

THE ASSIMILATION ARGUMENT AND THE ROLLBACK ARGUMENT

Christopher Evan Franklin

~Penultimate Draft~

Pacific Philosophical Quarterly 93:3, (2012): 395-416.

For final version go to

<http://onlinelibrary.wiley.com/doi/10.1111/j.1468-0114.2012.01432.x/abstract>

Seth Shabo has presented a new argument that attempts to codify familiar worries about indeterminism, luck, and control. His ‘Assimilation Argument’ contends that libertarians cannot distinguish overtly randomized outcomes from exercises of free will. Shabo claims that the argument possesses advantages over the *Mind* Argument and Rollback Argument, which also purport to establish that indeterminism is incompatible with free will. I argue first that the Assimilation Argument presents no new challenges over and above those presented by the Rollback Argument, and second that the Rollback Argument itself neither presents a deep challenge to, nor raises the cost of, accepting libertarianism.

1. Introduction

In a set of recent articles Seth Shabo has set forth a new argument that attempts to codify familiar worries about indeterminism, luck, and control.¹ His ‘Assimilation Argument’ aims to show that ‘libertarians cannot plausibly distinguish cases where [free will] is ostensibly present from ones where it is clearly absent’.² Since libertarianism’s inception, critics have worried that indeterminism, so far from enhancing control, is inimical to it.³ Some of the most well-known instances of this worry actually find their source in one of libertarianism’s most influential defenders, namely Peter van Inwagen.⁴ van Inwagen has developed this worry into two powerful arguments for the incompatibility of free will and indeterminism: the *Mind* Argument and the Rollback Argument.⁵ Shabo contends that the Assimilation Argument has virtues over both these arguments, in that it ‘better isolates the key challenge for libertarians’ than does the *Mind* Argument and provides stronger support for incompatibilism about free will and indeterminism than does the Rollback Argument.⁶

I have two aims in this paper. The first is to dispatch with the Assimilation Argument—in particular I will show that this argument presents no new challenges over and above those presented by the Rollback Argument.⁷ The second is to dispatch with the Rollback Argument itself. Before presenting and appraising these arguments I will offer a schema of a libertarian theory. I do not intend my defense of libertarianism to depend on this particular theory, but rather offer it as a concrete guide as we wade through these abstract matters.

2. Libertarianism and the Power to Settle

I here develop a familiar event-causal version of libertarianism.⁸ My aim in this paper is not to argue that this theory is true, but rather that the Assimilation Argument and Rollback Argument have no force against it. My reason for developing an event-causal rather than agent-causal theory is partly due to the fact that worries about luck and indeterminism have often been thought to be *more* difficult for event-causal libertarians.⁹ Therefore, if my event-causal version can escape both objections, presumably agent-causal libertarianism can also escape.¹⁰

Event-causal libertarians appropriate the causal theory of action, according to which an event (a bit of behavior or a mental episode) is an action if and only if the event is caused, in the appropriate manner,¹¹ by agent-involving mental states and/or events.¹² Proponents of the causal theory of action disagree over exactly which states and events of an agent must be involved in the causation of action, though the usual suspects include beliefs, desires, reasons, and intentions. Consider the well-known case of a thief¹³ deliberating about whether to rob a poor box at a local parish, and who has reasons to steal (including his great need for money) and reasons to refrain (including his promise to his mother that he would lead a good life), and who decides to steal. According to the causal theory of action, the thief's decision is an action if and only if it is caused, in

the appropriate manner, by mental states and events of the agent—such as his desiring to increase his total net worth and his belief that he could realize this end by so acting.¹⁴

Event-causal libertarians maintain that this is a satisfactory analysis of *action*, but not *free* action. They are, after all, incompatibilists and deterministic agents clearly can satisfy the causal theory of action. Thus, event-causal libertarians believe more must be added in order to transform the causal theory of action into a theory of free action. Central among the required additions is indeterminism: an action is free only if it is undetermined.¹⁵ I believe the best way to integrate the requirement of indeterminism with the causal theory of action is to require that the causal relation that obtains between the agent's non-actional states and events (e.g. desires and reasons) and free choice be nondeterministic.¹⁶ Supposing that the thief's choice to steal satisfies this requirement, then there is a possible world with the same past and laws of nature up until the moment of choice in which the thief does not choose to steal. In this world the thief's desires and reasons that are actually causally efficacious (causing his choice to steal) are causally silent.¹⁷

By so locating indeterminism, we can furnish agents with the ability to do otherwise.¹⁸ Just as an agent's acting on this theory is reducible to the causal interplay of the agent's mental states and events, an agent's possessing abilities to act is reducible to the causal powers of the agent's mental states. The thief's ability to choose consists in certain causal powers of his mental states and events, as does his ability to do otherwise. The thief's having the ability to refrain from stealing (partly) consists in there being a possible world that shares the same past and laws and yet in which the thief's desire to honor his promise to his mother and his belief that stealing would violate this promise jointly cause him to refrain from stealing.

Event-causal libertarianism requires, at minimum, that free choice be nondeterministically caused by agent-involving mental states and events. While I believe more must be added before this theory is defensible, its current form will be sufficient to dismantle the Assimilation Argument and

Rollback Argument. One last comment about this theory is needed. Shabo maintains that an adequate theory of free will must provide an analysis of the power to settle which of the options the agent chooses.¹⁹ Thus, free will requires not only the ability to do otherwise, but also the ability to settle which of two possible courses of action one selects. Unlike the ability to choose to steal and the ability to choose to refrain from stealing, the power to settle which of these two options one pursues is a power that, if exercised, is exercised regardless of the choice made. Supposing the thief possesses and exercises this power, he not only exercises his power to choose to steal, but also his power to settle which course becomes actual. And had he chosen to refrain, he would have exercised this *same* power. For ease, let us refer to this simply as the ‘power to settle’ (I will say more about the nature of this power below). Shabo argues that it is the power to settle that is incompatible with indeterminism, and so it will be important to see how I account for it. On the theory I offer, an agent *S* exercises the power to settle in performing action \square at *t* if and only if he exercises his ability to \square at *t* and had the ability at *t* to \square at *t* (where ‘ \square ’ \neq ‘ \square ’).²⁰ Thus, the power to settle which decision one makes is reducible to exercising an ability to make a choice and possessing the ability to do otherwise.²¹

3. The Rollback Argument

In presenting the Rollback Argument, van Inwagen asks us to imagine an agent, Alice, who is deliberating about whether to choose to tell the truth or to lie, and possesses the ability to choose to tell the truth and the ability to choose to lie (2000, p. 14). Suppose that Alice chooses to tell the truth. Since Alice had the ability to choose to lie, her choice to tell the truth was undetermined (van Inwagen is assuming that the ability to do otherwise is incompatible with determinism). Next we are asked to imagine that God has caused the world to ‘rollback’ to precisely the state of the universe at the moment before Alice chose to tell the truth, let the world evolve from that point, and that we

are in a position to observe this replay. Since Alice's choice is undetermined she might choose to lie or she might choose to tell the truth in the 'replay.' Suppose that God does this a thousand times: a thousand times he causes the world to rollback to precisely the moment before Alice made her choice and then allows things to proceed. Furthermore, imagine that, after watching 726 replays, in about half of the replays Alice has chosen to tell the truth and in the other half she has chosen to lie. After watching each of these replays, van Inwagen claims:

we shall be faced with the inescapable impression that what happens in the seven-hundred-and-twenty-seventh replay will be due simply to chance.... [W]hat other conclusion can we accept about the seven-hundred-and-twenty-seventh replay (which is about to commence) than this: each of the two possible outcomes of this replay has an objective, 'ground-floor' probability of 0.5—and there's nothing more to be said? And this, surely, means that, in the strictest sense imaginable, the outcome of the replay will be a matter of chance. (2000, p. 15)

And if either of Alice's actions would be simply a matter of chance, then Alice lacks free will. After all, as van Inwagen wonders, "If [Alice] was faced with telling the truth and lying, and it was a mere matter of chance which of these things she did, how can we say that...she was able to tell the truth and able to lie? How could anyone be able to determine the outcome of a process whose outcome is a matter of objective, ground-floor chance?" (2000, p. 16).²²

If van Inwagen is correct that Alice lacks the ability to choose to tell the truth and the ability to choose to lie, then it follows, from my theory, that she is not free. But since Alice seems to lack these abilities simply because her choice was undetermined, libertarianism appears incoherent: its requiring indeterminism for free will precludes the very possibility of the existence of free will—libertarians seem to mistake a sufficient condition for the absence of free will (i.e. indeterminism) for a necessary condition for its existence. But what exactly is the ability that van Inwagen claims Alice lacks and why should we think she lacks it? Shabo contends that it is not simply the ability to tell

the truth (forthcoming, p. 8), but instead the ‘power [to settle] whether she voluntarily and intentionally lies or voluntarily and intentionally tells the truth’ (forthcoming, p. 9).²³ According to Shabo, indeterminism threatens free will because it threatens the power to settle. Note that this power does not merely reduce to the power to perform some action. It is possible that Alice possessed the ability to choose to lie and that she exercised this ability, yet her choice be causally determined. In such a case Alice will have failed to manifest the power to settle—in such a case Alice possesses the power to choose to lie, but lacks the power to settle which choice she makes (recall that we are assuming incompatibilism about free will and determinism). In Shabo’s words, there is a ‘logical gap’ (forthcoming, p. 11) between making a choice and exercising the power to settle over that choice: having the ability to \square is not sufficient for having the power to settle whether one \square -s.

Shabo does not think that the Rollback Argument proves that indeterminism is incompatible with the power to settle, but he does think it ‘[forces] a reasonable question: if nothing prior to the decision varies between [the various rollbacks] – if there is no difference in the agent’s states or actions – in virtue of what does the outcome qualify as *up to her*? What makes it an exercise of her *power over* the course she adopts?’ (2011, p. 113). The Assimilation Argument is offered as proof that libertarians *cannot* offer a plausible answer to this question.

4. The Assimilation Argument

In preparation for the Assimilation Argument, consider the following five cases:²⁴

Case 1: A particle in a remote region of spacetime has an objective, .5 probability of swerving one way, and an equal probability of swerving another way. No other event in the universe influences this probability distribution or which way the particle swerves, given this distribution. The particle ends up swerving in the first of the two possible ways.

Case 2: A similar particle in a non-remote region of spacetime will swerve in one direction or another, with each outcome having an equal, objective probability. Scientists are using a device to detect the particle's behavior. If the particle swerves one way, the device transmits one signal to a receiver implanted in Alice's brain; the receiver then causally determines Alice's brain to go into a state that is intrinsically indistinguishable from an ordinary, executive intention to tell the truth in response to the question she has been asked. If the particle swerves the other way, the device transmits a different signal to the receiver, which causally determines Alice's brain to go into a state intrinsically indistinguishable from an ordinary, executive intention to lie. In the event, the particle swerves in the first direction, with the result that Alice exhibits truth-telling behavior, behavior that is caused in a normal way by the relevant neural state.

Case 3: This case is exactly like Case 2, except that the device has been miniaturized (swerving particle and all) and placed inside Alice's brain, where it induces one neural state or the other, eliminating the need for a remote receiver.

Case 4: The scientists introduce an upgraded device. The upgraded device harnesses indeterminacy in Alice's own neural pathways, indeterminacy that must be present if her action is to meet the libertarian's conditions for free will. The upgraded device can target pathways that are causally inert, pathways that are involved in reflex-type behavior, or pathways that are involved in the formation of intentions. When the pathways in the last of these regions are targeted, the upgraded device charges a particle (or group of particles) in that pathway, causing that particle(s) to behave like the particle in Case 3. That is, the particle(s) will swerve one way or another, with each outcome having an equal, objective probability. In this case, however, the outcome is an intention to lie or an intention to tell

the truth—or at least a neural state that is intrinsically indistinguishable from one of these intentions or the other.

Case 5: There is no input from any external device. Due to its ordinary, endogenous workings, Alice's brain state as she mulls her options is such that lying and telling the truth are equally probable outcomes. She intends to make up her mind, and this intention causally contributes to her deciding to tell the truth, a decision that is based on her reasons for doing so.²⁵ (forthcoming, pp. 11-12)

The cases advance from cases in which it is obvious that no one exercises free will to paradigm cases in which an agent satisfies libertarianism, and the challenge before libertarians is to explain where randomness ceases and freedom begins. According to the Assimilation Argument they cannot satisfactorily discharge this explanatory burden:

- (A1) In each of Cases 1–3, it isn't up to Alice (or anyone) which of the two causally possible outcomes described in the case actually ensues.
 - (A2) It's up to Alice which outcome ensues in Case 4 only if it's up to her in Case 3.
 - (A3) It's up to Alice which outcome ensues in Case 5 only if it's up to her in Case 4.
- Therefore,
- (A4) The actual outcome in Case 5—Alice's deciding to tell the truth for the appropriate reasons—isn't up to her. (forthcoming, p. 15-16)²⁶

Shabo contends that the Assimilation Argument goes beyond the Rollback Argument in the following manner: while the Rollback Argument poses a deep question about the compatibility of free will and indeterminism, the Assimilation Argument shows that libertarians cannot answer this question satisfactorily.

The contentions premises are (A2) and (A3). These premises are true, respectively, only if there is no relevant difference between cases 3 and 4 or cases 4 and 5. I will now argue that there is a relevant difference between cases 4 and 5 and thus that (A3) is false.

5. Farewell to the Assimilation Argument

The crucial difference between cases 4 and 5 is that Alice *acts* in case 5, but not in case 4. Consider the last sentence of Shabo’s description of case 4: ‘In this case, however, the outcome is an intention to lie or an intention to tell the truth—or at least a neural state that is intrinsically indistinguishable from one of these intentions or the other.’²⁷ The first disjunct is false. An event’s being an action depends on its causal etiology—specifically an action is an event that has agent-involving mental states and events in its causal history.²⁸ In case 4, Alice’s “choosing”, her neural state that is intrinsically indistinguishable from her intention to tell the truth, does not have the right causal etiology to count as an action; her neural event is caused by *a device*, not by any of her mental states, and thus there is no choice or intention at all. However, in case 5, Alice’s choice is genuine because it is caused, in the appropriate manner (we can assume), by her agent-involving mental states and events. Therefore, (A3) seems false since there is a relevant difference between cases 4 and 5.

The difference between cases 4 and 5 is not just some minor detail, but is a free-will-relevant difference. An agent exercises free will only if he acts. Since Alice does not act in case 4 we know that she does not exercise free will in case 4; but Alice does act in case 5 and so the obstacle to Alice’s exercising free will in case 4 is absent in case 5. The difference between these cases consists in the fact that a necessary condition for exercising free will is absent in case 4 but present in case 5. So we have arrived at an explanation of how overtly randomized outcomes differ from exercises of free will: an overtly randomized outcome is not an action while an exercise of free will must be action. Consequently, we seem to have met the Assimilation Argument’s challenge ‘to explain how putative exercises of free will are relevantly different from *overtly* randomized outcomes.’²⁹

One might worry that my claim concerning case 4 above that Alice’s ‘neural event is caused by *a device*, not by any of her mental states, and thus there is no choice or intention at all’ depends on rejecting the widely held view that the mental is multiply realizable.³⁰ Since the mental is multiply realizable (and as we are assuming token identity physicalism), it is possible for token mental states

to be identical to token physical states of different *types*. Suppose that in case 5 that Alice's reason-state R that favors her telling the truth is identical to neural state N. Given that the mental is multiply realizable, one might argue that in case 4 R is identical to a certain state of the device D. One might think that the device is integrated into Alice's brain in a way that qualifies it as a prosthetic device. Consequently, the crucial difference between cases 4 and 5 concerns the physical states that R is identical to. In case 5 R is identical to N and in case 4 it is identical to D. But now my contention that there is a free-will-relevant difference between cases 4 and 5 is called into question. I claimed that the difference between cases 4 and 5 is that Alice makes a genuine choice in case 5, but not in case 4, and I defended this claim by arguing that in case 4 Alice's neural event that is intrinsically indistinguishable from a choice was caused by D and not by any of her mental states. But this claim seems false: for D just is R and so D's causing this neural event just is R's causing this neural event. We have, then, no reason to think that in case 4 the neural event that is intrinsically indistinguishable from a choice is not a genuine choice, just as it is in case 5. Therefore, so the objection goes, there is not a free-will-relevant difference between cases 4 and 5.

The contention that R is identical to D in case 4 requires that we assume that R and D are functionally equivalent. The problem with assessing this assumption is that Shabo's description of case 4 leaves the matter underdetermined: there are not enough details in case 4 to determine whether R and D are functionally equivalent, e.g. whether R and D have the same characteristic inputs and outputs. However, we can dispatch this objection without having to bury ourselves in the details. Either R is identical to D or it is not. If R is not identical to D, then my original claim holds: in case 4 Alice's neural state that is intrinsically indistinguishable from her choice in case 5 is caused by D and not R, and so is not a genuine choice. It follows that there is a free-will-relevant difference between cases 4 and 5.

Consider now the first horn, under which R is identical to D. This means that R and D are functionally equivalent and suggests that Alice makes a genuine choice in both cases 4 and 5. This further suggests that there is not a free-will-relevant difference between these cases. The main difference between cases 4 and 5 merely has to do with which physical states R is identical to. But this difference does not make a difference, since in each case Alice's reasons that favor telling the truth are what cause her choice to tell the truth. But there now emerges a free-will-relevant difference between cases 3 and 4. Cases 3 and 4 both feature a neural state that is intrinsically indistinguishable from an ordinary choice and that has, somewhere in its causal etiology, the activity of a device implanted in Alice's brain by scientists. Under the present horn I contend that this neural state is a genuine choice in case 4 but not case 3, and what accounts for this difference is the manner in which this device operates in the different cases.

Consider how case 3 unfolds. There is a single particle located in Alice's brain that might swerve one of two possible directions given the past and laws of nature. There is also a device located in Alice's brain that will detect which direction the particle swerves and will cause Alice to be in a neural state that is intrinsically indistinguishable from a choice to tell the truth if the particle swerves in one direction and will cause Alice to be in a neural state that is intrinsically indistinguishable from a choice to lie if the particle swerves in the other direction. As it so happens, the particle swerves in the first direction, the device detects this, and as a result (deterministically) causes Alice to be in a neural state that is intrinsically indistinguishable from a choice to tell the truth. Call this neural state N*. Although N* is intrinsically indistinguishable from a choice to tell the truth, it is not a choice to tell the truth. Two events can be intrinsically indistinguishable and yet only one an action, since part of what makes an event an action depends on the event's causal history. N* is not a genuine choice because it does not have any mental state among its proximate causal antecedents. The device causes Alice to be in state N* completely independent of her reasons

for telling the truth.³¹ What determines the activity of the device is the swerving of a single particle, not Alice's weighing various reasons for or against telling the truth, nor her perceptions concerning the ethics or wisdom of truth-telling, nor her taking there to be certain reasons in favor of this action, etc. And given this difference between how the device functions in cases 3 and 4, it cannot plausibly be maintained that the state of the activity of the device in case 3 is identical to R. Therefore, there is a free-will-relevant difference between cases 3 and 4, since Alice makes a choice to tell the truth in case 4 but does not make any such choice in case 3.

Given a degree of under-description in Shabo's cases, it is unclear where the libertarian should stake his claim that there is a free-will-relevant difference in this series of cases, because it is unclear whether R is identical to D in case 4. However, whether or not R is identical to D in case 4, the Assimilation Argument fails. If R is not identical to D, then there is a free-will-relevant difference between cases 4 and 5, and thus premise (A3) is false. If R is identical to D, then there is a free-will-relevant difference between cases 3 and 4, and thus premise (A2) is false. Either way the Assimilation Argument is unsound. The remainder of my discussion of the Assimilation Argument will take place under the assumption that D is not identical to R, and thus that Alice acts in case 5 but not in case 4. However, my subsequent defense of the contention that there is a free-will-relevant difference between cases 4 and 5 applies *mutatis mutandis* to cases 3 and 4 if we instead assume instead that R is identical to D.

Perhaps surprisingly Shabo never disputes my contention that Alice acts in case 5 but not in case 4.³² Instead, he argues that this difference does not suffice to dismantle the Assimilation Argument: 'From the fact that the causally undetermined outcome in Case 5 is Alice's decision to tell the truth, a mental act she performs for reasons she has, it doesn't follow that she settles which outcome ensues, or which set of reasons she acts on. Why think, then, that the outcome in Case 5 is up to her?'³³ But this is the old bait and switch. Shabo begins by asking one question, 'What is the

relevant difference between cases 4 and 5?’ but criticizes our answer because it fails to answer a different question, ‘In virtue of what does Alice exercise the power to settle in case 5?’ Shabo appears to think his switching questions is justified because he assumes that an adequate explanation of the difference between cases 4 and 5 ‘would plausibly close the logical gap identified [above].’³⁴ In other words, Shabo assumes that an answer to the first question is satisfactory only if it is a satisfactory answer to the second question. But this is not the case—answers to these questions can and do come apart.

To appreciate this, consider the Manipulation Argument, which has the following core:

1. Agents who are severely manipulated are not free.
2. There is no relevant difference between “ordinary” deterministic agents and severely manipulated agents.
3. Therefore, deterministic agents are not free.

Defenses of premises (1) and (2) take the form of presenting cases in which an agent putatively satisfies all the necessary compatibilist conditions for freedom and yet, due to manipulation, is not free.³⁵ Incompatibilists then argue that there is no relevant difference between this form of manipulation and ordinary deterministic settings. In some of these manipulation cases, the agent’s mental states are controlled and induced each moment by manipulators. In response to such cases, some compatibilists have maintained that such manipulated agents are not persons and that being a person is necessary for exercising free will.³⁶ Suppose an incompatibilist replies that this difference does not show that the agent in the ordinary deterministic world is free, for there is a logical gap between being a person and exercising free will. It is clear how compatibilists should respond: they should concede that there is such a gap, but maintain that it is irrelevant. The original compatibilist reply defeats this instance of the Manipulation Argument because it shows that the cases are not similar, and thus that premise (2) is false.³⁷ Compatibilists can defeat the Manipulation Argument without thereby showing that agents in deterministic worlds can be free.

The Assimilation Argument is structurally similar to the Manipulation Argument. It, like the Manipulation Argument, begins with a case (or cases) in which freedom is uncontroversially absent. It also, like the Manipulation Argument, depends on a premise concerning the relevant similarity between the cases presented, and so it, like the Manipulation Argument, fails if there is a free-will-relevant difference between the cases.³⁸ Consequently, since the Manipulation Argument can be dismantled without closing the relevant logical gap, it seems that the Assimilation Argument can also be dismantled without closing the relevant logical gap—we can explain why cases 4 and 5 differ in a free-will-relevant way without thereby closing the logical gap. Hence, in order to defeat the Assimilation Argument it is sufficient to show that Alice acts in case 5 but not in case 4.

But arguably cases 4 and 5 differ in even more ways than this, and perhaps these additional differences do close the logical gap. Given the role of the upgraded device in case 4, Alice will not act either when the device causes her to ‘choose’ to lie or to ‘choose’ to tell the truth.³⁹ The device appears to rob Alice of the ability to choose to lie and the ability to choose to tell the truth, at least for the time being. After all, given the workings of the device it is impossible for Alice to choose to tell the truth or choose to lie; the device is interfering with and preventing her from exercising her own agency.⁴⁰ So in case 4 Alice not only fails to act, she also lacks the ability to choose to tell the truth and the ability to choose to lie. In case 5, however, Alice acts (she chooses to tell the truth) and so presumably she also has the ability to choose to tell the truth.⁴¹ Since Alice’s choice was undetermined, there is a possible world with the exact same past and laws of nature in which Alice does not choose to tell the truth. Moreover, we have been asked to assume that there is such a world in which she chooses to lie.⁴² Libertarians will maintain that because of the location of indeterminism *and* other features of Alice (specifically the causal powers of her mental states), she also possesses the ability to choose to lie. So not only do cases 4 and 5 differ with respect to Alice’s acting, they also differ with respect to what abilities she possesses. And according to the libertarian

theory I offered above, these differences *are* sufficient to close the logical gap. I claimed that an agent exercises his power to settle by making a decision, possessing the ability to make that decision, and possessing the ability to do otherwise. In case 5 Alice satisfies all these conditions, whereas in case 4 she fails to satisfy any of them; or, at the very least, the Assimilation Argument does not present any obstacle to making these claims.⁴³

Shabo is sensitive to this kind of reply, for after considering a similar objection to the Assimilation Argument, he offers the following ‘variant’ on van Inwagen’s Rollback Argument, which he claims is ‘effective’ against this kind of objection⁴⁴:

Suppose that God sets up two cosmic video displays side by side. The first shows replays of the original Case 4; the second replays the original Case 5. Each display is programmed to repeat a five-second clip of the run-up to Alice’s decision, replete with stages that are causally undetermined by their predecessors. Now suppose that we are able to zoom in—way in—on the action. On each display, we see images of neural activity in the deliberative centers of Alice’s brain. On the first monitor, we see a graphical representation (in a cloud of red and blue dots) in these neural centers as Alice mulls her options; then the upgraded device emits an energy pulse, which is immediately followed by a particular activity signature (always the same) in the affected pathways, and then by one decision “node” or the other “lighting up.” Events on the second display unfold in much the same way, except that the device is absent. First we see the same activity in Alice’s deliberative centers as on the first display; then the same (or a seemingly indistinguishable) activity signature, only without the mediation of the device, followed by one decision node or the other lighting up.⁴⁵

Shabo goes on, ‘Given that the only observable difference between the two cases is that the device operates in Case 4 but not in Case 5, it’s hard to see the outcomes in Case 5 as up to her when the outcomes in Case 4 so clearly aren’t.’⁴⁶ In this variant on the Rollback Argument, we have two

displays on which we watch two different rollbacks: rollbacks of case 4 and rollbacks of case 5. Alice does not exercise free will in case 4. Cases 4 and 5, so the argument goes, are relevantly similar. Therefore, Alice does not exercise free will in case 5.

But this argument fails for the same reason that the Assimilation Argument fails: it turns on an implausible claim of similarity. Cases 4 and 5, as we have seen, are importantly different since in case 4 Alice does not choose to tell the truth, does not have the ability to make this choice, and does not have the ability to choose to lie, whereas in case 5 Alice chooses to tell the truth, possesses the ability to choose to tell the truth, and, arguably, possesses the ability to choose to lie. Hence, even if there are no ‘observable’ differences between cases 4 and 5, there are free-will-relevant differences.

I conclude, therefore, that the Assimilation Argument does not present any serious worry for libertarians. First, we can readily distinguish cases 4 and 5 in a way that allows us to explain how exercises of free will differ from overtly randomized outcomes, since Alice acts in case 5 but not case 4. Moreover, Alice arguably differs in additional ways that do close the logical gap: Alice does not merely choose to tell the truth, she also possess the ability to do this and to choose to lie. Finally, we saw that trying to employ a variant on the Rollback Argument to defend the conclusion that there is no relevant difference simply reiterates the mistakes that plagued the Assimilation Argument.

6. Farewell to the Rollback Argument

But this conclusion may not go far in assuaging critics’ suspicion that libertarianism is incoherent. For many, the Rollback Argument itself establishes that libertarianism is incoherent, or at least is plagued by deep and potentially intractable conceptual difficulties. Rather than trying to use the Assimilation Argument to establish the conclusion of the Rollback Argument (as Shabo did), it

might seem that the Rollback Argument directly shows that free will is incompatible with indeterminism. So let us consider this argument directly.⁴⁷

What exactly is the Rollback Argument? It is really more of a thought experiment than an argument. We are asked to imagine ourselves watching replays of Alice's deciding what to do. Since, *ex hypothesi*, Alice's original choice to tell the truth was undetermined, in each replay it is possible that she not make this choice. According to the thought experiment, in some of the replays she chooses to lie instead. van Inwagen believes that after watching these replays for some time, we will be left with the inescapable impression that Alice's choice, whatever it is, is simply a matter of chance. Moreover, it is maintained that Alice's choice is simply a matter of chance in a sense that is incompatible with freedom.⁴⁸

The problem with this thought experiment is that it is metaphysically impossible. It asks us to imagine that God continually rolls back time just before Alice makes a decision, lets things proceed, and that, given the presence of indeterminism, we get *different* outcomes. But if God did this, Alice would *always* make the same decision, since it is metaphysically impossible for one and the same world to have different futures. Possible worlds are composite objects. Although it is contentious just what sort(s) of object(s) composes possible worlds, the most widely espoused views include propositions, states of affairs, and concrete objects.⁴⁹ The identity conditions of a possible world are solely determined by the objects that compose it. Supposing that possible worlds are composed of states of affairs, what distinguishes possible world W_1 from possible world W_2 is solely a function of the states of affairs that compose each world. If W_1 and W_2 are composed of all and only the same states of affairs, then W_1 and W_2 are identical. Importantly, a possible world has all its components essentially: a possible world could not have been different. Consider then a world that is partly composed of the state of affairs of *Alice choosing to tell the truth*. This is an essential part of this world. It is metaphysically impossible for this world to be actual and the state of affairs *Alice*

choosing to tell the truth not obtain. But this is exactly what van Inwagen's thought experiment asks us to imagine. It asks us to image that one and the same possible world could have different components. Therefore, van Inwagen's thought experiment is metaphysically impossible.

But perhaps this reply seems shallow. After all, Alice's choice is undetermined, and consequently there are *different* possible worlds that share the exact same past and laws of nature as the actual world in which there are different outcomes. Similar to what Shabo imagines in his variant on the Rollback Argument, we can imagine being presented with an array of monitors that present dramatic (yet accurate) representations of the outcome of Alice's identical process of deliberation in different possible worlds.⁵⁰ And perhaps this thought experiment, will, just as much as van Inwagen's original Rollback Argument, leave us with the inescapable impression that Alice's choice is simply a matter of chance.⁵¹

It may leave some with such an impression, but a libertarian would have to get himself rather muddled to be worried by *this* thought experiment. We are asked to imagine observing many different possible worlds, for ease let us suppose a hundred. All of these worlds share the exact same past and laws of nature as the actual world, and in about fifty of these worlds Alice chooses to tell the truth and in the rest she chooses to lie. If this leaves a libertarian with the inescapable impression that Alice's choice was simply a matter of chance, then he must have never understood what indeterminism was in the first place.

To appreciate this point, suppose we consider all the possible worlds that share the same past and laws of nature as the actual world and that, unlike our original thought experiment, in all these worlds Alice chooses to tell the truth. *This* will lead libertarians to have the inescapable impression that Alice was not free. Because of this libertarians believe that freedom requires variability between worlds that share the same past and laws. In other words, libertarians contend

that if we watched Alice's decision on various monitors, we would think she is free only if she made different choices in at least some of these worlds.

Enter the proponent of the Rollback Argument, who presents just such a scenario and tells us that precisely because there is variability from monitor to monitor we ought to doubt Alice's freedom. It is hard to see why this ought to cause any libertarian who understands what indeterminism is any pause for concern. To claim that a free action must be undetermined just is to claim that when an agent acts freely there are possible worlds that share the same past and laws of nature as the actual world in which the agent does not perform this action.⁵² Thus, if one is inclined to think indeterminism is necessary for freedom, the mere fact that there is variability between the outcomes of Alice's deliberation should not cause one any concern; indeed, it should function in the exact opposite way.

But perhaps the Rollback Argument is not aimed at committed libertarians, but rather agnostics, those who have yet to make up their minds about the truth of incompatibilism about free will and determinism.⁵³ Or perhaps if the Rollback Argument does not refute libertarianism it does raise the cost of accepting it. Consider again the Manipulation Argument. It is widely acknowledged that it is possible for agents who satisfy compatibilist conditions of freedom to be subject to certain forms of manipulation that we pre-theoretically thought were incompatible with freedom.⁵⁴ And although this may not refute compatibilism, it does raise its cost. Suppose we interpret the Rollback Argument along similar lines, the idea being: of course libertarians can *assert* that since Alice satisfies their theory's conditions she is free, just as compatibilists can assert that the aforementioned manipulated agents satisfy *their* theory's conditions and so are free; and while this line of response is open to both parties, it does not eliminate the fact that these arguments show that there is something surprising and untoward about the relevant theory—it does not eliminate the fact that the proponent of each view has a bullet to bite.

I do not think this analogy holds. There is a crucial difference between the Rollback Argument and Manipulation Argument. The Manipulation Argument shows that compatibilism has a surprising and unintuitive *consequence*: agents who are severely manipulated can nevertheless be free. It begins with a case in which, pre-theoretically, we would judge that the agent is not free, and then shows that there is no relevant difference between the manipulated agent and ordinary deterministic agents. More specifically, the argument seeks to reveal that certain activities that would seem incompatible with the exercise of free agency (i.e. certain forms of manipulation) turn out to be compatible with exercises of free agency *if* we accept compatibilism. Crucially, one can perfectly well understand what compatibilism is without realizing that it has this consequence. It is because the argument shows that compatibilism has this untoward consequence that the argument raises the cost of accepting compatibilism.⁵⁵ But the Rollback Argument is not like this. The Rollback Argument does not draw out a consequence of libertarianism, but simply *describes* libertarianism and asserts that after understanding what this theory is, we ought to conclude that it is deeply problematic, or perhaps metaphysically incoherent. One cannot understand what libertarianism is without realizing that there will be observed variability from monitor to monitor. To fail to appreciate this is to fail to understand indeterminism, and one cannot understand libertarianism without understanding indeterminism. Thus, to present the Rollback Argument is simply to describe libertarianism in a rather colorful way. But one cannot raise the cost of libertarianism by simply describing it.

The true analogue of the Rollback Argument is the following. Suppose that after a long and tiresome period of deliberation in which Alice felt deeply torn between two choices, she chooses to tell the truth, and her choice was causally determined. Now imagine that you have before you thousands of monitors that show dramatic representations of what Alice does in worlds that share

the same past and laws as the actual world. To your great astonishment she *always* makes the same choice. Surely this will leave you with the inescapable impression that Alice is not free.

It does leave me with such an impression; perhaps it also leaves you with such an impression. But regardless, this thought experiment cannot be used as an *argument* against compatibilism because it simply describes compatibilism in a rather colorful way. Therefore, I doubt that libertarians should be worried about the Rollback Argument. This argument does not threaten their theory since it neither presents a deep problem for libertarianism nor raises its cost. The argument, at best, expresses an intuition that indeterminism is incompatible with free will. But to have an intuition is not to have an argument.⁵⁶ Libertarians cannot explain away this intuition unless they are told *why* indeterminism is inimical to control. The Rollback Argument at best expresses the worry that indeterminism is inimical to control, neither explaining nor defending it; and without an explanation or defense, there is nothing left for libertarians to respond to.⁵⁷

Department of Philosophy
Biola University

¹ Seth Shabo, “Why Free Will Remains a Mystery”, *Pacific Philosophical Quarterly* 92 (2011): pp. 105-125; “Free Will and Mystery: Looking Past the *Mind* Argument”, *Philosophical Studies* (forthcoming). I will assume, as does Shabo, the truth of incompatibilism—the thesis that free will is incompatible with determinism. However, I will also sometimes use ‘incompatibilism’ to refer to the thesis that free will is incompatible with *indeterminism*. All latter uses of the term will be made explicit. To my knowledge no one has ever argued that free will is incompatible with indeterminism strictly speaking. The truth of indeterminism merely requires that there be at least one event that might not have occurred given the past and laws. It is very hard to see how the mere truth of this thesis could threaten the existence of free will. Rather, philosophers have worried that events that are undetermined cannot be instances of exercises of free will—it is in this sense that philosophers have argued that free will is incompatible with indeterminism. All uses of ‘incompatibilism about free will and indeterminism’ should be interpreted in this latter, more restricted fashion.

² Shabo, “Free Will and Mystery: Looking Past the *Mind* Argument”, p. 2. The page numbers to Shabo “Free Will and Mystery: Looking Past the *Mind* Argument” are taken from the online version published at <http://www.springerlink.com/content/u00761235558lh42/>.

³ See Ishtiyaque Haji, “Indeterminism and Frankfurt-type Examples”, *Philosophical Explorations* 2 (1999): pp. 42-58; “Control Conundrums: Modest Libertarianism, Responsibility, and Explanation”, *Pacific Philosophical Quarterly* 82 (2001):

pp. 178-200; R. E. Hobart, “Free Will as Involving Determination and Inconceivable Without It”, *Mind* 43 (1934): pp. 1-27; David Hume, *A Treatise of Hume Nature*, 2nd edition, eds. L.A. Selby-Bigge and P.H. Nidditch (Oxford: Oxford University Press, 1740 (1978)); Alfred R. Mele, “Ultimate Responsibility and Dumb Luck”, *Social Philosophy & Policy* 16 (1999): pp. 274-93; and *Free Will and Luck* (New York: Oxford University Press, 2006).

⁴ Peter van Inwagen, *An Essay on Free Will* (Oxford: Clarendon Press, 1983); and “Free Will Remains a Mystery”, *Philosophical Perspectives* 14 (2000): pp. 1-19. van Inwagen thinks that these arguments fail, but he claims that he does not know why (cf. van Inwagen, “Free Will Remains a Mystery”, p. 18). Thus, he thinks that free will is a mystery.

⁵ There is actually quite a bit of difficulty in relating the various arguments put forth in defense of incompatibilism about free will and indeterminism. van Inwagen first presented the *Mind* Argument (so named because its proponents were so frequently published in the philosophy journal *Mind*) in his *An Essay on Free Will*. The argument presented there had three instances, which are logically independent of each other—the failure of one does not entail the failure of any of the others. However, most discussions of the *Mind* Argument restrict themselves to the third instance (see for example Randolph Clarke, *Libertarian Accounts of Free Will* (New York: Oxford University Press, 2003); Alicia Finch and Ted Warfield, “The Mind Argument and Libertarianism”, *Mind* 107 (1998): pp. 515-528; Christopher Evan Franklin, “Farewell to the Luck (and *Mind*) Argument”, *Philosophical Studies* 156 (2011): pp. 199–230; and Shabo, “Free Will and Mystery: Looking Past the *Mind* Argument”). Also, when presenting the Rollback Argument almost twenty years later, van Inwagen claims that this argument is ‘one way to formulate the *Mind* Argument’ (van Inwagen, “Free Will Remains a Mystery”, p. 10). He never explains in what sense these four ways of formulating the *Mind* Argument are four formulations of the *same* argument (beside the fact that instances of the arguments supposedly appeared in *Mind*). The only thing I can tell that they have in common is their conclusion: that free will is incompatible with indeterminism.

Attempts at taxonomy are further exacerbated by bringing the Luck Argument into the picture, which itself has numerous instances. In Franklin, “Farewell to the Luck (and *Mind*) Argument”, pp. 199-202, I suggested, for the sake of clarity, that ‘the *Mind* Argument’ be used exclusively to refer to the third instance of the *Mind* Argument, and that ‘the Luck Argument’ be used to refer to arguments that share the following core premises:

- (i.) If an action is undetermined, then it is a matter of luck.
- (ii.) If an action is a matter of luck, then it is not free.

On this taxonomy, the first instance of the *Mind* Argument and the Rollback Argument end up counting as instances of the Luck Argument since they share this argument form.

But what about the second instance of the *Mind* Argument? This argument turns out, as we will see, to bear important similarities to Shabo’s Assimilation Argument. We might then have ‘the Assimilation Argument’ refer to any argument that shares the following core premises:

- (i.) There are no relevant differences between undetermined actions and overtly randomized outcomes.
- (ii.) No one is free with respect to overtly randomized outcomes.

On this taxonomy we arrive at three main arguments for incompatibilism about free will and indeterminism: the *Mind* Argument, the Luck Argument, and the Assimilation Argument.

⁶ Shabo, “Free Will and Mystery: Looking Past the *Mind* Argument”, p. 2; cf. Shabo, “Why Free Will Remains a Mystery”, p. 113.

⁷ I will not call into question Shabo’s contention that the Assimilation Argument better isolates the main challenge for libertarianism than the *Mind* Argument, though I am inclined to disagree.

⁸ See Clarke, *Libertarian Accounts of Free Will*, chapters 3-5 for an important discussion of this theory. Although Clarke defends this theory against certain objections, he ultimately rejects it (Clarke, *Libertarian Accounts of Free Will*, chapter 6). I defend this account against Clarke’s, as well as others, criticisms in Christopher Evan Franklin, “The Problem of Enhanced Control”, *Australasian Journal of Philosophy* 89 (2011): pp. 687-706.

⁹ Cf. Timothy O’Connor, *Persons and Causes: The Metaphysics of Free Will* (New York: Oxford University Press, 2000); and Derk Pereboom, *Living without Free Will* (New York: Cambridge University Press, 2001).

¹⁰ van Inwagen actually offers the Rollback Argument not as an argument that shows that libertarianism is false, but as an argument ‘for the conclusion that the concept of *agent causation* is entirely irrelevant to the problem of free will (“Free Will Remains a Mystery”, p. 11). van Inwagen’s contention is that *if* event-causal libertarians cannot address the Rollback Argument, then neither can agent-causal libertarians. I believe that this conclusion is correct, but I also contend that event-causal libertarians can dismantle the Rollback Argument. Also, van Inwagen is mistaken to assume that the concept of agent-causation is relevant to the problem of free will only if it shows how free will is compatible with indeterminism (van Inwagen, “Free Will Remains a Mystery”, p. 13)—for there are many difficulties under the heading ‘the problem of free will,’ and it is certainly possible that the concept of agent-causation can help with these other difficulties. Roderick Chisholm, “Freedom and Action”, in *Freedom and Determinism*, ed. Keith Lehrer (New York: Random House, 1966: pp. 11-44); Clarke, *Libertarian Accounts of Free Will*, chapters 8-9; and O’Connor, *Person and Cause* illustrate ways in which this concept can be useful even if it does not help in dismantling the Rollback Argument.

¹¹ This qualification is required to rule out cases of causal deviance. One respect in which the account I offer is only a schema is that I leave the content of this all important clause unspecified. See John Bishop, *Natural Agency: An Essay on the Causal Theory of Action* (Cambridge: Cambridge University Press, 1989) for one of the most careful and detailed attempts to explain what exactly ‘appropriate manner’ amounts to.

¹² Bishop, *Natural Agency*; Myles Brand, *Intending and Acting: Toward a Naturalized Action Theory* (Cambridge: MIT Press, 1984); Donald Davidson, *Essays on Actions and Events* (New York: Oxford University Press, 1980); Berent Enc, *How We Act: Causes, Reasons, and Intentions* (New York: Oxford University Press, 2003); Alvin Goldman, *A Theory of Human Action* (Princeton: Princeton University Press, 1970); and Alfred R. Mele, *Springs of Action* (New York: Oxford University Press, 1992).

¹³ van Inwagen, *An Essay on Free Will*, p. 144.

¹⁴ Perhaps the causal theory of action is best stated as a sufficient condition, allowing that events caused by conscious, rational substances (if possible) would also suffice for action. I will, for the sake of simplicity, set aside this possible complication.

¹⁵ Libertarians (e.g. Robert Kane, *The Significance of Free Will* (New York: Oxford University Press, 1996)) usually draw a distinction between directly and indirectly free actions. Directly free actions must be undetermined, but indirectly free actions can be determined so long as they are appropriately connected (e.g. caused) by directly free actions. Our discussion will be restricted to directly free actions.

¹⁶ In *Libertarian Accounts of Free Will*, p. 29, Clarke refers to this account as *centered* event-causal libertarianism. Centered accounts contrast with *deliberative* accounts (e.g. Daniel Dennett, “On Giving Libertarians What They Say They Want”, in his *Brain Storms: Philosophical Essays on Mind and Psychology* (Montgomery: Bradford books, 1978: pp. 286-299); John Martin Fischer, “Libertarianism and Avoidability: A Reply to Widerker”, *Faith and Philosophy* 12 (1995): pp. 119-125; and Mele, “Ultimate Responsibility and Dumb Luck”) that locate indeterminism earlier in the action-sequence, such as the process of deliberation. In Franklin, “Farwell to the Luck (and *Mind*) Argument”, pp. 206-208, I argue that libertarians should prefer centered to deliberative accounts.

¹⁷ To be clear, the event-causal libertarian account I advocate does not envision a single set of reasons having two possible outcomes. For example, it is not the case that the thief’s reasons for stealing might cause his choice to steal or might cause his choice to refrain. Rather, on this account, given the presence of indeterminism, different reason-states might be causally efficacious. If the thief decides to steal, then his reasons that favor this choice will be causally efficacious. Alternatively, if the thief decides to refrain, then other reasons—namely those favoring refraining—will be causally efficacious. I am grateful to Neal Tognazzini for helping me see the need to make this point explicit.

¹⁸ In Franklin, “The Problem of Enhanced Control”, pp. 691-699, I argued that free will is constituted by an agent’s abilities and opportunities to exercise his abilities. Moreover, I argued that it is the opportunity to do otherwise, rather than the ability to do otherwise, that is incompatible with determinism. I believe this view provides a more satisfactory analysis of both free will and the problem facing compatibilists. But since nothing of present concern will turn on this point, and since both Shabo and van Inwagen assume that free will is solely built up out of abilities, I will join them in making this assumption.

¹⁹ This is Shabo’s attempt to capture van Inwagen’s notion of ‘having a choice’ (Shabo, “Free Will and Mystery: Looking Past the *Mind* Argument”, p. 7).

²⁰ The ability to \square might simply be the ability not to \square .

²¹ In “Why Free Will Remains a Mystery”, p. 121 n. 4, Shabo points out the possibility of interpreting the power to settle in this fashion. I offered this interpretation of ‘having a choice’ in Franklin, “Farewell to the Luck (and *Mind*) Argument”, p. 227. van Inwagen endorses this analysis of ‘having a choice’ in his “A Promising Argument”, in *The Oxford Handbook of Free Will*, 2nd edition, ed. Robert Kane (New York: Oxford University Press, 2011: 475-483). In “Free Will and Mystery: Looking Past the *Mind* Argument”, p. 11 n. 14, Shabo maintains that I claim in “Farewell to the Luck (and *Mind*) Argument” that ‘having a choice about which choice one makes’ is incoherent. But I never claim this. I only claim that I cannot make sense of ‘having a choice about which choice one makes’ *unless* we understand it merely to mean that the agent had the ability to the make the choice and the ability not to make the choice (“Farewell to the Luck (and *Mind*) Argument”, p. 227).

²² This concludes Shabo’s presentation of the Rollback Argument, but I do not think it exhausts van Inwagen’s argument. Indeed, as I argue below, it is hard to see how this is even an *argument*. That van Inwagen did not intend this thought experiment to stand alone gains support from the fact he further defended his claim that no one has the ability to determine the outcome of an indeterministic process. Just after presenting this rollback scenario, he says he will offer ‘an example designed to convince us’ that no one has the ability to choose A if it is a matter of chance whether one will choose A or B (“Free Will Remains a Mystery”, p. 17). Here is the example: Consider a situation in which the friend of a candidate running for public office knows some discreditable fact that, if revealed, will cost the candidate the election. The candidate asks his friend to promise not to reveal this fact. Suppose the friend knows (imagine God told him) that

there is an objective, ground-floor probability of 0.43 of his revealing the discreditable fact to the public, and an objective, ground-floor probability of 0.57 that he will remain silent. Is the friend in a position to promise the candidate that he will not reveal the fact? It seems not, van Inwagen contends, since of a million perfect duplicates of the friend, placed in exactly the same situation, 43% will reveal the secret and 57% will remain silent. ‘But’, van Inwagen goes on, ‘if [the friend believes] that [he is] able to keep silent, [he] should, it would seem, regard [himself] as being in a position to make this promise. What more [does he] need to regard [himself] as being in a position to promise to do X than a belief that [he is] able to do X? Therefore, in this situation, [he] should not regard [himself] as being able to keep silent’ (“Free Will Remains a Mystery”, p. 17).

In “Farewell to the Luck (and *Mind* Argument”, pp. 217-219, I argued that this argument fails because it assumes something libertarianism is not committed to: namely that indeterminism occurs *after* the agent’s decision. The promising case assumes that the friend cannot now do anything about a later undetermined action. For example, the friend’s deciding to keep the promise will not affect the later probabilities of whether he will in fact keep the promise. But libertarians are not committed to this view of indeterminism. In “A Promising Argument”, p. 477, van Inwagen concedes that his promising argument fails for this reason.

Shabo is not alone in overlooking this important part of the Rollback Argument (see also Laura Waddell Ekstrom, “Free Will, Chance, and Mystery”, *Philosophical Studies* 113 (2003): pp. 153-80). I will follow Shabo’s presentation of the Rollback Argument, assuming that this argument consists of no more than is presented above. Hence, it is important to bear in mind that my critique of the Rollback Argument will only be a critique of the Rollback Argument as presented in the text.

²³ It is unclear to me why Shabo thinks this. van Inwagen never makes such a distinction and I find it somewhat hard to believe that he meant us to have in mind this more sophisticated ability without ever flagging our attention to it.

²⁴ Shabo assumes, ‘for ease of exposition’, that mental states are token identical to neural states (cf. Shabo, “Why Free Will Remains a Mystery”, pp. 114 and 123 n. 19), but encourages the reader to substitute ‘is constituted by’, ‘supervenes on’ or ‘is correlated with’ for ‘is identical to’. I will follow Shabo in assuming token-identity physicalism.

²⁵ van Inwagen’s second instance of the *Mind* Argument (*An Essay on Free Will*, pp. 129-142, esp. pp. 131-134) bears interesting similarities to Shabo’s argument. van Inwagen invites us to imagine a ‘freakish demon’ who has connected an invisible wire to our thief’s brain. The other end of the wire is connected to a piano and by playing the piano in certain ways the demon can control the thief’s inner life. We are then to imagine that the demon is replaced with an indeterministic mechanism in the piano that nondeterministically causes keys to be depressed and, via the wire, sends signals to the thief’s brain, causing corresponding behavioral movements. This case, in which it is obvious that the thief is not free, is modified, step by step, until we arrive at a ‘normal’ instance of the thief seemingly performing an undetermined action. One crucial difference between this argument and Shabo’s is that the former seeks to show that undetermined *action* is impossible. The Assimilation Argument, however, only attacks the possibility of undetermined *free* action. Interestingly, as we will see, both arguments fail for similar reasons: in each argument, we can distinguish the normal case from the other cases by the fact that only in the normal case is the agent’s action caused, in the appropriate way, by agent-involving mental states and events. See van Inwagen, *An Essay on Free Will*, pp. 134-142 for his reply to the second instance of the *Mind* Argument.

²⁶ The presentation of the Assimilation Argument varies somewhat from Shabo, “Why Free Will Remains a Mystery” to Shabo, “Free Will and Mystery: Looking Past the *Mind* Argument”, although these differences appear immaterial. I have elected to follow more closely the presentation of the argument in Shabo, “Free Will and Mystery: Looking Past the *Mind* Argument”, but my criticisms apply equally to the argument in Shabo, “Why Free Will Remains a Mystery”.

²⁷ Shabo, “Free Will and Mystery: Looking Past the *Mind* Argument”, p. 12.

²⁸ It is important to note that, given the causal theory of action, two events can be intrinsically indistinguishable and yet only one an action. This possibility arises since (some of) the mark(s) of action are extrinsic, such as its causal history.

²⁹ Shabo, “Why Free Will Remains a Mystery”, p. 106.

³⁰ I am grateful to an anonymous referee for raising this objection.

³¹ In “Free Will and Mystery: Looking Past the *Mind* Argument”, p. 13, Shabo suggests that this is an accurate description of case 3: ‘Alice doesn’t act in Case 3, the idea is, because the device causally determines her decision (or decision-like neural state) based on the behavior of the particle, and completely independent of her reason for making that decision.’

³² Shabo, “Free Will and Mystery: Looking Past the *Mind* Argument”, pp. 13-14. Cf. Shabo, “Why Free Will Remains a Mystery”, pp. 115-116.

³³ Shabo, “Free Will and Mystery: Looking Past the *Mind* Argument”, p. 13.

³⁴ Shabo, “Free Will and Mystery: Looking Past the *Mind* Argument”, p. 11.

³⁵ The two most important recent defenses of the Manipulation Argument are found in Mele, *Free Will and Luck*, pp. 188-195; and Pereboom, *Living without Free Will*, pp. 110-117. Defenses of the Manipulation Argument primarily differ in

how they defend (1), specifically concerning the mechanisms of manipulation that occur in the offered thought experiment.

³⁶ Harry Frankfurt, *The Importance of What We Care About* (New York: Cambridge University Press, 1988: p. 53). Cf. David Shoemaker, “Caring, Identification, and Agency”, *Ethics* 114 (2003): p. 117.

³⁷ My claim here is not that all instances of the Manipulation Argument fail, but merely that this instance of the Manipulation fails.

³⁸ These arguments do not fail merely because there is some difference between cases. The difference must be a free-will-relevant difference. One obvious way in which a difference is a free-will-relevant difference is that in one case a condition necessary for exercising free will obtains and yet this condition does not obtain in the second case.

³⁹ Again, I am assuming that R is not identical to D. If one thinks this assumption is false, my present argument can be modified without loss of force to apply to the differences between cases 3 and 4.

⁴⁰ Presumably, if the device were removed, Alice would regain these abilities. I would prefer to construe the device as robbing Alice of her opportunities to exercise her abilities to make the relevant choices, but as noted above (n. 18), I am here using ‘ability’ in a way that incorporates the notion of opportunity.

⁴¹ Is performing an action \square sufficient for possessing the ability to \square ? Consider someone who has never thrown a dart before and on his first try hits the bull’s-eye, but never succeeds in doing this again, despite numerous tries. It does not strike me as implausible to say that he lacks the ability to hit the bull’s-eye. But even if \square -ing is not sufficient for possessing the ability to \square , it usually is. And given the absence of special circumstances in case 5, we are entitled to infer from Alice’s making the choice that she possessed the ability to make this choice. Moreover, I cannot think of any cases in which it is plausible to say that an agent chooses to \square and yet lacks the ability to choose to \square .

⁴² This is an additional supposition. From the fact that Alice’s choice was undetermined it does not *follow* that there is a possible world with the same laws and past in which she performs a different action. For example, it is possible that Alice’s choice at t was undetermined because she might have died at t . In worlds in which this alternative possibility is realized Alice will not *do* otherwise (although she will be otherwise).

⁴³ Obviously some will dispute these claims. My point is that the Assimilation Argument presents no reason to doubt them.

⁴⁴ Shabo, “Free Will and Mystery: Looking Past the *Mind* Argument”, p. 14.

⁴⁵ Shabo, “Free Will and Mystery: Looking Past the *Mind* Argument”, p. 14.

⁴⁶ Shabo, “Free Will and Mystery: Looking Past the *Mind* Argument”, p. 15.

⁴⁷ As mentioned in n. 22 above, I do not think the argument I am about to consider is the entire Rollback Argument.

⁴⁸ For an important discussion of the Rollback Argument that attempts to show that it turns on an equivocation concerning the notion of ‘chance’, see Ekstrom, “Free Will, Chance, and Mystery”.

⁴⁹ These views are defended respectively in Robert Adams, “Actualism and Thisness”, *Synthese* 49 (1981): pp. 3-41; Alvin Plantinga, *The Nature of Necessity* (New York: Oxford University Press, 1974); and David Lewis, *On the Plurality of Worlds* (Malden: Blackwell, 1986).

⁵⁰ It is important that the monitors only depict *dramatic* representations (similar to how movies can be dramatic representations of actual historical events) of the outcome of Alice’s deliberation in these possible worlds, rather than the literal events themselves. If the monitors depict the literal events in the alternative possible worlds, then there are causal interactions between different possible worlds. But this is metaphysically impossible. However, if the monitors merely show dramatic representations of the events (again similar to how movies can be dramatic representations of actual historical events), then there is nothing metaphysically impossible about this case. I am grateful to an anonymous referee for bringing the need for this qualification to my attention.

⁵¹ Anecdotally, I have witnessed philosophers, after realizing that van Inwagen’s thought experiment is impossible, move from being quite worried about the Rollback Argument to unsure about its force. So perhaps van Inwagen’s argument gains illicit support from his original, problematic formulation.

⁵² I am actually sliding between two claims—namely ‘that there was a 0.5 probability that Alice would choose to tell the truth’ and ‘that Alice’s choice to tell the truth was undetermined’. These claims are not equivalent. An event is undetermined if and only if the chance of its not occurring is more than 0 and less than 1, such as 0.01. However, these claims can be treated as equivalent for present purposes. If the Rollback Argument depends on a particular assignment of probabilities to Alice’s choosing to tell the truth, such as 0.5, then the Rollback Argument is invalid. This argument is supposed to show that *indeterminism* is incompatible with free will and thus it cannot depend on assigning Alice’s choice a particular probability between 0 and 1. Thus, although I am sliding between two claims, it is a slide that the Rollback Argument itself makes.

⁵³ Important recent discussions of the appropriate audience of philosophical arguments are found in Mele, *Free Will and Luck*, especially pp. 190-195; and Peter van Inwagen, *The Problem of Evil* (New York: Oxford University Press, 2006: lecture 3).

⁵⁴ Cf. Gary Watson, “Free Action and Free Will”, *Mind* 96 (1987): pp. 145-172.

⁵⁵ See John Martin Fischer, “The Zygote Argument Remixed”, *Analysis* 71 (2011): pp. 267-272 where he contends that the Manipulation Argument does not even raise the cost of compatibilism. My argument assumes that the Manipulation Argument raises the cost of compatibilism because the objection I am responding to assumes this.

⁵⁶ van Inwagen does have an argument, it is his promising argument (“Free Will Remains a Mystery”, pp. 17-18). It is (partly) for these reasons that I believe it is a mistake to interpret the Rollback Argument as exhausted by the rollback thought experiment. However, adding the promising argument will, in the end, not help since it is invalid. See Franklin, “Farewell to the Luck (and *Mind*) Argument”, pp. 226-228; and van Inwagen, “A Promising Argument”, p. 477).

⁵⁷ I am grateful to Seth Shabo and an anonymous referee and editor for the *Pacific Philosophical Quarterly* for their probing comments on earlier drafts of this paper.