) precludes her from endorsing any such principle. More generally, any reason for doubting the compatibility of

---

[15] I certainly grant that gaining knowledge can *enhance* one's abilities; see Shabo (2014).

(A*) and (B*) is irrelevant to the present discussion *if* it similarly casts doubt on the compatibility of (A) and (B).

Turning now to a slightly different issue, one general concern with *The Argument Against Pre-CDFs* is that it does not explain *why* Pre-*CDF*s do not exist. For, even if it *could* be shown within the Molinist framework that Pre-*CDF*s are incompatible with divine freedom (and we assumed that God is in fact free), we would not thereby have an explanation for *why* Pre-*CDF*s do not exist, but only a demonstration *that* Pre-*CDF*s do not exist. An explanation for *why* Pre-*CDF*s do not exist would presumably appeal to the nature of propositions, truthmaking, and so forth.[16]

Admittedly, Molina does appeal to the doctrine of supercomprehension in order explain why, logically prior to God's actualizing some world, God does not know which world God actualizes. Accordingly, in order to know directly (i.e. not through, say, the testimony of another reliable source) not just what a possible person could do, but rather what they *would* do under any possible circumstance, one must supercomprehend the essence of that person which in turn requires infinitely surpassing that person, at least in intellect.[17] Molina (1988, 174/*Concordia* 4.52.13) thus appeals to this doctrine to mark a relevant difference between God and free creatures:

> [W]hile the divine intellect and knowledge surpass in perfection by an infinite distance
> each *created* faculty of choice which they contain eminently in themselves and which for

---

[16] The point I am making here mirrors Patrick Todd's (2014) claim (which is made in a more detailed and eloquent manner) that the alleged incompatibility of divine foreknowledge and creaturely freedom does not explain *why* divine foreknowledge does not exist, but only demonstrates *that* it does not exist (given the assumption that we are in fact free).

[17] For further discussion of this doctrine, see Craig (1988: 178–183), Freddoso (1988: 51–53), and Flint (1998: 56–57).

this reason they comprehend in a certain infinitely more eminent way than that in which

it is knowable, they do not likewise surpass the *divine* will in perfection or comprehend it

in a more eminent way than that in which it is knowable itself.

Let us grant just for a moment that the doctrine of supercomprehension is not implausible. It is

important to see that *even if* God does not know which world she actualizes logically prior to

actualizing some world, it does not follow that Post-*CDF*s exist. To suppose otherwise is to a

conflate metaphysics with epistemology. For, recall that the Pre-*CDF*/Post-*CDF* distinction

concerns what, if anything *makes CDF*s true. By contrast, the doctrine of supercomprehension, at

best, points to which logical moment God *knows* the true *CDF*s, irrespective of whether they are

Pre-*CDF*s or Post-*CDF*s.

   More importantly, however, I find the doctrine of supercomprehension to be, in the words of

Flint (1998, 56, fn 26), "murky and unhelpful", to say the least. For, it is difficult to see, among

other things, why one must infinitely surpass some person in intellect in order to know (directly)

certain truths about that person. At any rate, until we are given a contemporary defense of the

doctrine that appeals to metaphysical and epistemological concepts with which we are at least

somewhat familiar, our credence in the doctrine of supercomprehension should be low.

   In the next section, I will defend Derk Pereboom's epistemic openness requirement for

rational deliberation, and then amend it in certain respects for our purposes here. Then, in the

final section I will show that, given the truth of (A*) within the Molinist framework, the Molinist

God fails to satisfy the epistemic openness requirement, and thus cannot rationally deliberate

about which world to actualize.

5.5 The Epistemic Openness Requirement for Rational Deliberation

At a bare minimum, rational deliberation is a mental process that consists of figuring out what to do. Figuring out what to do seems to presuppose that there is more than one possible action under consideration. In order for a possible action ϕ to be under consideration for a subject S with respect to figuring out what to do, it seems that S must not believe that S cannot ϕ. Must S instead believe that S can in fact ϕ? It seems sufficient for S to simply be agnostic about whether or not S can ϕ (Kapitan 1986; Pettit 1989). Rational deliberation also seems to have an important connection to the reasons an agent has for performing certain actions under consideration. The exact connection between deliberation and reasons for actions is an issue I will explore both here and in the next section.

Now, Pereboom (2008, 294) defends the following epistemic-openness requirement for rational deliberation:[18]

> **(S)** In order to deliberate rationally among distinct actions $A_1 \dots A_n$, for each A_i, S cannot be certain of the proposition that she will do A_i nor of the proposition that she will not do A_i; and either (a) the proposition that she will do A_i is consistent with every proposition that, in the present context, is settled for her, or (b) if it is inconsistent with some such proposition, she cannot believe that it is.

The word 'settled' in (S) is used as a technical term to be understood in the following manner (Pereboom 2008, 294):

---

[18] While I will focus on Pereboom (2008), see also Pereboom (2014, ch. 5). For other defenses of an epistemic openness requirement, see Hampshire and Hart (1958), Ginet (1962), Taylor (1964, 75), Kaufman (1966), and Kapitan (1986, 235–241).

(Settled) A proposition is settled for an agent just in case she believes it and disregards any doubt she has that it is true, e.g., for the purpose of deliberation.

In support of (S), Pereboom (2008, 295) says that the best explanation for why he cannot rationally deliberate about (e.g.) presently becoming a mercenary in Africa is that he fails to satisfy the condition in (S). Pereboom does not satisfy the condition in (S) because the proposition that he will now become a mercenary is inconsistent with his (occurent or dispositional) beliefs concerning his values and his character.

Dana Nelkin (2011, 137) thinks, however, that there are counterexamples to (S):[19]

[S]uppose that you are engaged in an enjoyable activity (watching the seventh game of the world series, reading the seventh Harry Potter book, dining at your favorite restaurant), and you receive a call from your friend in need. It seems possible that you could know what you are going to do, and even be certain of what you will do (help your friend) and so, *a fortiori*, be settled with respect to this fact, without having made the decision or formed the intention to do it. I am not sure why this isn't coherent: one simply has no doubts about what one is going to choose, and still not yet have chosen from among one's alternatives.

Given Pereboom's (2008, 293) insistence that if an agent is certain of what she will do, the agent cannot still figure out what to do, Nelkin suggests that we should thus understand Pereboom's

---

[19] See also Clarke's (1992) rejection of an epistemic openness requirement for rational deliberation, as well as Henden's (2010) response.

conception of 'figuring out what to do' as a mere epistemic notion of figuring out what one will in fact do. Nelkin (2011, 138), by contrast, maintains that the notion of 'figuring out what to do' involves something more robust, viz. an aim at adopting reasons, or aiming at the best (or a good) action, where this notion of 'figuring out what to do' is consistent with having no doubts about what one is going to do.

I offer two responses to Nelkin in defense of (S). First, we need not interpret Pereboom's notion of 'figuring out what to do' as a mere epistemic activity of finding out what one will in fact do. Rather, Pereboom can agree with Nelkin that 'figuring out what to do' involves aiming at adopting reasons for action. However, if there is absolutely no doubt in S's mind that S will in fact ϕ, all that is left for S to do is find out what reasons S has for doing what S has no doubt that S will in fact do, viz. ϕ. This does not sound like deliberation at all precisely because it is already settled for S that S will ϕ.[20] Rather, figuring out what to do seems to involve figuring out what the reasons are for ϕ-ing *in order to determine whether or not to ϕ*. This reading of Pereboom's conception of 'figuring out what to do' involves more than just an epistemic activity.

Second, in Nelkin's case of receiving a call from a friend, notice that even if it is settled for you that you are going to help your friend, typically there are still many *ways* in which you can help your friend. Consequently, there is still much to deliberate about. For example, suppose I receive a call from my friend who is in need of help while I'm dining at my favorite restaurant. Even if I am certain that I will help my friend, I might still deliberate about any of the following things: whether to get dessert before helping my friend, whether to have one final word with the person with whom I am dining before helping my friend, how long I am willing to spend with my friend, and whether I will return to the plans I had in mind for the evening after helping my

---

[20] This point may be applied to cases that Clarke (1992, 108) and Pendergraft (2014, 348–349) offer against (S). Henden (2010, 323–324) makes a similar point in response to Clarke's case.

friend. So, I claim that Nelkin's example does not undermine (S) once we see that it is easy to conflate deliberation about *how* to help your friend with deliberation about *whether* to help your friend (cf. Kapitan 1991, 110).

As further support for this claim, consider an alternative scenario where I must push a button exactly at time *t* in order to save my friend's life, despite the fact that pushing the button at *t* will cause me some pain. Since I believe that I am a morally decent person, it is settled for me that I will in fact press the button at *t* in order to save my friend's life. Moreover, it doesn't appear to me that I can deliberate about whether to push the button at *t*, even if I can still inquire into all of the reasons I in fact have for pressing the button at *t*. Why is it clearer in this case rather than in Nelkin's case that one cannot rationally deliberate? It is precisely because, in my case, it is more difficult to conflate deliberation about *how* to help my friend with deliberation about *whether* to help my friend since there just aren't multiple ways in which I can help my friend. It seems, then, that Nelkin has not offered a strong reason for rejecting (S).

Suppose, however, that you remain convinced of Nelkin's alleged counterexample to (S). It doesn't follow that a weaker analogue of (S) is also false, according to which absolute certainty about the fact that one will ϕ is incompatible with rationally deliberating about whether to ϕ. Nelkin (2011, 127) notes that such a proposal purchases immunity from counterexamples at the price of explanatory power since we are certain of very little. But this is of course an important concession for our concerns here since, for any proposition *God* believes to be true, God is surely certain of that proposition. Hence, despite my previous remarks, I will grant for the sake of argument that (S) is false and instead affirm the following weaker analogue of (S) that even Nelkin seems to admit is immune from counterexamples:

(S*) In order to deliberate rationally among distinct actions $A_1 \ldots A_n$, for each A_i, S cannot be certain of the proposition that she will do A_i nor of the proposition that she will not do A_i; and either (a) the proposition that she will do A_i is consistent with every proposition of which S is certain, or (b) if it is inconsistent with some such proposition, she cannot believe that it is.

Now, I need to modify (S*) in order to apply an epistemic openness requirement to the Molinist God. For, recall that the Molinist God proceeds in logical stages rather than temporal ones. We thus need an epistemic openness requirement that takes such logical stages into account:

(S**) In order to deliberate rationally among distinct actions $A_1 \ldots A_n$, for each A_i, at logical stage l, S cannot be certain of the proposition that she does A_i at logical stage l^* (such that l^* is posterior to l) nor of the proposition that she does not do A_i at l^*; and either (a) the proposition that she does A_i at l^* is, at l, consistent with every proposition of which S is certain, or (b) if it is inconsistent with some such proposition, at l, she cannot believe that it is.

We are finally ready to apply the results of this section, in conjunction with my defense of (A*), to Molinism.

5.6 The Molinist God Cannot Rationally Deliberate

At a first glance, one might think that only an open theist God can satisfy the condition in (S**). But that is a mistake. To illustrate, consider the simple foreknowledge view, according to which,

at all points in time, God has complete foreknowledge of everything that will occur, including of what God will do. Despite this exhaustive foreknowledge at all times, the simple foreknowledge God nevertheless proceeds in non-temporal, logical or explanatory stages in order to deliberate about what to do in the future (Hunt 1993; cf. Zimmerman 2012). Accordingly, even on the simple foreknowledge view there are logical moments at which God is not certain of what God will do at a later time (or what God does at a later logical or explanatory moment). Hence, the simple foreknowledge God can satisfy the condition in (S**), despite having exhaustive foreknowledge at all times. Besides open theism and the simple foreknowledge view, one could arguably consistently maintain that God satisfies the condition in (S**) given other accounts of divine providence, such as divine timelessness (Stump and Kretzmann 1981) and theological determinism (Pereboom 2011). However, as I will now show, the Molinist God does not satisfy the condition in (S**).

Suppose that the circumstances God is in that are logically prior to God's decision of which world to actualize are such that creaturely world-type T_1 is true. Furthermore, given the existence of Pre-*CDF*s (recall my defense of (A*)), suppose that God knows the truth of at least the following three *CDF*s logically prior to God's decision of which world to actualize:

(I) If T_1 were true, God would freely actualize world W_1.

(II) If T_2 were true, God would freely actualize world W_2.

(III) If T_3 were true, God would freely actualize world W_3.

So, logically prior to God's free knowledge, she knows both that T_1 is true, and also that (I) is true. Given that God's knowledge is closed under entailment, God also knows logically prior to

God's free knowledge that God actualizes W_1.[21] In other words, logically prior to her decision of what to do, God is certain of the truth of the proposition that God actualizes W_1. So God fails to satisfy the condition in (S**) that is required for rational deliberation. Generalizing from this case, it follows that the Molinist God cannot rationally deliberate about which world to actualize.

Somewhat ironically, it has been widely assumed that the "doctrine of divine middle knowledge…affords room for a sort of deliberation on God's part logically prior to His decision to actualize a world" (Craig 1991, 278). But, as I have argued, matters seem to be quite the reverse; it is precisely the doctrine of divine middle knowledge, and certain consequences thereof such as the existence of Pre-*CDF*s, that in fact *preclude* God from possessing the ability to rationally deliberate.

How costly it is for Molinism that God cannot rationally deliberate? If, as some think, rational deliberation is required for intentional action, this consequence would indeed be a significant cost.[22] There are further reasons for taking God's ability to rationally deliberate to be a *desideratum* for a theory of divine providence. David Basinger (1986, 171) goes so far as to say that he is aware of no theist that has ever granted that God's choices of what to do "were *never* formulated as the result of any sort of temporal or timeless deliberation on his part", and then further suggests that most theists would not view God's responses to prayer to be a non-deliberative manner.

While Kvanvig (2011, 105) states that an omniscient and omnipotent being has no need to deliberate, Kvanvig also says that we need a deliberational *model* of God's activity in creation. Such a model will be adequate only if God's omniscience is explanatorily dependent upon God's

---

[21] There is apparently no knowledge left for God to have that is to be categorized under God's free knowledge given the existence of Pre-*CDF*s. This is an interesting consequence that I will not further explore here.

[22] In defense of the claim that intentional action requires rational deliberation, see Kapitan (1991; 1994), Basinger (1986), and Peterson, et al. (2009, 81). Against this position, see Reichenbach (1984), and Hunt (1992; 1996).

creative activity (Kvanvig 2011, 108). But the Molinist cannot secure even this condition for an adequate deliberational model since the existence of Pre-*CDF*s seems to render God's omniscience explanatorily *independent* of God's creative activity precisely because God knows which world God actualizes logically prior to God's decision to actualize a world. So, I have still shown that, in accordance with Kvanvig's remarks, given the existence of Pre-*CDF*s within the Molinist framework, Molinism cannot offer an adequate deliberational model of God's interaction with the world.[23]

Now, while I cannot possibly offer an adequate assessment of these issues here, it is safe to say that there is no clear consensus as to how costly it is for a theory of divine providence that it implies that God cannot rationally deliberate, or that it cannot offer an adequate deliberational model of God's activity in creation. I want to conclude, however, by considering an argument the Molinist might be tempted to endorse in order to explain why, unlike for finite beings like us, God has no use for rational deliberation, even understood as a progression in logical or explanatory moments rather than temporal ones from a state of indecision to a further state of resolving such indecision. I call this *The No Use Argument*:

The No Use Argument

11. If there is some world that God has most reason, all things considered, to actualize, then God can know what which world that is without deliberating.

12. There is some world that God has most reason, all things considered, to actualize.

---

[23] It is worth nothing that the notion of deliberation Kvanvig seems to be focusing on is a *temporal* progress from a state of indecision to further state in which the indecision is resolved. I think that we ought to think of the kind of deliberation that we undergo as strongly analogous to a progression in *logical* moments or stages (rather than temporal ones) from a state of indecision to a further state in which the indecision is resolved. Hence, Kvanvig's remarks are consistent with the claim that it would be useful for God to rationally deliberate, *if* the notion of rational deliberation employed here is understood analogically (cf. Swinburne 1993).

13. Therefore, God can know which world God has most reason, all things considered, to actualize without deliberating. [(11) & (12)]

14. The sole purpose of rational deliberation is to aid one in figuring out what one has most reason, all things considered, to do.

15. If (13) and (14) are true, then God has no use for rational deliberation.

16. Therefore, God has no use for rational deliberation. [(13)–(15)]

Premise (12) might be thought to be plausible under the assumption that there is a best possible world. But that assumption is highly contentious. Instead, there is good reason to think that either there is a 'tie at the top', an infinite hierarchy of better worlds, or set of worlds that are incommensurable with one another in terms of their overall value (Senor 2008, 187–194). Suppose, however, that there *is* a best possible world. Does the existence of such a world show that there is something that God has most reason, all things considered, to do? Not necessarily. If Molinism is true, then, possibly, God cannot actualize the best possible world because it is not feasible, and thus God would not have *any* reason to actualize that world.[24]

Now, there *may* be a best feasible world even if the best possible world is not feasible. But this would only be a contingent matter. For, possibly, God is stuck with a (contingently) true creaturely world-type such that there is no best feasible world. In that case, according to Molinism it is possible that (12) is false and also possible that (12) is true. But in that case, the Molinist is not rationally permitted to accept (12), barring some sort extraordinary evidence for the claim that actual world happens to be the best feasible one.[25]

---

[24] I am implicitly appealing to a principle that is defended, among others, by Haji (2012, 24): "If *S* has most reason to do something, *A*, and, thus, if *S* reasons-wise ought to do *A*, then *S* can do *A*."

[25] It is surely also too big of a concession for the Molinist to assert that God just *happens* to have no use for rational deliberation, but very well could have if some other creaturely world-type were in fact true.

Premise (14) is dubious as well. There are arguably many times in our life in which the course of action we choose matters immensely, *despite* the fact that the relevant courses of action available to us are incommensurable (Raz 1997). Hence, considering all of the reasons for choosing a particular course of action in order to figure out what to do seems to play an integral role with respect to which course of action one ends up choosing.[26]

At this point, the Molinist might give up *The No Use Argument* and argue instead that even if rational deliberation can be employed when there is nothing that an agent has most reason, all things considered, to perform, rational deliberation would only be useful in such a situation for finite beings like us rather than for God who is immediately aware of all of the reasons there are for performing each available action. But this line of reasoning seems to wrongly assume that rational deliberation is only an epistemic activity of coming to see what reasons one has for performing certain actions. As I have suggested in the previous section, however, the notion of figuring out what to do involves more than that. So, even if God—an omniscient being—is immediately aware of all of the reasons for performing certain actions, God may still undergo the process of weighing these reasons against one another as an integral role in figuring out what to do, even when she lacks most reason, all things considered, to actualize any specific world.

5.7 Conclusion

I have argued that *CDF*s are pre-volitional for God within the Molinist framework. This is because *CCF*s are pre-volitional for free creatures, and there is no relevant difference between God and free creatures. The burden is now on a Molinist who says otherwise to show that there

---

[26] Pendergraft (2014, 344) seems to accept premise (14), although perhaps a more charitable interpretation suggests he is only defending that claim that God has no use for rational deliberation *if* there is something God has most reason, all things considered, to perform. I accept this claim, but do not think it will help the Molinist given that, as explained above, premise (12) is possibly false if Molinism is true.

is in fact such a difference. I then defended an epistemic openness requirement for rational deliberation against Nelkin's criticisms, though I ultimately upheld (S*) and (S**) which are both immune from Nelkin's alleged counterexample. I then showed that, as a result of Molinism's commitment to Pre-*CDF*s, only the Molinist God fails to satisfy the epistemic openness requirement for rational deliberation according to (S**). Moreover, I argued against the claim that God has no use for rational deliberation, although I do not expect to have settled that issue here. We may thus tentatively conclude that the nature of *CDF*s may come at a significant cost to Molinism.

6

Molinists (Still) Cannot Endorse the Consequence Argument

Molinism is primarily motivated by an attempt to reconcile a robust account of divine providence with libertarianism. Libertarianism is often motivated by an alternative possibilities claim, according to which free will requires the ability to do otherwise (alternative possibilities). Thus, one of the more powerful arguments for libertarianism is the consequence argument for the conclusion that determinism is incompatible with the ability to do otherwise. Ken Perszyk (2003) has argued, however, that the Molinist cannot endorse the consequence argument. This is due to a structurally similar argument for the incompatibility of true Molinist counterfactuals of freedom (*CCF*s) and the ability to do otherwise. If Perszyk is correct, subscribing to Molinism thereby comes at a serious cost.[1] Edward Wierenga (2011) has responded to Perszyk's challenge, claiming that on the proper understanding of *CCF*s, there is a relevant difference between the consequence argument and the anti-Molinist argument. I argue that even on Wierenga's understanding of *CCF*s, there is in fact no relevant difference between the two arguments. Moreover, I strengthen Perszyk's challenge by highlighting further relevant similarities between *CCF*s and facts about the laws. I will not attempt to summarize everything that Perszyk says in defense of his challenge. Rather, I intend to flesh out the core aspects of Perszyk's challenge in a way that bears directly upon Wierenga's reply, as well as subsequent discussion thereof.

---

[1] For the purposes of this paper, I set aside Perszyk's (2000) ingenious attempt to merge Molinism with compatibilism.

6.1 Perszyk's Challenge

Turning to Peter van Inwagen's familiar, modal version of the consequence argument, consider the following one-place operator:

Np p and no one[2] has or ever had a choice about whether p.

And, consider the following two rules of inference:

Alpha $\Box p \supset Np$

Beta $(Np\ \&\ N(p \supset q)) \supset Nq$[3]

Take 'H' to refer to the complete state of the world at some time in the past, 'L' to refer to the complete list of laws of nature, and 'P' to refer to any truth. With these notions in hand, here is the modal version of the consequence argument:

1. $\Box((H\ \&\ L) \supset P)$ Determinism

2. $\Box(H \supset (L \supset P))$ 1

---

[2] 'No one' is to be read as 'no human being'. Hence, all subsequent discussion of Np should be taken to exclude the choices of God, angels, etc.

[3] Although it is widely accepted that McKay and Johnson (1996) have demonstrated that Beta as originally construed by van Inwagen (1983) is invalid, there are many other inference rules that are arguably immune for McKay and Johnson's criticism—including an inference rule advocated by McKay and Johnson (1996, 118–121). For other revisions to Beta as originally proposed by van Inwagen (1983), see Finch and Warfield (1998), van Inwagen (2000), Huemer (2000), and O'Connor (2000, ch. 1). Nothing I wish to argue for below depends upon which revised inference rule the proponent of the consequence argument wishes to accept. As a result, the reader may interpret N (and thus Beta) throughout the paper according to van Inwagen's (2000, 8) revised account of N, such that the revised version of Beta is immune from the counterexample proposed by McKay and Johnson. Moreover, van Inwagen's (2000) revised version of N does not in any way alter the dialectic below between Molinists and the consequence argument.

3.	N(H ⊃ (L ⊃ P))	2, Alpha
4.	N(H)	Fixity of the Past
5.	N(L ⊃ P)	3,4 Beta
6.	N(L)	Fixity of the Laws
7.	N(P)	5,6 Beta

Since 'P' may refer to any true proposition, including propositions about how we in fact act, (7) entails that an agent S lacks the ability to do other than what she in fact does. Thus, the consequence argument shows us that if determinism is true, then S does not have the ability to do otherwise. I follow Perszyk in abbreviating the argument as follows:

The Consequence Argument

1A. N((H & L) ⊃ P)

2A. N(H & L)

3A. N(P)

In response to the consequence argument, many have argued that either the argument equivocates between two senses abilities or the argument is simply invalid. I'll focus my attention on David Lewis' (1981) response.[4] Lewis (1981, 120) distinguishes between a weak and strong thesis regarding having a choice about whether *p* (or, as Lewis puts it, 'rendering a proposition false'), which can be formulated as follows:

---

[4] See also Gallois (1977), Lehrer (1980), and Fischer (1986). While Lewis does not explicitly discuss rule Beta in his (1981), one remark by Lewis (1981, 120) suggests that he accepts the validity of Beta even under the strong interpretation and that (2A) is false under the strong interpretation (to be discussed below). Even so, my comments below suggest what a compatibilist such as Lewis *should* say in response to the consequence argument as formulated above, given Lewis' distinction between the weak and strong thesis.

Weak Thesis An agent *S* has a **W-choice** about whether *p iff S* was able to *φ* such that, if *S φ*-ed, *p* would have been falsified (though not necessarily by *S*'s *φ*-ing or by any event that was caused by *S*'s *φ*-ing).

Strong Thesis An agent *S* has an **S-choice** about whether *p iff S* was able to *φ* such that, if *S φ*-ed, *p* would have been falsified by *S*'s *φ*-ing or by some event that was caused by *S*'s *φ*-ing.

On the one hand, if the term 'choice' referred to in the N operator is understood to be an S-choice, then compatibilists deny the validity of Beta. For, compatibilists agree that we lack an S-choice about whether the past or the laws obtain. We cannot, for instance, *break* or *violate* the laws of nature. However, we *do* have an S-choice about the truth-value of certain propositions. For instance, I could have thrown a rock at a window, such that this action would have caused an event which in turn would have falsified the true proposition '*The window remains intact*'.[5] But this is inconsistent with Beta according to the S-choice interpretation. So Beta must be invalid if the term 'choice' referred to in the N operator is understood to be an S-choice. On the other hand, if the term 'choice' referred to in the N operator is understood to be a W-choice, then compatibilists deny premise (2A). This is because, under the W-choice interpretation, compatibilists deny either N(H) or N(L). Lewis specifically denies N(L).

Turning now to Molinism, Perszyk claims that an exactly parallel argument can be formulated for the incompatibility of true *CCF*s and the ability to do otherwise. To see how the argument goes, take the following to stand for a true *CCF*: ($\alpha \ \square\rightarrow \beta$), such that '$\alpha$' refers to a

---

[5] Lewis (1981, 119).

maximally specified description of the circumstances a free creature is in, and 'β' refers to that creature's freely performing an action in those circumstances. The reason the antecedent of a *CCF* must be maximally specific is that ordinary counterfactuals don't permit strengthening the antecedent; adding more to the antecedent can change the truth-value of the counterfactual.[6] Consequently, it is only the maximally described *ultima facie* counterfactuals, rather than the non-maximally described *prima facie* counterfactuals, that are *useful* to God for deciding what to do.[7] So, an example of a *CCF* goes as follows:

      Adam If Adam were in the garden, he would freely eat the forbidden fruit.

Once again, it is crucial that "the garden" is shorthand for a maximally specified description of certain circumstances. Since the consequent says that Adam would *freely* eat the forbidden fruit, it follows that in those very circumstances Adam can refrain from eating the forbidden fruit. This is because it is an assumption of the Molinist framework that freedom requires the ability to do otherwise, or at least that the kind of freedom specified in *CCF*s entails that an agent has the ability to do otherwise.

    With this preliminary understanding of a *CCF* in place, here is Perszyk's anti-Molinist argument:[8]

      The Anti-Molinist Argument

      1B.   N$(((\alpha \ \Box\!\!\rightarrow \beta) \ \& \ \alpha) \supset \beta)$

---

[6] Lewis (1973, 31–32).
[7] Zimmerman (2009, 56–59).
[8] Perszyk (2003, 135).

2B. $N((\alpha \:\square\!\!\rightarrow \beta) \:\&\: \alpha)$

3B. $N(\beta)$

Since 'β' refers to a person's *freely* performing an action in certain circumstances, it follows that that person is able to do otherwise in the exact same circumstances. In that case, $N(\beta)$ entails a contradiction, falsifying the Molinist's assumption that there are brute, contingently true counterfactuals with *free* actions in their consequents. Now, Wierenga agrees (as do I) that (1B) is eminently plausible.[9] So, given the truth of (1B), the Molinist must claim that Beta is invalid and/or that (2B) is false. If Beta is valid, then (2B) is false. If (2B) is false and $N(\alpha)$ is true, then it is false that $N(\alpha \:\square\!\!\rightarrow \beta)$.

It is easy to see that at least sometimes we have no choice about which circumstances we are placed in, such as the circumstance in which we have our very first (alleged) free choice about what to do. Of course, it doesn't follow from the fact that person *S* has no choice about α that $N(\alpha)$ since some *other* human being may have a choice about whether to place *S* in the circumstances referred to in α. However, we may focus our attention on *CCF*s such that, clearly, *S and no other human being* has a choice about α. When we focus on such a *CCF*, we can establish that $N(\alpha)$. So, assuming that such a *CCF* is the one at issue in *The anti-Molinist Argument*, then either Beta is invalid or it is false that $N(\alpha \:\square\!\!\rightarrow \beta)$. A Molinist who endorses the consequence argument as formulated above must accept the validity of Beta. So, in order to avoid $N(\beta)$, the Molinist must deny that $N(\alpha \:\square\!\!\rightarrow \beta)$. This is exactly what Molinists have done

---

[9] Wierenga (1991, 428; 2011, 134).

in print. They have claimed that if β, then some free creature has counterfactual power over the truth of ($\alpha \Box\!\!\rightarrow \beta$). For example, consider Thomas Flint's (2011, 43) remarks on the matter:

> The truth of those counterfactuals…does not interfere with the relevant creatures' freedom to do otherwise. Even though ($C \Box\!\!\rightarrow Z$) is true and [free creature] A is in C, A still is able to do $\sim Z$. Had she done $\sim Z$, ($C \Box\!\!\rightarrow Z$) wouldn't have been true. Instead, ($C \Box\!\!\rightarrow \sim Z$) would have been true. So A has the power to do something ($\sim Z$), such that, had she done it, a counterfactual that is true—namely, ($C \Box\!\!\rightarrow Z$)—would have been false.

Flint does not say that if A had done $\sim Z$, then ($C \Box\!\!\rightarrow Z$) would have been falsified by A's $\sim Z$-ing or by some event that was caused by A's $\sim Z$-ing. Rather, Flint is only asserting the weaker claim that if A had done $\sim Z$, then ($C \Box\!\!\rightarrow Z$) would have been false. In other words, Flint is only saying that A has a W-choice about whether ($C \Box\!\!\rightarrow Z$).

Flint is not saying that A has an S-choice about whether ($C \Box\!\!\rightarrow Z$), and for good reason. Within the Molinist framework, *CCF*s are true logically prior to God's decision of which world to actualize. *A fortiori*, *CCF*s are true logically prior to any creature's free choice.[10] Consequently, no *CCF* has its truth-value even partly in virtue of a creature's free choice or even partly in virtue of some event that was caused by a creature's free choice. In other words, within the Molinist framework, N($\alpha \Box\!\!\rightarrow \beta$) is true if the term 'choice' referred to in the N operator is

---

[10] Molinists must assert this, among other reasons, in order to reply to the so-called 'too late' objection. See Adams (1977, 112–113), Kenny (1979, 70), Wierenga (1989, 148–150), Mares and Perszyk (2011, 99–101).

understood to be an S-choice, and N($\alpha \ \square\rightarrow \beta$) is false if the term 'choice' referred to in the N operator is understood to be a W-choice.[11]

So the Molinist can consistently accept the validity Beta and deny that N($\alpha \ \square\rightarrow \beta$) if the term 'choice' referred to in the N operator is understood to be a W-choice. However, the compatibilist also accepts the validity of Beta under the W-choice interpretation. The Molinist must thus affirm both N(H) and N(L) under the W-choice interpretation, contrary to the compatibilist's position. Alas, here is the central tension. Under the W-choice interpretation, the Molinist must affirm both N(H) and N(L), and yet also *deny* N($\alpha \ \square\rightarrow \beta$). It is thus incumbent upon the Molinist to provide some relevant difference between ($\alpha \ \square\rightarrow \beta$) and both H and L. Wierenga attempts to do exactly that.

6.2 Wierenga on *CCF*s and Initial Segments of Worlds

Recall that since ordinary counterfactuals do not permit strengthening the antecedent, it is only the *ultima facie* counterfactuals that are useful to God. Thus, in order for *Adam* to be the kind of *CCF* that is useful to God, "the garden" must be shorthand for a maximally specified description of certain circumstances at a time. Does that mean that *Adam* itself is true in "the garden"? No. We are stipulating that Adam has the ability to do otherwise in "the garden". But doing otherwise in "the garden" is incompatible with the truth of *Adam*. It is for this reason that Wierenga introduces the notion of an initial segment that is shared by distinct worlds:

> For any world *W* and time *t*, there exists a state of affairs, $\sum_{(W,\ t)}$, which is an initial
>
> segment of *W* terminating at *t*, and which is included in *W*. (Wierenga 2011, 127)

---

[11] Perszyk (2003, 136).

Worlds W and W' share an initial segment up until a time t if and only if $\sum_{(W, t)} = \sum_{(W', t)}$. (Wierenga 2011, 128)

A proposition p is true in an initial segment $\sum_{(W, t)}$ if and only if it is not possible that $\sum_{(W, t)}$ obtain and p be false. (Wierenga 2011, 128)

To illustrate the relationship between *CCF*s and initial segments of worlds, we can reformulate *Adam* in a way that incorporates Wierenga's terminology. I'll refer to this *CCF* as *Fruit*:

> **Fruit** If Adam were in the circumstances such that $\sum_{(W, t)}$ obtained at t, he would freely eat the forbidden fruit at t.

Adam's being in the circumstances such that $\sum_{(W, t)}$ obtains at t does not *entail* that Adam eats the forbidden fruit precisely because *Fruit* itself is not true in $\sum_{(W, t)}$ at t. Adam's being in the circumstances such that $\sum_{(W, t)}$ obtains at t is consistent with Adam's refraining from eating the forbidden fruit. Thus, the following counterfactual is possibly true:

> **No Fruit** If Adam were in the circumstances such that $\sum_{(W, t)}$ obtained at t, he would freely refrain from eating the forbidden fruit at t.

Suppose that Adam in fact eats the forbidden fruit in $\sum_{(W, t)}$ at t, and that W is actual. In that case, *Fruit* is true at t in the actual world (viz. W). Let's also stipulate that the world in which Adam does other than what he in fact does at $\sum_{(W, t)}$ at t is world W^*. In that case, *No Fruit* is true at t in W^*. W and W^* share $\sum_{(W, t)}$ as an initial segment. Moreover, while *Fruit* is true at t in W and *No*

Fruit is true at t in W^*, neither *Fruit* nor *No Fruit* is true at t in $\sum_{(W,\,t)}$. With this understanding of the relationship between *CCF*s and initial segments of worlds, we are now in a position to see how Wierenga responds to Perszyk's challenge.

Wierenga reformulates the consequence argument with the employment of an accidental necessity operator 'AN'. This notion of accidental necessity is defined partly in terms of the notion of an initial segment:

> A proposition p is *accidentally necessary* at a time t in a world W if and only if p is contingent and $\sum_{(W,\,t)}$ includes p. (Wierenga 2011, 133)

Wierenga also adopts an analogue rule Beta that employs AN:

> Beta*: $(AN(p)\ \&\ \Box(p \supset q)\ \&\ \sim \Box(q)) \supset AN(q)$. (Wierenga 2011, 133)

Here, then, is the version of the consequence argument and the version of the anti-Molinist argument that Wierenga (2011, 133–134) discusses:

> ***The Consequence Argument****
>
> 1C. $\Box((H\ \&\ L) \supset P)$
>
> 2C. $AN(H\ \&\ L)$
>
> 3C. $AN(P)$

The Anti-Molinist Argument*

1D. $\Box\,(((\alpha\,\Box\!\rightarrow\beta)\,\&\,\alpha)\supset\beta)$

2D. $AN((\alpha\,\Box\!\rightarrow\beta)\,\&\,\alpha)$

3D. $AN(\beta)$

Wierenga claims that (2D) is false because it's false that $AN(\alpha\,\Box\!\rightarrow\beta)$. Why? Since *CCF*s are not included in initial segments of worlds, it is false by definition that $AN(\alpha\,\Box\!\rightarrow\beta)$. Wierenga (2011, 134–135) says the following in support of the claim that the compatibilist cannot employ a similar strategy to reject (2C):

> If we are restricting our attention to worlds in which determinism is *true*, then we are only looking at worlds in which (H & L) is true; we are only looking at worlds that have that history and those laws in their initial segments. So constructing the Consequence Argument and the anti-Molinist argument in terms of accidental necessity allows the Molinist to accept the former while rejecting the latter.

Contrary to Wierenga's claim, I think that Wierenga's strategy for rejecting *The Anti-Molinist Argument** can in fact be adopted by a compatibilist in order to reject *The Consequence Argument**. While I will attempt to show this by focusing on Lewis (1981) who denies the fixity of the laws, I note in passing that what I say on behalf of Lewis could plausibly be said *mutatis mutandis* on behalf of a compatibilist who denies the fixity of the past.[12]

---

[12] For a denial of the fixity of the past, see, e.g., Saunders (1968) and Perry (2008).

6.3 Compatibilism and Accidental Necessity

Before turning to what compatibilists should say in regards to the notion of accidental necessity, recall that Perszyk's challenge, in a nutshell, is this. Molinists claim that we have a W-choice about whether ($\alpha \;\square\!\!\rightarrow \beta$). So (2B) is false since it is false that N($\alpha \;\square\!\!\rightarrow \beta$). Compatibilists claim either that we have a W-choice about whether H or that we have a W-choice about whether L. So (2A) is false since either it is false that N(H) or it is false that N(L). If the Molinist objects to the compatibilist's denial of (2A), it is difficult to see how the Molinist can consistently deny (2B) since there seems to be no relevant difference between the ways in which (2A) and (2B) are rejected by the Molinist and the compatibilist, respectively. Does Wierenga's discussion of accidental necessity help us see how the Molinist can deny (2B) while consistently objecting to the compatibilist's denial of (2A)? This is doubtful. For, as I now intend to show, Wierenga has not shown us how the Molinist can deny (2D) while consistently objecting to the compatibilist's denial of (2C).

Suppose Sadie φ-s at t in the actual world, W_1, and that determinism is true in W_1. A compatibilist such as Lewis holds that if Sadie freely φ-s at t, Sadie could have refrained from φ-ing at t. Suppose that W_2 is the relevant world in which Sadie refrains from φ-ing at t. Lewis holds that Sadie does not have a W-choice about the past, but that Sadie does have a W-choice about whether the actual laws of nature obtain. What, then, is Lewis likely to say about *The Consequence Argument** which employs the notion of accidental necessity? Lewis (1981, 114) says that 'If I had raised my hand, the intrinsic state of the world long ago would have been no different…[and] L would not have been true'. What I think this comment suggests is that Lewis would (and should) say that facts about which laws of nature obtain are not true in initial segments—just as Wierenga claims that *CCF*s are not true in initial segments.

To illustrate, consider again the actual, deterministic world W_1 in which Sadie φ-s at t, and the merely possible, deterministic world W_2 in which Sadie refrains from φ-ing at t. Suppose that 'L_1' refers to the complete list of deterministic laws of nature that obtain in W_1 (the actual world), and that 'L_2' refers to the complete list of deterministic laws of nature that obtain in W_2. A compatibilist such as Lewis should say that W_1 and W_2 share an initial segment, $\sum_{(W,\,t)}$, up until t. To be sure, L_1 is true at t in W_1 (the actual world), and L_2 is true at t in W_2. However, neither L_1 nor L_2 is included at t in $\sum_{(W,\,t)}$. In other words, while the past (or 'the intrinsic state of the world long ago') is accidentally necessary, facts about the laws are by definition not accidentally necessary given that they are not included in initial segments of worlds. Thus, once the consequence argument is construed in terms of accidental necessity, a compatibilist such as Lewis can employ the exact same strategy to deny (2C) that the Molinist can employ to deny (2D).

In order for Wierenga to claim that the strategy that the Molinist employs to deny (2D) cannot similarly be employed by the compatibilist to deny (2C), Wierenga needs to tell us what the relevant difference is between L and $(\alpha \;\square\!\rightarrow \beta)$, such that the former is to be included in initial segments, and the latter is not. And, it is important to see that Wierenga has not done this. In order to reject the *Anti-Molinist Argument**, Wierenga can start with the assumption that there exists true *CCF*s, which in turn rules out *CCF*s from initial segments, which in turn rules out *CCF*s from being accidentally necessary given Wierenga's definition of accidental necessity. But the compatibilist can do the exact same thing. In order to reject the *Consequence Argument**, the Lewisian compatibilist can start with the assumption that we have the ability to do otherwise in deterministic worlds, which in turn rules out facts about the laws from initial segments, which in

turn rules out facts about the laws from being accidentally necessary given Wierenga's definition of accidental necessity.

I have claimed that Wierenga has not shown how the Molinist can deny (2B)/(2D) while consistently objecting to the compatibilist's denial of (2A)/(2C). This is because Wierenga has not shown what the relevant difference is between *CCF*s and facts about the laws; introducing a notion of accidental necessity does not illuminate what the relevant difference is such that it is true that AN(L) and false that AN($\alpha \;\square\!\!\rightarrow \beta$). Similarly, with regards to Perszyk's challenge as originally formulated, we have not been shown what the relevant difference is between *CCF*s and facts about the laws such that it is true that N(L) and false that N($\alpha \;\square\!\!\rightarrow \beta$). The burden is thus squarely on the Molinist to provide some relevant difference between these two kinds of facts in order to adequately answer Perszyk's challenge. Until that is done, it is fair to say that the Molinist *still* cannot endorse the consequence argument.[13] In the final section, I intend to strengthen Perszyk's challenge by noting two important similarities between *CCF*s and facts about the laws in order to further motivate the claim that N($\alpha \;\square\!\!\rightarrow \beta$) *iff* N(L).

6.4 *CCF*s and Facts about the Laws: Strengthening the Similarity

The first similarity between *CCF*s and facts about the laws that I wish to highlight concerns how both kinds of facts play a significant role in ordering the closeness of worlds. To see that there is such a similarity, I must first summarize the anti-Molinist 'tie' argument, according to which

---

[13] Flint (2012) has recently discussed how the Molinist should respond to a number of anti-Molinist arguments that appeal to different accounts of accidental necessity. One thus might hope that Wierenga's strategy for responding to Perszyk's challenge may be vindicated by adopting some other account of accidental necessity. However, I am skeptical of such hope. Again, what the Molinist needs to do is provide some relevant difference between (L) and ($\alpha \;\square\!\!\rightarrow \beta$) such that it is true that N(L) and false that N($\alpha \;\square\!\!\rightarrow \beta$). Until that is done, the compatibilist will be able to maintain the same relation with respect to L and accidental necessity that the Molinist wishes to hold with respect to ($\alpha \;\square\!\!\rightarrow \beta$) and accidental necessity, no matter which notion of accidental necessity is adopted.

there cannot be true *CCF*s given the standard Lewis-Stalnaker semantics for counterfactuals.[14]

Here's how the argument goes. According to Lewis-Stalnaker semantics, the counterfactual (p $\Box\!\to q$) is non-vacuously true *iff q* is true in all of the closest p-worlds. Now, suppose that Gabriela is in fact never offered a bribe.[15] The Molinist holds that one of the following counterfactuals is true:[16]

Accept If Gabriela were offered a \$10,000 bribe in circumstances c, she would freely accept the bribe.

Reject If Gabriela were offered a \$10,000 bribe in circumstances c, she would freely reject the bribe.

Let's suppose that *Reject* is true. In that case, according to Lewis-Stalnaker semantics, 'Gabriela freely rejects the bribe' is true in every world within the set of closest worlds in which Gabriela is offered a \$10,000 bribe in circumstances c. But, according to a proponent of the tie argument, this is implausible. The closeness of worlds are ranked by their similarity. And, Gabriela's rejecting rather than accepting the bribe doesn't seem to have what it takes to make a world in

---

[14] See van Inwagen (1997), Gaskin (1998) and McCann (2011, 256–260). Notice that this anti-Molinist argument should not to be conflated with a similar, yet distinct argument that has come to be known as the 'might' argument. This argument also appeals to considerations about Lewis-Stalnaker semantics for counterfactuals, and is also defended in van Inwagen (1997).

[15] Making this assumption helps us to stay focused on distinctively Molinist views about counterfactuals; according to Stalnaker-Lewis semantics, the truth of p and q in the actual world entails (p $\Box\!\to q$), given unweakened (also known as strong) centering. The Molinist, however, thinks that a *CCF* can be true even if the antecedent and consequent of that *CCF* are both false. Moreover, even when we consider a *CCF* in which the antecedent and the consequent of that *CCF* are both true, the Molinist nevertheless holds that that *CCF* is true *logically prior* to the truth of the antecedent and consequent of that *CCF*. The Molinist must hold this in response to the 'too late' or 'not true soon enough' objection raised, e.g., by Adams (1977) and Kenny (1979). For a reply to this objection, see, e.g., Wierenga (1989, 148–150).

[16] Notice that "circumstance c" is shorthand for a maximally specified description of certain circumstances.

which she rejects the bribe in *c* closer to the actual world than a world in which Gabriela accepts

the bribe in *c*. And the same goes for Gabriela's accepting rather than rejecting the bribe if we

were to suppose that *Accept* is true. So, since within the set of closest worlds in which Gabriela is

offered a $10,000 bribe in circumstances *c*, some of those worlds are such that Gabriela freely

accepts the bribe and some of those worlds are such that Gabriela freely rejects the bribe (i.e.

since there is a *tie*), both *Accept* and *Reject* are false given Lewis-Stalnaker semantics for

counterfactuals.

The only response in the literature of which I am aware is Plantinga's (1974, 178), according

to which 'one feature determining the similarity of worlds is whether they share their

counterfactuals'.[17] According to this strategy, the Molinist can maintain that *Reject* is true

without abandoning Lewis-Stalnaker semantics for counterfactuals. For, if Plantinga is right,

then what makes a world in which Gabriela freely rejects the bribe in *c* closer to the actual world

than a world in which Gabriela freely accepts the bribe in *c* is that both the actual world and the

world in which Gabriela freely rejects the bribe in *c have the truth of Reject in common*.[18] Thus,

given Plantinga's strategy for rejecting the tie argument, we see how facts about the laws and

*CCF*s play a significant role in ordering the closeness of worlds.

I will now attempt to explain why I think the fact that both *CCF*s and facts about the laws

play a significant role in ordering the closeness of worlds supports the claim that N($\alpha \: \square \rightarrow \beta$) *iff*

N(L). An agent's abilities are sometimes understood in terms of an accessibility relation between

the agent and a world.[19] Accordingly, let's say that at *t* an agent *S* can φ at *t** only if there is

some world *w* in which *S* φ-s at *t**, and *w* is accessible to *S* to at *t*. Libertarians maintain that the

only worlds that are accessible to us are those that share the same laws and the same past up until

---

[17] See, e.g., Flint (1998, 135–136).
[18] See Mares and Perszyk (2011) for further discussion of the tie argument.
[19] See Lehrer (1976, 253–254) and Fischer and Todd (2011, 101–105).

the present moment. In fact, van Inwagen's (1983) second formulation of the consequence

argument—what he calls 'The Second Argument'—is formulated in terms of an agent's

accessibility to words, whereby the crucial premise is that "no person has access to any world in

which the laws of nature are different from what they are in the actual world" (92). In support of

this premise, van Inwagen merely reasserts the libertarian line that "no one can render a law of

nature false" (92). However, one way the libertarian might further motivate this crucial premise

is by claiming that the *distance* between the actual world and worlds with different laws entails

that worlds with different laws are not accessible to us.[20] Since facts about the laws play a

significant role in ordering the closeness of worlds, there is nothing we can do that would

amount to "hopping" over all of the worlds with the same laws (and the same past) to worlds

with *different* laws.[21]

If this admittedly rough sketch of some libertarian's intuitions is at all on target, the problem

for Molinists is that the very same thing can be said with respect to *CCF*s; worlds with different

true *CCF*s are not accessible to us since they are *too far*. Since *CCF*s play a significant role in

ordering the closeness of worlds, there is nothing we can do that would amount to "hopping"

over all of the worlds with the same true *CCF*s (and the same past and laws) to worlds with

different true *CCF*s.

To illustrate, suppose that Gabriela is in fact offered a bribe and she (allegedly) freely rejects

the bribe, and thus *Reject* is true. If Gabriela could have accepted the bribe, then there is a world

in which she accepts the bribe that was accessible to Gabriela. So Gabriela could have done

---

[20] To be clear, I'm not suggesting that the distance of worlds *makes it the case* that certain worlds are not accessible to us. Rather, I'm only suggesting that such distance *entails* that certain worlds are not accessible to us. After all, the distance of worlds is supposed to be arranged in such a way that tracks our intuitions about counterfactual claims given the Stalnaker-Lewis semantics for counterfactuals. Hence, the proper account of the distance of worlds is presumably partly determined by our abilities, and not the other way around.
[21] For a similar claim regarding the relevance of the distance of a world to an agent's abilities, see Spencer (2013, 157).

something that would amount to "hopping" over all of the worlds with the same true *CCF*s (and the same past and laws) to a world in which *Reject* is false and *Accept* is true. But if both *CCF*s and facts about the laws play a significant role in ordering the closeness of worlds, then why can't Gabriela *also* "hop" over all of the worlds with the same laws to worlds with different laws? After all, if the distance between worlds with different true *CCF*s doesn't preclude an agent's having access to worlds with different true *CCF*s, then we shouldn't think the distance between worlds with different laws precludes an agent's having access to worlds with different laws.

The second similarity between *CCF*s and facts about the laws that I wish to highlight holds only under the assumption that the laws are non-Humean. The similarity in question is that both kinds of facts hold or obtain independently of any act of a free creature's will. In order to illustrate this similarity, I will compare both kinds of facts with temporally relational (soft) facts. Soft facts are typically taken to be compatible with the ability to do otherwise. This compatibility, I suggest, can be explained by noting a relevant difference between soft facts and facts about non-Humean laws. This difference, however, also obtains between soft facts and *CCF*s. I now proceed to discuss soft facts and their bearing upon Perszyk's challenge.

Roughly 250 million years ago, the supercontinent Pangea existed. The proposition '*Pangea exists*' was thus true 250 million years ago at time t^*. This proposition is a temporally intrinsic (hard) fact. Suppose it was also true 250 million years ago at time t^* that '*Pangea exists 250 million years prior to Jones' sitting at time t*', such that time t is five minutes from now. This proposition is a temporally relational (soft) fact. Some libertarians do not view the existence of soft facts to be incompatible with the ability to do otherwise. However, we can construct an argument for the incompatibility of the existence of soft facts and the ability to do otherwise that

parallels the consequence argument. Take 'H' to refer to the complete state of the world at t^*, 250 million years ago. Take 'X' to refer to the proposition '*It was true 250 million years ago that Pangea exists 250 million years prior to Jones' sitting at time t*'. Lastly, take 'Y' to refer to the proposition '*Jones sits at t*'. Now, consider the following argument for the incompatibility of the existence of soft facts and the ability to do otherwise:

The Anti-Soft Facts Argument

1E. $N((H \& X) \supset Y)$

2E. $N(H \& X)$

3E. $N(Y)$

Premise (3E) entails that Jones is unable to refrain from sitting at t. Generalizing from (3E), it follows that the existence of soft facts are incompatible with the ability to do otherwise. Can the libertarian who endorses the consequence argument consistently accept the existence of soft facts? I think they can. To see this, let's first see how the libertarian who endorses *The Consequence Argument* cannot respond to *The Anti-Soft Facts Argument*.

If the laws are deterministic, then "our acts are the deductive consequences of the laws of nature and events in the remote past."[22] Similarly, given the existence of soft facts, our acts are the deductive consequences of soft facts and events in the remote past. In other words, it is indisputable that $\Box((H \& X) \supset Y)$. Thus, given that the libertarian who endorses *The Consequence Argument* accepts the validity of Alpha, such a libertarian must accept (1E). Since such a libertarian also accepts the validity of Beta, she must accept that *The Anti-Soft Facts*

---

[22] van Inwagen (1983, 16).

Argument is valid. So the only plausible way for such a libertarian to reject the soundness of *The Anti-Soft Facts Argument* is by denying (2E). Since the libertarian who endorses *The Consequence Argument* obviously cannot consistently deny N(H), she must thereby deny N(X). In other words, she must reject 'the fixity of soft facts'. Is there a relevant difference between X and L that can motivate the claim that N(L) is true but that N(X) is false? I think there is.

Patrick Todd (2013) has recently argued for a new (and, to my mind, plausible) way to cash out the soft/hard fact distinction, viz. in terms of ontological dependence. The soft/hard fact distinction has often been drawn in terms of entailment, an example of which by Todd (2013, 834) goes as follows:[23]

(ENT-S) A fact *F* at a time *t* is soft if and only if there is some 'basic' fact *F** about the future relative to *t* such that, necessarily, *F* is a fact only if *F** is a fact.

(ENT-H) A fact *F* at a time *t* is hard if and only if it is not the case that there is some 'basic' fact *F** about the future relative to *t* such that, necessarily, *F* is a fact only if *F** is a fact.

Todd (2013, 835) offers a counterexample to ENT-S in order to show that entailment is not enough to capture the soft/hard fact distinction:

If God has decreed that Jones is to sit at t, this entails that Jones sits at *t*—God's decrees infallibly come to pass. Given this fact, it follows that God has decreed that Jones will sit at *t* necessarily only if Jones sits at *t*. Moreover, that Jones sits at *t* is obviously a 'basic'

---

[23] See the collection of essays in Fischer (1989).

fact about *t*. However, it seems plainly incorrect to say that God's past *decrees* concerning the future *depend* on the future (relative to the time the decrees were made).

Since God's decrees are intuitively not soft facts, Todd (2013, 839) offers the following alternative account of a soft fact:

> **(IDT-S)** A fact *F* at a time *t* is soft if and only if *F* specifies an entity *E* as having a property *P* at *t*, and whether *E* counts as having *P* at *t* is at least in part determined by whether there exists an event or events in the future relative to *t*.

By employing the phrase 'determined by', Todd is appealing to the currently prevalent notion of ontological dependence that does not reduce to modal terms.[24] We need not get bogged down in the details of IDT-S or the contemporary debate about ontological dependence. The overall lesson is that if Todd is on the right track, then we should think that X—'*It was true 250 million years ago that Pangea exists 250 million years prior to Jones' sitting at time t*'—is at least partly determined by Jones' sitting at *t*. As a result, the libertarian can plausibly deny that N(X) given the following intuitive principle:

> ***The Choice Principle* (*Choice*)** For any proposition *p*, agent *S*, and time *t*, *S* has a W-choice at *t* about whether *p* *iff* either (i) *p* is true at least partly in virtue of *S*'s φ-ing at *t*, and *S* had the ability to refrain from φ-ing at *t*, or (ii) *p* is true at least partly in virtue of some event *e* that was caused by *S*'s φ-ing at *t*, and *S* had the ability to refrain from φ-ing at *t*, and if *S* had not φ-ed at *t*, *e* would not have occurred.

---

[24] See, e.g., Fine (1995).

The phrase 'in virtue of' is meant to track the same notion of ontological dependence that Todd is referring to when he employs the phrase 'determined by'. Now, *Choice* entails that it is false that N(X) if Jones was able to refrain from sitting at *t*, despite the fact that X. So the libertarian who adopts *Choice* has a plausible way to deny (2E). Let's now consider the implications of *Choice* for *The Consequence Argument* and *The Anti-Molinist Argument*.

Unless the compatibilist is a Humean about laws, she cannot adopt a similar strategy to deny N(L) and thus to deny (2A). This is because unless Humeanism about laws is true, L is not true even partly in virtue of any act of any person (or even partly in virtue of some event that was caused by some act of some person). So if *Choice* is true and Humeanism about laws is false, then N(L) is true. Now, I am admitting that if Humeanism about laws is true, then *The Consequence Argument* is probably not sound. But I think that Beebee and Mele (2002) have effectively demonstrated this anyway. Moreover, I think that what libertarians had (or, at any rate, *should* have had) in mind all along is that *non*-Humean laws are incompatible with the ability to do otherwise. For, consider what van Inwagen (2004, 349) says in defense of the premise (premise (6)) that "J was not able to render L false":

> The ability that premise (6)…says that J does not have is the ability to perform a miracle. And since it's entirely plausible to suppose that ordinary people in ordinary circumstances are not able to perform miracles, it's entirely plausible to suppose that (6) is true.

That we are unable in ordinary circumstances to perform miracles should, if anything, count against the existence of Humean laws, not the compatibility of Humean laws and the ability to do otherwise. For, there seems to be no relevant difference between Humean laws and soft facts. But surely the defender of *The Consequence Argument* does not wish to claim that her argument is sound *iff The Anti-Soft Facts Argument* is sound. Rather, it is better for the libertarian to instead concede that Humean laws are compatible with the ability to do otherwise.[25]

Let's simply grant for the sake of argument that the libertarian who endorses *The Consequence Argument* admits that the argument is not sound if Humeanism about laws is true. We thereby have a reason to think that soft facts are compatible with the ability to do otherwise and that (non-Humean) determinism is incompatible with the ability to do otherwise because only soft facts are true at least partly in virtue of what we freely do. This result only places the Molinist in a more difficult position with respect to Perszyk's challenge.

As we saw earlier, *CCF*s are true logically prior to God's decision of which world to actualize. *A fortiori*, *CCF*s are true logically prior to any creature's free choice. Consequently, just like facts about non-Humean laws, no *CCF* has its truth-value even partly in virtue of anyone's (any *creature's*) free choice or even partly in virtue of some event that was caused by anyone's free choice. Hence, while *Choice* renders the intuitively correct verdict that N(X) is false and N(L) is true (if Humeanism is false), *Choice* also renders the verdict that N($\alpha \ \square\!\rightarrow \beta$) is true. In other words, while *Choice* undermines *The Anti-Soft Facts Argument*, *Choice* gives us good reason to think that *The Consequence Argument* and *The Anti-Molinist Argument* are both sound.

---

[25] See Beebee and Mele's (2002, 207–208) discussion of the similarity between soft facts (or facts about the future) and facts about Humean laws.

6.5 Conclusion

I have argued that Wierenga has not succeeded in defusing Perszyk's challenge by appealing to accidental necessity. Whatever the Molinist wishes to say regarding accidental necessity and true *CCF*s, the compatibilist can maintain the same position regarding accidental necessity and, e.g. facts about the laws. In order for the Molinist to consistently endorse the consequence argument, she must say what the relevant difference is between true *CCF*s and facts about the laws (and facts about the past). And, this has not yet been done. Moreover, I have attempted to strengthen Perszyk's challenge by noting two salient similarities between *CCF*s and facts about the laws. According to the first similarity, *CCF*s and facts about the laws both play a significant role in ordering the closeness of worlds. According to the second similarity, *CCF*s and facts about the (non-Humean) laws both hold or obtain independently of any act of a free creature's will. I want to emphasize, however, that even if my attempts to strengthen Perszyk's challenge fail, the burden is still on the Molinist to highlight some relevant difference between the consequence argument and the anti-Molinist argument. Until this is done, it is fair to say that the Molinist still cannot endorse the consequence argument.

Bibliography

Adams, R. M. (1977). Middle Knowledge and the Problem of Evil. *American Philosophical Quarterly* 14: 109–117.

Adams, R. M. (1985). Plantinga on the Problem of Evil. In J. E. Tomberline & P. van Inwagen (eds.), *Alvin Plantinga*, 225–55. Dordrecht: D. Reidel.

Arntzenius, F. & Maudlin, T. (2002). Time Travel and Modern Physics. In C. Callender, (ed.) *Time, Reality and Experience*, 169–200. Cambridge: Cambridge University Press.

Ayer, A. J. (1954). Freedom and Necessity. In A.J. Ayer, *Philosophical Essays*, 271–284. London: Macmillan.

Baker, L. R. (2006). Moral Responsibility without Libertarianism. *Noûs* 40: 307–330.

Balaguer, Mark. 2009. *Free Will as an Open Scientific Problem*. Cambridge, MA: MIT Press.

Basinger, D. (1986). Omniscience and Deliberation: a Reply to Reichenbach. *International Journal for Philosophy of Religion* 20: 169–172.

Beebee, H. & Mele, A. (2002). Humean Compatibilism. *Mind* 111: 201–233.

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.

Bok, H. (1998). *Freedom and Responsibility*. Princeton: Princeton University Press.

Cain, J. (2014). A Frankfurt Example to End All Frankfurt Examples. *Philosophia* 42: 83–93.

Campbell, J. (1997). A Compatibilist Theory of Alternative Possibilities. *Philosophical Studies* 88: 319–330.

Clarke, R. (1992). Deliberation and Beliefs About One's Abilities. *Pacific Philosophical Quarterly* 73: 101–13.

Clarke, R. (2003). *Libertarian Theories of Free Will*. New York: Oxford University Press.

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.

Coffman, E. J., & Warfield, T. A. (2005). Deliberation and Metaphysical Freedom. *Midwest*

Studies in Philosophy 29: 25–44.

Cova, F. (2014). Frankfurt-Style Cases User Manual: Why Frankfurt-Style Enabling Cases Do Not Necessitate Tech Support. *Ethical Theory and Moral Practice* 17: 505–521.

Cohen, Y. (2015). Molinists (Still) Cannot Endorse the Consequence Argument. *International Journal for Philosophy of Religion* 77: 231–246.

Correia, F. & Schneider, B. (2012). Grounding: an Opinionated Introduction. In F. Correia & D. Schneider (eds.), *Metaphysical Grounding: Understanding the Structure of Reality*, 1–36. Cambridge: Cambridge University Press.

Cova, F. (2014). Frankfurt-Style Cases User Manual: Why Frankfurt-Style Enabling Cases Do Not Necessitate Tech Support. *Ethical Theory and Moral Practice* 17: 505–521.

Cowan, S. B. (2003). The Grounding Objection to Middle Knowledge Revisited. *Religious Studies* 39: 93–102.

Craig, W.L. (1987). Divine Foreknowledge and Newcomb's Paradox. *Philosophia* 17: 331–350.

Craig, W. L. (1988). *The Problem of Divine Foreknowledge and Future Contingents from Aristotle to Suarez*. Leiden: E.J. Brill.

Craig, W.L. (1991). *Divine Foreknowledge and Human Freedom*: *The Coherence of Theism: Omniscience*. Leiden: E. J. Brill.

Craig, W.L. (2001). Middle Knowledge, Truth–Makers, and the 'Grounding Objection'. *Faith and Philosophy* 18: 337–352.

Craig, W.L. (2007). Naturalism and Intelligent Design. In W. A. Dembski & M. Ruse (eds.), *Intelligent Design*, 58–71. Minneapolis: Fortress Press.

Cushman, F. (2008). Crime and Punishment: Distinguishing the Roles of Causal and Intentional Analyses in Moral Judgment. *Cognition*: 108: 353–380.

Davidson, D. (1963). Actions, Reasons, and Causes. *Journal of Philosophy* 60: 685–700.

Della Rocca, M. 1998. Frankfurt, Fischer and Flickers. *Noûs* 32: 99–105.

Demetriou, K. (2010). The Soft-Line Solution to Pereboom's Four-Case Argument. *Australasian Journal of Philosophy* 88: 595–617.

Dowe, P. (2000). The Case for Time Travel. *Philosophy* 3: 441–51.

Ekstrom, L. (2000). *Free Will*. Boulder, CO: Westview.

Fara, M. (2005). Dispositions and Habituals. *Noûs* 39: 43–82.

Fara, M. (2008). Masked Abilities and Compatibilism. *Mind* 117: 843–65.

Finch, A. & Rea, M. (2008). Presentism and Ockham's Way Out. In J. Kvanvig (ed.), *Oxford Studies in Philosophy of Religion vol. 1*, 1–17. New York: Oxford University Press.

Fine, Kit. (1995). Ontological Dependence. *Proceedings of the Aristotelian Society* 95: 269–290.

Fine, Kit. (2012). Guide to Ground. In F. Correia and D. Schneider (Eds.), *Metaphysical Grounding: Understanding the Structure of Reality*, 37–80. Cambridge: Cambridge University Press.

Fischer, J. M. (1986). Van Inwagen on Free Will. *Philosophical Quarterly* 36: 252–260.

Fischer, J. M. (ed.) (1989). *God, Foreknowledge, and Freedom*. Stanford: Stanford University Press.

Fischer, J.M. (1994). *The Metaphysics of Free Will: An Essay On Control*. Oxford: Blackwell Publishers.

Fischer, J.M. (2003). Responsibility and Alternative Possibilities. In D. Widerker & M. McKenna (eds.), *Moral Responsibility and Alternative Possibilities*, 27–52. Burlington, VT: Ashgate.

Fischer, J. M. (2004). Responsibility and Manipulation. *The Journal of Ethics* 8: 145–177.

Fischer, J.M. 2006. *My Way: Essays on Moral Responsibility*. New York: Oxford University Press.

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: 315–36.

Fischer, J. M. (2013). The Deterministic Horn of the Dilemma Defense: A Reply to Widerker and Goetz. *Analysis* 73: 489–496.

Fischer, J. M. & Ravizza, M. (1998). *Responsibility and Control: a Theory of Moral Responsibility*. Cambridge: Cambridge University Press.

Fischer, J.M. & Todd, P. (2011). The Truth About Freedom: A Reply to Merricks. *Philosophical*

Review 120: 97–115.

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.

Fischer, J. M. & Tognazzini, N. A. (2011). The Physiognomy of Responsibility. *Philosophy and Phenomenological Research* 82: 381–417.

Finch, A. & Warfield, T. (1998). The Mind Argument and Libertarianism. *Mind*, 107: 515–528.

Flint, T. P. (1998). *Divine Providence: The Molinist Account*. Ithaca: Cornell University Press.

Flint, T. P. (1999). A New Anti-Anti-Molinist Argument. *Religious Studies* 35: 299–305.

Flint, T. P. (2003). The Multiple Muddles of Maverick Molinism. *Faith and Philosophy* 20: 91–100.

Flint, T. P. (2009). Divine Providence. In T. P. Flint & M. C. Rea (eds.) *The Oxford Handbook of Philosophical Theology*, 262–285. Oxford: Oxford University Press.

Flint, T. P. (2011). Whence and Whither the Molinist Debate: A Reply to Hasker. In K. Perszyk (ed.), *Molinism: The Contemporary Debate*, 37–49. Oxford: Oxford University Press.

Flint, T. P. (2012). The Varieties of Accidental Necessity. In K.J Clark & M.C. Rea (eds.) *Reason, Metaphysics, and Mind: New Essays on the Philosophy of Alvin Plantinga*, 38–54. Oxford: Oxford University Press.

Forrest, P. (1985). Backward Causation in Defence of Free Will. *Mind* 94: 210–7.

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. (2011a). Masks, Abilities, and Opportunities: Why the New Dispositionalism Cannot Succeed. *The Modern Schoolman* 88: 89–103.

Franklin, C.E. (2011b). Neo-Frankfurtians and Buffer Cases: the New Challenge to the Principle of Alternative Possibilities. *Philosophical Studies* 152: 189–207.

Franklin, C.E. (2011c). The Problem of Enhanced Control. *Australasian Journal of Philosophy* 89: 687–706.

Freddoso, A. (1988). Introduction. In *Concordia. On Divine Foreknowledge* (*Part IV of the Concordia*). A. Freddoso (trs.), 1–81. Ithaca: Cornell University Press.

Funkhouser, E. (2009). Frankfurt Cases and Overdetermination. *Canadian Journal of Philosophy* 39: 341–370.

Gallois, A. (1977). Van Inwagen on Free Will and Determinism. *Philosophical Studies* 32: 99–105.

Gaskin, R. (1993). Conditionals of Freedom and Middle Knowledge. *Philosophical Quarterly* 43: 412–430.

Gaskin, R. (1998). Middle Knowledge, Fatalism and Comparative Similarity of Worlds. *Religious Studies* 34: 189–203.

Ginet, C. (1962). Can the Will Be Caused? *The Philosophical Review* 71: 49–55.

Ginet, C. (1966). Might We Have no Choice? In K. Lehrer (ed.), *Freedom and Determinism*, 87–104. New York: Random House.

Ginet, C. (1990). *On Action*. Cambridge: Cambridge University Press.

Ginet, C. 1996. In Defence of the Principle of Alternative Possibilities: Why I Don't Find Frankfurt's Argument Convincing. *Philosophical Perspectives* 10: 403–417.

Ginet, C. (2007). An Action Can be Both Uncaused and Up to the Agent. In C. Lumer & S. Nannini (eds.), *Intentionality, Deliberation, and Autonomy*, 243–256. Farnhnam, UK: Ashgate.

Goetz, S. 2005. Frankfurt-Style Counterexamples and Begging the Question. *Midwest Studies in Philosophy* 29: 83–105.

Goetz, S. (2008). *Freedom, Teleology, and Evil*. London: Continuum.

Griffith, M. (2010). Why Agent-Caused Actions Are Not Lucky. *American Philosophical Quarterly* 47: 43–56.

Haji, I. (1998). *Moral Appraisability: Puzzles, Proposals, and Perplexities*. New York: Oxford University Press.

Haji, I. 2009. *Incompatibilism's Allure: Principal Arguments for Incompatibilism*. Peterborough ON: Broadview Press.

Haji, I. (2012). *Reason's Debt to Freedom: Normative Appraisals, Reasons, and Free Will*. Oxford: Oxford University Press.

Haji, I. (2013). Historicism, Non-historicism, Or A Mix? *Journal of Ethics* 17: 185–204.

Haji, I. & Cuypers, S. (2007). Magical Agents, Global Induction, and the Internalism/ Externalism Debate. *Australasian Journal of Philosophy* 85: 343–371.

Haji, I. & McKenna, M. (2011). Disenabling Levy's Frankfurt-Style Enabling Cases. *Pacific Philosophical Quarterly* 92: 400–414.

Hampshire, S. & Hart, H.L.A. (1958). Decision, Intention and Certainty. *Mind* 67: 1–12.

Hartman, R. J. (Manuscript). Counterfactuals of Freedom and the Luck Objection to Libertarianism.

Hasker, W. (1986). A Refutation of Middle Knowledge. *Nous* 20: 545–557.

Hasker, W. (1989). *God, Time, and Knowledge*. Ithaca, NY: Cornell University Press.

Hasker, W. (1995). Middle Knowledge: A Refutation Revisited. *Faith and Philosophy* 12: 223–236.

Hasker, W. (2003). Counterfactuals and Evil: A Final Reply to R. Douglas Geivett. *Philosophia Christi* 5: 235–249.

Hasker, W. (2004). *Providence, Evil, and the Openness of God*. London: Routledge.

Hasker, W. (2011). The (Non-)Existence of Molinist Counterfactuals. In K. Perszyk (ed.) *Molinism: The Contemporary Debate*, 25–36. Oxford: Oxford University Press.

Henden, E. (2010). Deliberation Incompatibilism. *Dialectica* 64: 313–333.

Horgan, T. (1979). 'Could', Possible Worlds, and Moral Responsibility. *The Southern Journal of Philosophy* 17: 345–358.

Huemer, M. (2000). Van Inwagen's Consequence Argument. *Philosophical Review* 109: 525–544.

Hunt, D. P. (1992). Omniprescient Agency. *Religious Studies* 28: 351–369.

Hunt, D. P. (1993). Divine Providence and Simple Foreknowledge. *Faith and Philosophy* 10: 394–414.

Hunt, D. P. (1996). The Compatibility of Omniscience and Intentional Action: A Reply to Tomis Kapitan. *Religious Studies* 32: 49–60.

Hunt, D. P. (2000). Moral Responsibility and Unavoidable Action. *Philosophical Studies* 97: 195–227.

Hunt, D. P. (2003). Freedom, Foreknowledge and Frankfurt. In D. Widerker & M. McKenna (eds.), *Moral Responsibility and Alternative Possibilities*, 159–183. Burlington, VT: Ashgate.

Hunt, D. P. & S. Shabo. (2013). Frankfurt Cases and the (In)significance of Timing: A Defense of the Buffering Strategy. *Philosophical Studies* 164: 599–622.

Kane, R. (1985). *Free Will and Values*. Albany, NY: State University of New York Press.

Kane, R. (1996). *The Significance of Free Will*. New York: Oxford University Press.

Kapitan, T. (1986). Deliberation and the Presumption of Open Alternatives. *Southern Journal of Philosophy* 40: 230–251.

Kapitan, T. (1991). Agency and Omniscience. *Religious Studies* 27: 105–120.

Kapitan, T. (1994). The Incompatibility of Omniscience and Intentional Action: a reply to David P. Hunt. *Religious Studies* 30: 55–65.

Kapitan, T. (2011). A Compatibilist Reply to the Consequence Argument. In R. Kane (ed.), *The Oxford Handbook of Free Will 2nd ed*, 131–50. New York: Oxford University Press.

Kaufman, A. S. (1966). Practical Decision. *Mind* 75: 25–44.

Kaufman, A. S. (1991). Agency and Omniscience. *Religious Studies* 27: 150–120.

Kenny, A. (1979). *The God of the Philosophers*. Oxford: Oxford University Press.

Kittle, S. (2014). Vihvelin and Fischer on 'Pre-decisional' Intervention. *Philosophia* 42: 987–997.

Korsgaard, C. M. (1996). *Creating the Kingdom of Ends*. (Cambridge: Cambridge University Press).

Kvanvig, J. (2002). On Behalf of Maverick Molinism. *Faith and Philosophy* 19: 348–357.

Kvanvig, J. (2011). *Destiny and Deliberation*. Oxford: Oxford University Press.

Lehrer, K. (1968). Cans Without Ifs. *Analysis* 29: 29–32.

Lehrer, K. (1976). 'Can' in Theory and Practice: a Possible World Analysis. In M. Brand & D. Walton (eds.), *Action Theory*, 241–270. Dordrecht: D. Reidel.

Lehrer, K. (1980). Preferences, Conditionals and Freedom. In P. van Inwagen (ed.), *Time and Cause: Essays Presented to Richard Taylor*, 187–201. Dordrecht: D. Reidel.

Leon, F. & Tognazzini, N.A. (2010). Why Frankfurt-Examples Don't Need to Succeed to Succeed. *Philosophy and Phenomenological Research* 80: 551–565.

Levy, N. (2006). Determinist Deliberations. *Dialectica* 60: 453–459.

Levy, N. (2008). Counterfactual Intervention and Agents' Capacities. *The Journal of Philosophy* 105: 223–239.

Levy, N. (2012). Capacities and Counterfactuals: A Reply to Haji and McKenna. *Dialectica* 66: 607–620.

Levy, N. (2013). Manipulations for Manipulation Cases. Posted on Flickers of Freedom, June 12, 2013. Accessed on June 11, 2014 from: http://philosophycommons.typepad.com/flickers_of_freedom/2013/06/manipulations-for-manipulation-cases.html.

Levy, N. (2014a). Countering Cova: Frankfurt-Style Cases Are Still Broken. *Ethical Theory and Moral Practice* 17: 523–527.

Lewis, D.K. (1973a). Causation. *Journal of Philosophy* 70: 556–567.

Lewis, D.K. (1973b). *Counterfactuals*. Cambridge, MA: Harvard University Press.

Lewis, D. K. (1976). The Paradoxes of Time Travel. *American Philosophical Quarterly* 13: 145–52.

Lewis, D. K. (1981). Are We Free to Break the Laws? *Theoria* 47: 113–121.

Lewis, D. (1997). Finkish Dispositions. *Philosophical Quarterly* 47: 143–158.

List, C. (2014). Free Will, Determinism, and the Possibility of Doing Otherwise. *Noûs* 48: 156–178.

Lycan, W. G. (1997). *Consciousness*. Cambridge, MA: MIT Press.

Mackie, J.L. (1982). *The Miracle of Theism: Arguments For and Against the Existence of God*. Oxford: Clarendon Press.

Manley, D., & Wasserman, R. (2008). On Linking Dispositions and Conditionals. *Mind*, 117: 59–84.

Mares, E. & Perszyk, K. (2011). Molinist Conditionals. In K. Perszyk (ed.) *Molinism: The Contemporary Debate*, 97–117. Oxford: Oxford University Press.

McCann, H. (1998). *The Works of Agency*. Ithaca: Cornell University Press.

McKay, T. & Johnson, D. (1996). A Reconsideration of An Argument Against Compatibilism. *Philosophical Topics* 24: 113–122.

McKenna, M. (2003). Robustness, Control, and the Demand for Morally Significant Alternatives: Frankfurt Examples with Oodles and Oodles of Alternatives. In D. Widerker & M. McKenna (eds.), *Moral Responsibility and Alternative Possibilities*, 201–217. Burlington, VT: Ashgate.

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. *The Journal of Ethics* 16: 145–174.

Mele, A. R. (1992). *Springs of action: understanding intentional behavior*. New York: Oxford University Press.

Mele, A.R. (1995). *Autonomous Agents*. New York: Oxford University Press.

Mele, A.R. (2003). *Motivation and Agency*. 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.

Mele, A. R. & Robb, D. (1998). Rescuing Frankfurt-style Cases. *The Philosophical Review* 107: 97–112.

Mele, A. R., & Robb, D. (2003). Bbs, Magnets and Seesaws: the Metaphysics of Frankfurt-style Cases. In D. Widerker & M. McKenna (eds.), *Moral Responsibility and Alternative Possibilities*, 127–138. Burlington, VT: Ashgate.

Merricks, T. (2007). *Truth and Ontology*. Oxford: Oxford University Press.

Merricks, T. (2009). Truth and Freedom. *The Philosophical Review* 118: 29–57.

Molina, L. (1988). *Concordia* in *On Divine Foreknowledge* (*Part IV of the Concordia*). A. Freddoso (trs.). Ithaca: Cornell University Press.

Moore, G.E. (1912). *Ethics*. London: Williams and Norgate.

Murray, D., & Lombrozo, T. (Manuscript). Effects of Manipulation on Attributions of Causation, Free Will, and Moral Responsibility.

Nelkin, D. K. (2004). Deliberative Alternatives. *Philosophical Topics* 32: 215–240.

Nelkin, D. K. (2011). *Making Sense of Freedom and Responsibility*. New York: Oxford University Press.

O'Connor, T. (1992). The Impossibility of Middle Knowledge. *Philosophical Studies* 66: 139–166.

O'Connor, T. (2000). *Persons and Causes*. New York: Oxford University Press.

O'Connor, T. (2009). Agent-Causal Power. In T. Handfield (ed.), *Dispositions and Causes*, 189–214. New York: Oxford University Press.

Otte, R. (2009). Transworld Depravity and Unobtainable Worlds. *Philosophy and Phenomenological Research* 78: 165–177.

Palmer, D. (2011). Pereboom on the Frankfurt Cases. *Philosophical Studies* 153: 261–272.

Palmer, D. (2013). The Timing Objection to Frankfurt Cases. *Erkenntnis* 78: 1011–1023.

Palmer, D. (2014). Deterministic Frankfurt cases. *Synthese* 191: 3847–3864.

Pendergraft, G. (2014). Against Deliberation Restrictions. *Religious Studies* 50: 341–357.

Pereboom, D. (2001). *Living without Free Will*. Cambridge: Cambridge University Press.

Pereboom, D. (2003). Source Incompatibilism and Alternative Possibilities. In D. Widerker & M. McKenna (eds.), *Moral Responsibility and Alternative Possibilities*, 185–199. Burlington, VT: Ashgate.

Pereboom, D. (2008). A Compatibilist Account of the Epistemic Conditions on Rational Deliberation. *Journal of Ethics* 12: 287–306.

Pereboom, D. (2011a). Free-Will Skepticism and Meaning in Life. In R. Kane (ed.), *The Oxford Handbook of Free Will 2nd ed.*, 407–424. New York: Oxford University Press.

Pereboom, D. (2011b). Theological Determinism and Divine Providence. In K. Perszyk (ed.) *Molinism: The Contemporary Debate*, 262–280. Oxford: Oxford University Press.

Pereboom, D. (2014). *Free Will, Agency, and Meaning in Life*. New York: Oxford University Press.

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.

Perszyk, K. (2000). Molinism and Compatibilism. *International Journal for Philosophy of Religion* 48: 11–33.

Perszyk, K. (2003). Molinism and the Consequence Argument: A Challenge. *Faith and Philosophy* 20: 131–151.

Perszyk, K. (2011). Introduction. In K. Perszyk (ed.) *Molinism: The Contemporary Debate*, 1–24. Oxford: Oxford University Press.

Perszyk, K. (2013). Recent Work on Molinism. *Philosophy Compass* 8: 755–770.

Peterson, M., Hasker, W., Reichenbach, B. R. & Basinger, D. (2009). *Reason and Religious Belief: an Introduction to the Philosophy of Religion*. Oxford: Oxford University Press.

Pettit, P. (1989). Determinism with Deliberation. *Analysis* 49: 42–44.

Phillips, J., & Shaw, A. (2014). Manipulating Morality: Third-party Intentions Alter Moral Judgments by Changing Causal Reasoning. *Cognitive Science* 38: 1–28.

Plantinga, A. (1979). *The Nature of Necessity*. Oxford: Clarendon Press.

Plantinga, A. (1985). Replies to My Colleagues. In J. E. Tomberlin & P. van Inwagen (eds.) *Alvin Plantinga*, 313–396. Dordrecht: D. Reidel.

Plantinga, A. (1986). On Ockham's Way Out. *Faith and Philosophy* 3: 235–269.

Plantinga, A. (2009). Transworld Depravity, Transworld Sanctity, and Unobtainable Essences. *Philosophy and Phenomenological Research* 78: 178–191.

Raven, M. J. (2012). In Defence of Ground. *Australasian Journal of Philosophy* 90: 687–701.

Raz, J. (1997). Incommensurability and Agency. In R. Chang (ed.), *Incommensurability, Incomparability, and Practical Reason*, 110–28. Cambridge, MA: Harvard University Press.

Robinson, M. (2012). Modified Frankfurt-type Counterexamples and Flickers of Freedom.

Philosophical Studies 157: 177–194.

Rosen, G. (2010). Metaphysical Dependence: Grounding and Reduction. In B. Hale & A. Hoffmann (eds.), *Modality: Metaphysics, Logic, and Epistemology*, 109–136. New York: Oxford University Press.

Reichenbach, B. R. (1984). Omniscience and Deliberation. *International Journal for Philosophy of Religion* 16: 225–236.

Rogers, K. (2007). God is Not the Author of Sin: an Anselmian Response to Hugh McCann. *Faith and Philosophy* 24: 300–309.

Sartorio, C. (2011). Actuality and Responsibility. *Mind* 120: 1071–1097.

Sartorio, C. (2013). Making a Difference in a Deterministic World. *The Philosophical Review* 122: 189–214.

Sartorio, C. (Forthcoming). Vihvelin on Frankfurt-style Cases and the Actual-Sequence View. *Criminal Law and Philosophy*, 1–14.

Saunders, J. T. (1968). The Temptation of Powerlessness. *American Philosophical Quarterly* 5: 100–108.

Schaffer, J. (2009). On What Grounds What. In D. Chalmers, D. Manley, & R. Wasserman (eds.), *Metametaphysics: New Essays on the Foundations of Ontology*, 347–383. New York: Oxford University Press.

Searle, J. R. (2001). *Rationality in Action*. Cambridge, MA: MIT Press.

Senor, T. D. (2008). Defending Divine Freedom. In J. Kvanvig (ed.), *Oxford Studies in Philosophy of Religion vol. 1*, 168–195. Oxford: Oxford University Press.

Shabo, S. (2010). Uncompromising source incompatibilism. *Philosophy and Phenomenological Research* 80: 349–383.

Shabo, S. (2014). It Wasn't Up to Jones: Unavoidable Actions and Intensional Contexts in Frankfurt Examples. *Philosophical Studies* 169: 379–399.

Smith, M. (2003). Rational Capacities, Or: How to Distinguish Recklessness, Weakness, and Compulsion. In S. Stroud S. & C. Tappolet (eds.), *Weakness of Will and Practical Irrationality*, 17–38. New York: Oxford University Press.

Spencer, J. (2013). What Time Travelers Cannot Not Do (but Are Responsible for Anyway). *Philosophical Studies* 166: 149–162.

Stump, E. & Kretzmann, N. (1981). Eternity. *Journal of Philosophy* 78: 429–458.

Strawson, G. (1994). The Impossibility of Moral Responsibility. *Philosophical Studies* 75: 5–24.

Stump, E. (1996). Libertarian Freedom and the Principle of Alternate Possibilities. In D. Howard-Snyder & J. Jordan (eds.), *Faith, Freedom and Rationality*, 73–88. Totowa, NJ: Rowman and Littlefield.

Stump, E. (2003). Moral Responsibility Without Alternative Possibilities. In D. Widerker & M. McKenna (eds.), *Moral Responsibility and Alternative Possibilities*, 139–158. Burlington, VT: Ashgate.

Suarez, F. (1741). *De Scientia Quam Deus Habet de Futuris Contingentibus*. In *Opera Omnia*, vol. 10. Venice edition.

Swinburne, R. (1993). *The Coherence of Theism*. Oxford: Clarendon Press.

Tännsjö, T. (1989). Soft Determinism and How We Became Responsible for the Past. *Philosophical Papers* 18: 189–201.

Taylor, R. (1964). Deliberation and Foreknowledge. *American Philosophical Quarterly* 1: 73–80.

Taylor, R. (1966). *Action and Purpose*. Englewood Cliffs, NJ: Prentice-Hall, Inc.

Taylor, R. (1991). *Metaphysics*, *4ᵗʰ ed.* Englewood Cliffs, N.J.: Prentice-Hall.

Todd, P. (2013). Soft Facts and Ontological Dependence. *Philosophical Studies* 164: 829–844.

Todd, P. (2014). Against Limited Foreknowledge. *Philosophia* 42: 523–538.

van Inwagen, P. (1983). *An Essay on Free Will*. Oxford: Oxford University Press.

Van Inwagen, P. (1997). Against Middle Knowledge. In P. A. French, T. W. Uehling Jr., & H. K. Wettstein (eds.) *Midwest Studies in Philosophy*, 21, 225–236. Notre Dame: University of Notre Dame Press.

Van Inwagen, P. (2000). Free Will Remains a Mystery. In J.E. Tomberlin (ed.) *Philosophical Perspectives* 14, 1–19). Oxford: Blackwell.

Van Inwagen, P. (2004). Freedom to Break the Laws. *Midwest Studies in Philosophy* 28: 334–350.

Vihvelin, K. (2000). Freedom, Foreknowledge, and the Principle of Alternative Possibilities. *Canadian Journal of Philosophy* 30: 1–23.

Vihvelin, K. (2004). Free Will Demystified: a Dispositional Account. *Philosophical Topics* 32: 427–450.

Vihvelin, K. (2008). Foreknowledge, Frankfurt, and the Ability to do Otherwise: A Reply to Fischer. *Canadian Journal of Philosophy* 38: 343–372.

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.

Wallace, R. J. (1999). Three Conceptions of Rational Agency. *Ethical Theory and Moral Practice* 2: 217–242.

Waller, B. N. (1985). Deliberating About the Inevitable. *Analysis* 45: 48–52.

Watson, G. (1975). Free Agency. *The Journal of Philosophy* 83: 517–522.

Whittle, A. (2010). Dispositional Abilities. *Philosophers' Imprint* 10: 1–23.

Widerker, D. (1995a). Libertarian Freedom and the Avoidability of Decisions. *Faith and Philosophy* 12: 113–118.

Widerker, D. (1995b). Libertarianism and Frankfurt's Attack on the Principle of Alternative Possibilities. *Philosophical Review* 104: 247–261.

Widerker, D. & Goetz, S. (2013). Fischer Against the Dilemma Defence: the Defence Prevails. *Analysis* 73: 283–295.

Wierenga, E. (1989). *The Nature of God: An Inquiry Into the Divine Attributes*. Ithaca: Cornell University Press.

Wierenga, E. (1991). Prophecy, Freedom, and the Necessity of the Past. *Philosophical Perspectives* 5: 425–445.

Wierenga, E. (2001). Providence, Middle Knowledge, and the Grounding Objection. *Philosophia Christi* 3: 447–457.

Wierenga, E. (2011). Tilting at Molinism. In K. Perszyk (ed.), *Molinism: The Contemporary Debate*, 118–139. Oxford: Oxford University Press.

Wyma, K. (1997). Moral Responsibility and Leeway for Action. *American Philosophical Quarterly* 34: 57–70.

Williamson, T. (2002). Necessary existents. In A. O'Hear (Ed.), *Logic, Thought, and Language*, 233–251. Cambridge: Cambridge University Press.

Wolf, S. (1990). *Freedom Within Reason*. New York: Oxford University Press.

Zimmerman, D. (2009). Yet Another Anti-Molinist Argument. In S. Newlands & L. M. Jorgensen (Eds.), *Metaphysics and the Good: Themes from the Philosophy of Robert Adams*, 33–94. Oxford: Oxford University Press.

Zimmerman, D. (2012). The Providential Usefulness of 'Simple Foreknowledge'. In K.J. Clark & M. Rea (Eds.), *Reason, Metaphysics, and Mind: New Essays on the Philosophy of Alvin Plantinga*, 174–202. Oxford: Oxford University Press.

Yishai A. Cohen

CONTACT

Department of Philosophy
Syracuse University
541 Hall of Languages
Syracuse, NY 13244-1170

yishai.a.cohen@gmail.com
yishaicohen.com
214.315.5687

EDUCATION

Syracuse University
　　Ph.D. Candidate, Philosophy (Expected date: May 2016)
　　Dissertation Title: *Incompatibilist Alternative Possibilities*
Southern Methodist University
　　B.A., Philosophy, Religious Studies (2010)

RESEARCH

Areas of Specialization
　　Free Will & Moral Responsibility, Metaphysics, Philosophy of Religion

Areas of Competence
　　Ethics

ARTICLES

1. Forthcoming. "Leeway Compatibilism and Frankfurt-Style Cases" *Thought*.
2. Forthcoming. "Moral Obligations: Actualist, Possibilist, or Hybridist?" (with Travis Timmerman) *Australasian Journal of Philosophy*.
3. Forthcoming. "Counterfactuals of Divine Freedom" *International Journal for Philosophy of Religion*.
4. 2015. "Endless Future: A Persistent Thorn in the *Kalām* Cosmological Argument" *Philosophical Papers* 44 (2): 165–187.
5. 2015. "The Manipulation, at the Very Least, Undermines Classical Compatibilism" *Philosophia* 43 (2): 291–307.
6. 2015. "Molinists (Still) Cannot Endorse the Consequence Argument" *International Journal for Philosophy of Religion* 70 (3): 231–246.
7. 2015. "Reasons-Responsiveness and Time Travel" *Journal of Ethics and Social Philosophy* (2015): 1–7. [Discussion Note]
8. 2015. "Agential-Settling Requires a Conscious Intention" *Journal of Cognition and Neuroethics* 3 (1): 140–155.
9. 2014. "Don't Count on Taurek: Vindicating the Case for the Numbers Counting" *Res Publica* 20 (3): 245–261.
10. 2013. "Skeptical Theism and the Threshold Problem", *Forum Philosophicum* 18 (1): 73–92.

REVIEWS	1. 2014. Review of Robert C. Roberts, *Emotions in the Moral Life*, Cambridge: Cambridge University Press, 2013, *Faith and Philosophy* 31 (3): 337–40.

IN PROGRESS	1. "Fischer's Deterministic Frankfurt-Style Argument" (conditional acceptance at *Erkenntnis*)
	2. "Actualism Has Control Issues" With Travis Timmerman (under review)
	3. "Manipulating Deliberators" (under review)
	4. "Theistic Modal Realism and Gratuitous Evil?" (under review)

APA PRESENTATIONS	1. February 2015. "A Grounding Solution to the Davidsonian Challenge." Colloquium Presentation, Pacific APA, Vancouver, British Columbia. Comments by Alfred Mele.
	2. December 2014. "Manipulating Deliberators: a Challenge for Compatibilists, a Dilemma for Hard Incompatibilists." Colloquium Presentation, Eastern APA, Philadelphia, PA. Comments by Seth Shabo.
	3. April 2014. "Actualism's Ad Hoc Problem." (co-authored with Travis Timmerman) Colloquium Presentation, Pacific APA, San Diego, CA.

CONFERENCE PRESENTATIONS	1. February 2016. "Theistic Modal Realism and Gratuitous Evil?" Ryerson University Philosophy of Religion Work-In-Progress Group.
	2. September 2015. "Counterfactuals of Divine Freedom." Ryerson University Philosophy of Religion Work-In-Progress Group.
	3. April 2015. "Virtue Ethics as a Substantive Normative Ethical Theory." (co-authored with Travis Timmerman) 9th Annual Felician Ethics Conference, Felician College.
	4. November 2014. "A Grounding Solution to the Davidsonian Challenge." Syracuse University Working Papers, Syracuse, NY.
	5. November 2014. "Free Will Primitivism." Society of Christian Philosophers, Eastern Meeting, Lewiston, NY.
	6. October 2014. "Agential Settling Requires Intentionality." Free Will conference at the Center for Cognition and Neuroethics, University of Michigan-Flint.
	7. August 2014. "Normativity without Metaphysical Vagueness." (co-authored with Travis Timmerman) Poster session at the Seventh Annual Rocky Mountain Ethics Conference, University of Colorado Boulder.
	8. August 2014. "The Manipulation Argument, At the Very Least, Undermines Leeway Compatibilism." Free Will Summer School, Center for Consciousness Studies, Moscow.

9. October 2013. "Dual Obligations Hybridism." (co-authored with Travis Timmerman) 65th Annual Northwest Philosophy Conference, Pacific University/ June 2013. British Postgraduate Association Conference, University of Glasgow.
10. November 2012. "Street's Evolutionary Debunking Argument and Epistemic Possibility." Illinois Philosophical Association, University of Illinois Urbana-Champaign.
11. November 2012. "Molinism's Divine Problems." Syracuse University Working Papers, Syracuse, NY.
12. March 2012. "Minimalism, Moral Realism, and Persons." Kent State Graduate Conference, Kent State University.
13. June 2012. "Skeptical Theism and the Threshold Problem." Annual Conference of the Australasian Philosophy of Religion Association, Australian Catholic University/ February 2012. 36th annual Midsouth Philosophy Conference, Rhodes College.

COMMENTS
1. August 2014. Comments on David Killoren's "Robust Moral Realism: An Excellent Religion." Seventh Annual Rocky Mountain Ethics Conference, University of Colorado Boulder.
2. April 2013. Comments on Andrew Jones' "Brutal Persistence." Syracuse Graduate Conference, Syracuse University.
3. April 2012. Comments on Melissa Frankel's "Berkeley on Divine Archetypes." Syracuse University.
4. February 2012. Comments on Tully Borland's & T. Allan Hillman's "God's Hiddenness, Coerced Persons, and the Threat of Moral Nihilism." 36th annual Midsouth Philosophy Conference, Rhodes College.

TEACHING
PHI 107: Theories of Knowledge and Reality (Fall 2014, Spring 2015)
PHI 383 (Online): Free Will (Summer 2015)

TEACHING ASSISTANT
PHI 107: Theories of Knowledge and Reality (Spring 2013)
PHI 191: Ethics and Contemporary Issues (Fall 2013, Spring 2014)
PHI 197: Human Nature (Spring 2012)

PROFESSIONAL SERVICE
2015– Philosophy TV, Managing Editor (philostv.com)
Journal Referee
2015: *Res Philosophica*, *Neuoroethics*, *Philosophia*, *Philosophical Papers*
2016: *Thought*

Conference Referee
>2012. Ohio Philosophical Association

INSTITUTIONAL SERVICE

2014–2015. Graduate Student President, Syracuse University Philosophy
>Department

GRANTS AND AWARDS

2015. Syracuse University Graduate Fellowship
2015. Departmental Summer Research Fellowship (Syracuse)
2013. Departmental Summer Research Fellowship (Syracuse)
2012. Syracuse University Graduate Fellowship

PROFFESSIONAL ACTIVITY

2015. Free Will Summer School, Center for Consciousness Studies of
>Moscow State University (John Martin Fischer, Matthew Talbert)
2012, 2013. Analytic Theology Summer Workshop, Southern Methodist
>University (William J. Abraham, Michael C. Rea)

AFFILIATIONS

The American Philosophical Association
Society for Philosophy of Agency

GRADUATE COURSEWORK

('*' indicates courses audited)
Syracuse University

Free Will	M. Heller (Fall 2013)
Metaethical Naturalism	H. Paakkunainen (Spring 2013)
Leibniz	K. McDaniel (Spring 2013)
Metaphysics	A. Gallois (Fall 2012)
Moral and Political Philosophy	H. Paakkunainen (Fall 2012)
Independent Study: Expressivism	H. Paakkunainen (Spring 2012)
Independent Study: Causation	M. Heller (Spring 2012)
History of Philosophy	K. Richardson (Spring 2012)
Husserl's Logical Investigations	K. McDaniel (Spring 2012)
Ontology	M. Heller (Fall 2011)
Lang/Epist/Mind/Metaphysics	A. Gallois (Fall 2011)
Logic and Language	T. McKay (Fall 2011)

Cornell University

*Free Will	D. Pereboom (Fall 2013)
Metaphysics (On Grounding)	T. Sider (Spring 2013)
Metaphysics (On Properties)	T. Sider (Fall 2011)

Rochester University

Causation	B. Weslake (Fall 2012)

REFERENCES   Mark Heller   André Gallois.

Professor of Philosophy   Professor of Philosophy

Syracuse University   Syracuse University

heller@syr.edu   agallois@syr.edu

John Martin Fischer

Distinguished Professor of Philosophy

University of California, Riverside

john.fischer@ucr.edu

TEACHING   Robert Van Gulick

REFERENCE Professor of Philosophy

Syracuse University

rnvangul@syr.edu

*To receive confidential letters of recommendation, one may also contact Lisa Farnsworth, the office coordinator of Syracuse University's philosophy department: lfarnswo@syr.edu (315-443-2245)