A COMPANION TO FREE WILL

EDITED BY
Joseph Campbell, Kristin M. Mickelson,
and V. Alan White



WILEY Blackwell

19

Libertarianism[1]

MARK BALAGUER

DEPARTMENT OF PHILOSOPHY, CALIFORNIA STATE UNIVERSITY, LOS ANGELES

This chapter is about the libertarian view of free will.[2] I'll start by saying a few words about how libertarianism ought to be defined. Then I'll formulate some classic arguments against libertarianism, and finally, I'll develop a version of event-causal libertarianism and argue (very briefly) that if we adopt this version of the view then we can respond to the arguments against libertarianism.

1 Defining Libertarianism

Libertarianism has traditionally been defined as the conjunction of two theses, namely, (a) the thesis that human beings have free will, and (b) *incompatibilism* – i.e., the thesis that free will is incompatible with determinism. But this definition doesn't do a very good job of capturing everything that libertarians are committed to (and I think it also fails to get at the *bottom-level* commitments of libertarians). The first step toward getting a better definition is to articulate the kind of freedom that libertarians believe in. We can do this as follows:

> A person is *libertarian-free* (or for short, *L-free*) if and only if she makes at least some decisions that are such that (a) they are both undetermined and appropriately non-random, and (b) the indeterminacy is relevant to the appropriate non-randomness in the sense that it *generates* the non-randomness, or *procures* it, or *enhances* it, or *increases* it, or something along these lines.

More needs to be said about what appropriate non-randomness consists in. Different libertarians will say different things about this, but all should agree that it consists in some sort of *agent-involvedness*. For example, one might say that it consists in the agent *controlling* which options is chosen, or *authoring* the choice, or being the *source* of the choice, or making a *rational* choice, or some combination of these things. Also, many libertarians would follow Kane (1985, 1996) in requiring *plural* control (or authorship or whatever) – i.e., in requiring it to be the case that even if the agent had chosen differently, she still would have controlled it or authored it, or whatever.

A Companion to Free Will, First Edition. Edited by Joseph Campbell, Kristin M. Mickelson, and V. Alan White.
© 2023 John Wiley & Sons, Inc. Published 2023 by John Wiley & Sons, Inc.

Downloaded from https://onlinelibrary.wiley.com/doi/ by Mark Balaguer , Wiley Online Library on [12/09/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Also, more needs to be said about clause (b) of the definition of L-freedom. I intend this clause to be read very loosely; the idea is that the indeterminacy has to be relevant to the non-randomness (i.e., to the agent-involvedness) in *some way or other*. Different libertarians can fill this in in different ways. Some might say that indeterminacy is a *necessary condition* for non-randomness; others might deny this and maintain that, in some cases, the fact that a decision is undetermined somehow *causes* that decision to be non-random. I'll say more below about what *I* think libertarians should say about this, but for now, I want to leave it open.

Given the above definition of L-freedom, we can define libertarianism as the conjunction of the following two theses:

> *HB-libertarianism*: Human beings are L-free.
> *C-libertarianism*: Free will is L-freedom.[3,4]

According to this new definition, libertarianism is a stronger view than it is on the traditional definition. To appreciate this, notice first that this new view (i.e., the conjunction of HB-libertarianism and C-libertarianism) clearly entails both conjuncts of the traditional view – i.e., that humans have free will and that incompatibilism is true. But the new view commits to much more than the traditional view does. According to the traditional definition, libertarians are committed to the thesis that human beings have an indeterministic kind of free will. But L-freedom isn't just indeterministic; there are further requirements that need to be met for an indeterministic kind of free will to count as a kind of L-freedom; in particular, the indeterminacy needs to generate (or procure, or whatever) an appropriate sort of non-randomness, or agent-involvedness.

Given that my definition takes libertarianism to be a stronger view than the traditional definition does, I should say a few words to motivate the idea that libertarians really are committed to what I'm saying they're committed to. The first point to note here is that libertarians are clearly committed to the thesis that human beings have a kind of free will that involves appropriate non-randomness, or agent-involvedness. If all of my decisions are controlled by Martians via remote control, then I don't have free will in the libertarian sense of the term (or any other sense, for that matter), and this is true even if there are indeterminacies in the causal pathways from the Martians' manipulations to my decisions. So indeterminacy is clearly not enough for libertarian-freedom; it also needs to be the case that our decisions are appropriately non-random – i.e., they need to be made *by us*. Second, it's also clear that libertarians are committed to the idea that the non-randomness (or agent-involvedness) has to be generated (or procured, or enhanced, or whatever) by the indeterminacy. Without this, the following kind of freedom would seem to count as a kind of libertarian-freedom:

> A person is *virtually-Hume-free* iff (i) there are insignificant indeterminacies in her decision-making processes—i.e., indeterminacies that almost never have any effect on which options are chosen; and (ii) her decisions are "for-all-practical-purposes determined" by her desires—i.e., they're probabilistically caused by her desires with a high degree of certainty (in particular, the objective moment-of-choice probabilities of her choices are always at least 0.9999999).

If we're virtually Hume-free, then our decisions are undetermined, and they also seem to be non-random in an agent-involving way (because they're "virtually determined" by our desires, and so they *come from us* in an obvious way[5]). But virtual-Hume-freedom is clearly not a *libertarian* sort of freedom, and the reason is that the indeterminacy here is entirely irrelevant to the non-randomness. Indeed, if virtual-Hume-free decisions are non-random – i.e., if they're made *by us* – then this is *in spite* of the fact that they're undetermined. Libertarians

Downloaded from https://onlinelibrary.wiley.com/doi/ by Mark Balaguer , Wiley Online Library on [12/09/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

are clearly committed to something stronger than this; they're committed to the idea that the indeterminacy is *relevant* in some way to the non-randomness – i.e., to the idea that the indeterminacy generates the non-randomness, or procures it, or enhances it, or some such thing.

So I think that my definition of libertarianism is better than the traditional definition because it's more informative – i.e., it does a better job of capturing what libertarians are actually committed to. Moreover, my definition also does a better job of capturing the *bottom-level* commitments of libertarians. To appreciate this, notice that while the two conjuncts of the traditional definition are not independent of one another, the two conjuncts of my definition *are* independent of one another. One of the conjuncts (HB-libertarianism) is a contingent claim about the nature of the actual world, in particular, human decision-making processes; and the other conjunct (C-libertarianism) is a *conceptual* claim – it's an answer to the conceptual-analysis question 'What is free will?' But the answer to the conceptual-analysis question is clearly not relevant to the question of whether HB-libertarianism is true because the term 'free will' doesn't appear at all in the definition of HB-libertarianism. Moreover, *prima facie*, it seems that the answer to the contingent, actual-world question of whether HB-libertarianism is true isn't relevant to the conceptual-analysis question either. Now, you might try to resist this last claim by arguing that 'free will' should be defined as the freedom-like ability that humans actually have – whatever that turns out to be. But I think it can be argued that (a) this is in fact *not* the right way to define 'free will,' and (b) even if it is, the question of whether HB-libertarianism is true is still not relevant to the conceptual-analysis question 'What is free will?'[6]

2 Zeroing in on HB-Libertarianism

The question of whether libertarianism is true reduces to the following two questions:

The HB-question: Is HB-libertarianism true?
The C-question: Is C-libertarianism true?

These are both difficult, multi-faceted questions, and I don't have the space to do justice to both of them here. Thus, in what follows, I'm going to focus on the HB-question and ignore the C-question. I do this because I find the HB-question much more interesting. The reason I'm less interested in the C-question is that I think the answer to this question is (a) entirely settled by empirical facts about the usage and intentions of ordinary folk concerning words like 'free,' 'can,' 'could have,' etc., and (b) not relevant in any non-trivial way to metaphysical questions about the nature of human decision-making processes. Both of these claims are extremely controversial; I can't give the complete argument for them here, but I'd like to say a few words to clarify my view.[7]

The first point to notice here is that there are multiple kinds, or concepts, of freedom. For instance, there's L-freedom (defined above); and Hume-freedom (i.e., roughly, the ability to do what you want, or to act and choose in accordance with your desires); and Frankfurt-freedom (i.e., roughly, the ability to control, with second-order attitudes, which of your first-order desires will affect your behavior); and so on. (There are also multiple kinds, or concepts, of moral responsibility; e.g., there's L-responsibility (which requires L-freedom), and Hume-responsibility (which requires Hume-freedom), and Frankfurt-responsibility (which requires Frankfurt-freedom), and so on.) Now, in themselves, these are all perfectly good concepts. If you have Platonist leanings, you can think of them as existing "side-by-side" in Platonic heaven; if you don't have such leanings, you can tell whatever side-by-side story you

Downloaded from https://onlinelibrary.wiley.com/doi/ by Mark Balaguer , Wiley Online Library on [12/09/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

like. Either way, the point is that these are all perfectly good concepts that a community of rational, moral creatures could usefully employ.

(It's important to note that I'm not yet saying anything controversial. By merely claiming that these concepts exist, and that we can introduce terms of art that express them, all I'm really saying is that people can introduce new terms and define them however they like. For instance, if we wanted to, we could introduce the term 'one-footed freedom' to denote the ability to make a decision while standing on one foot. One-footed freedom isn't a very interesting kind of freedom, and it's clearly not the kind of freedom that's picked out by the English expression 'free will,' but there's nothing *wrong* with it, as a concept. So when I say that L-freedom and Hume-freedom and so on are "perfectly good concepts that a community of rational creatures could employ," I don't mean to suggest that any of these concepts captures the correct definition of the English expression 'free will.' Moreover, I also don't mean to suggest that these concepts are all instantiated, or even that they *could* be instantiated. As we'll see in Section 4, some people think that L-freedom is *incoherent* and that it's *impossible* for a person to be L-free. But even if this were true, it wouldn't follow that there was no such thing as the concept of L-freedom. Indeed, the claim that L-freedom is incoherent straightforwardly *depends on* the claim that there *is* a concept of L-freedom; for if there were no such concept, then there wouldn't be anything there to be incoherent, and the sentence 'L-freedom is incoherent' wouldn't be true – indeed, it would be gibberish.[8])

In any event, given that we've got all of these concepts to work with, when philosophers ask the conceptual-analysis question 'What is free will?,' we can understand them as asking this: Which of the various concepts of freedom – or if you'd rather, which of the various *freedom-like* concepts, e.g., L-freedom, Hume-freedom, etc. – is identical to the concept of free will? In other words, they're asking which of these concepts provides the correct analysis of the concept of free will, or the correct definition of 'free will.' (Likewise, when philosophers ask what moral responsibility is, they're asking which of the various concepts of responsibility – or which of the various *responsibility-like* concepts – is identical to the concept of moral responsibility; i.e., they're asking which of these concepts provides the correct analysis of the concept of moral responsibility.) But what sorts of questions are these? What kinds of facts could determine the answers to these questions? After all, as I just pointed out, all of the concepts of freedom and responsibility are, in themselves, perfectly good. And it's not as if some of them are glowing with a special hue in Platonic heaven. So what could make two of these concepts stand out as *the* concepts of free will and moral responsibility?

One view here is that the relevant facts – the facts that determine which concepts count as the concepts of free will and moral responsibility – are facts about *us*. In particular, on this view, the concept of free will is just the concept that *we have in mind* when we think and talk about free will, and the concept of moral responsibility is the concept that we have in mind when we think and talk about moral responsibility. I think there are good reasons to think that (a) this view of conceptual analysis is true, and (b) it implies that the C-question is not relevant in any non-trivial way to questions about the nature of human decision-making processes.[9] Now, of course, both of these claims require argument, and unfortunately, I can't say any more about this topic here, but this at least clarifies why I think the C-question is less interesting than the HB-question.

3 Arguments for HB-Libertarianism?

I don't think there are any good arguments for HB-libertarianism. Some might think that the fact that we have an "intuition" that we're free gives us good reason to believe that we're L-free, but

Downloaded from https://onlinelibrary.wiley.com/doi/ by Mark Balaguer , Wiley Online Library on [12/09/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

this strikes me as highly immodest. HB-libertarianism requires the truth of indeterminism, and indeterminism involves a substantive, controversial thesis of subatomic physics. The idea that our intuitions could track the truth about this seems to me about as plausible as the idea that our intuitions could track the truth about the number of planets that are orbiting Alpha Centauri.

One might also try to argue for HB-libertarianism from the following two claims: (i) we're morally responsible for our actions, and (ii) moral responsibility requires L-freedom.[10] But the idea that moral considerations can give us good reason to believe a controversial thesis of sub-atomic physics is no more plausible than the idea that our intuitions can give us good reason to believe such a thesis. To say a bit more about this, I think the problem with the argument from (i) and (ii) is that while both of these premises can seem plausible when considered in isolation, it's not plausible to suppose that we currently have good reason to believe that they're *both* true. In particular, if we had good reason to believe thesis (ii), that would immediately give us good reason to think that we had no reason to believe thesis (i), for in this scenario, the truth of (i) would depend on a controversial thesis of subatomic physics.

4 Arguments Against HB-Libertarianism

Prima facie, the arguments *against* HB-libertarianism seem much stronger than the arguments in favor of that view. Here are three arguments against HB-libertarianism that, I think, a lot of people find convincing:

(1) *The argument from determinism*: HB-libertarianism is false because it requires a certain sort of indeterminism, and there are good empirical reasons to think that this sort of indeterminism isn't true.

(2) *The luck argument (aka, the Mind argument, or the argument from randomness, or the Hobbes-Hume-Hobart argument)*: Even if the relevant sort of indeterminism is true, it doesn't matter because there's a clear sense in which undetermined events are *random*, or *lucky*. In other words, they occur by *chance* – i.e., they *just happen*. Thus, if we introduce an undetermined event into a decision-making process, that would seem to either (a) increase the level of randomness (or luckiness) in that process or (b) leave the level of randomness (or luckiness) alone (if the indeterminacy ends up not mattering). So it's hard to see how the introduction of an undetermined event into a decision-making process could increase *non-random*ness (or non-luckiness, or agent-involvedness, or whatever). Thus, since this is precisely what HB-libertarianism requires, it seems that HB-libertarianism is false and, indeed, arguably incoherent.

(3) *The argument from naturalism*: HB-libertarianism is incompatible with a naturalistic, scientifically reputable view of the world. In particular, it seems to be incompatible with a materialistic view. Therefore, since materialism is true, HB-libertarianism is false. (This argument is almost never discussed in the literature, but it's, so to speak, "in the air" – and, I think, in lots of people's heads.)

Among these three arguments, the luck argument has received the most attention, especially recently. Numerous versions of this argument have appeared in the literature, spread out over hundreds of years, under many different names.[11] There are also many responses to this argument.[12] In fact, I think it's fair to say that the various different versions of libertarianism have been developed in connection with attempts to respond to the luck argument. The most

Downloaded from https://onlinelibrary.wiley.com/doi/ by Mark Balaguer , Wiley Online Library on [12/09/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

important distinction here is between *agent-causal* views, *event-causal* views, and *non-causal* views. Agent-causal libertarians maintain that when humans make L-free decisions, they agent-cause these decisions to occur, where agent-causation is a kind of causation that obtains not between two events (as is the case with event-causation) but between an agent and an event (and according to agent-causal views, agent-causation is not reducible to event-causation).[13] Event-causal libertarians, on the other hand, maintain that all causation reduces to ordinary event-causation and that our L-free decisions are probabilistically caused by agent-involving events.[14] And non-causal libertarians maintain that our L-free decisions are uncaused.[15] These three views have been developed, by and large, as ways of getting around the luck objection. I don't have the space to discuss all of them here, so I will focus on event-causal libertarianism. Indeed, I'll focus on one specific version of event-causalism, namely, the one that I think provides the right response to the luck argument.[16]

I'll spend most of the rest of this chapter responding to the luck argument, but at the end, I'll say a few words about the arguments from determinism and naturalism. Indeed, my response to the luck argument will bring with it a response to the argument from naturalism; for the version of libertarianism that I'll be developing here is entirely materialistic and scientifically reputable. In particular, I'll be assuming that human decisions are *neural* events.

(In addition to being materialistic and event-causal, the version of libertarianism I'll be developing here is *centered* – i.e., it places the important indeterminacy at the moment of choice. This is in contrast to *deliberative* libertarianism, which places the important indeterminacy *prior* to the moment of choice. Deliberativism has been developed by Dennett (1978) and Mele (1995), but to the best of my knowledge, no one has ever *endorsed* this view.[17] This, I think, is for good reason; I think there are serious problems with deliberativism.)

5 Responding to the Luck Argument

5.1 Torn Decisions

In order to respond to the luck argument, it's important for libertarians to focus on decisions of the right kind – namely, what I have elsewhere called *torn decisions*. The notion of a torn decision can be defined as follows:

> A *torn decision* is a decision in which the person in question has reasons for multiple options, feels torn as to which option is best (and has no conscious belief as to which option is best), and decides without resolving the conflict, i.e., decides while feeling torn.

I think we make decisions like this several times a day about things like whether to have eggs or cereal for breakfast, or whether to work out before leaving for work, or whatever. But we can also make torn decisions in connection with big life-changing decisions; e.g., you might have a good job offer in a city you don't like, and you might have a deadline that forces you to decide while feeling utterly torn.

Torn decisions should be distinguished from three other kinds of decisions. First, they should be distinguished from *leaning decisions*; these are decisions in which the agent chooses while leaning toward one of her live options, whereas in a torn decision, the agent feels *completely* torn. Second, torn decisions should be distinguished from *Buridan's-ass decisions*; these are similar to torn decisions except that the various tied-for-best options are more or less indistinguishable, and because of this, the agent *doesn't feel torn*. (For instance,

Downloaded from https://onlinelibrary.wiley.com/doi/ by Mark Balaguer , Wiley Online Library on [12/09/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

if you want a can of tomato soup, and there are ten cans of the same kind on the shelf, you won't feel torn – you'll just grab one and be on your way.[18]) Third, torn decisions should be distinguished from what Kane (1996) calls *self-forming actions*, or SFAs. The most important difference here is that, unlike SFAs, torn decisions are not defined as being undetermined. They're defined in terms of their phenomenology. Thus, we know from experience that we do make torn decisions, and it's an empirical question whether any of these decisions are undetermined.

My reason for thinking that libertarians should focus on torn decisions is not that I think that decisions of these other kinds can't be free, or L-free. My reason is just that if we focus on torn decisions, then, as we'll see in what follows, we can get a very clear response to the luck argument.

5.2 *TDW-indeterminism*

The next thing we need to do, in order to get a response to the luck argument, is to get clear on the relevant kind of indeterminacy – i.e., the kind of indeterminacy that needs to be present in torn decisions in order for decisions of this kind to be L-free. This sort of indeterminacy can be defined as follows:

> A torn decision is *wholly undetermined* at the moment of choice—or, for short, *TDW-undetermined*—iff the actual objective moment-of-choice probabilities of the various reasons-based tied-for-best options being chosen match the reasons-based probabilities (or the phenomenological probabilities), so that these moment-of-choice probabilities are all roughly even, given the complete state of the world and all the laws of nature, and the choice occurs without any further causal input, i.e., without anything else being significantly causally relevant to which option is chosen.

It's important to note that this sort of indeterminacy is compatible with various features of the decision being fully determined. Suppose, e.g., that I'm about to make a torn decision between options A and B. It could be determined that (i) I'm going to make a torn decision (i.e., I'm not going to refrain from choosing); and it could be determined that (ii) I'm going to choose between A and B (i.e., I'm not going to choose some third option that I don't like as much); and it could be determined that (iii) the moment-of-choice probabilities of A and B being chosen are both 0.5 (i.e., they're not 0.7 and 0.3, or 1 and 0, or anything else). All of this is perfectly consistent with the decision being TDW-undetermined. All that needs to be undetermined, in order for the choice to be TDW-undetermined, is *which tied-for-best option is chosen*.

It's also important to note that TDW-indeterminacy lies at one end of a spectrum of possible cases and that there are *degrees* of the kind of indeterminacy I'm talking about here. To see what I've got in mind by this, suppose that Ralph makes a torn decision to order chocolate pie instead of apple pie. Since this is a torn decision, we know that given all of Ralph's conscious reasons and thought, he feels completely neutral between his two tied-for-best options. But it could be that factors external to Ralph's conscious reasons and thought (e.g., unconscious compulsions, or wholly non-mental brain events that precede the decision in his head) causally influence the choice and wholly or partially determine which option is chosen. Indeed, there's a continuum of possibilities here. At one end of the spectrum, which option is chosen is TDW-undetermined, so the moment-of-choice probabilities of the two tied-for-best options being chosen are 0.5 and 0.5, and nothing else significantly causally influences which option is chosen. At the other end of the spectrum, the choice is fully

Downloaded from https://onlinelibrary.wiley.com/doi/ by Mark Balaguer , Wiley Online Library on [12/09/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

determined – i.e., factors external to Ralph's conscious reason and thought come in and, unbeknownst to Ralph, cause him to choose chocolate. And in between, there are possible cases where the moment-of-choice probabilities are neither 0.5 and 0.5 nor 1 and 0 – i.e., where they're 0.8 and 0.2, or 0.7 and 0.3, or whatever; in these cases, external factors causally influence the choice without fully determining it. In cases like this, we can say that which option is chosen is *partially determined and partially undetermined*.[19]

5.3 *The Central Libertarian Thesis*

Let *TDW-indeterminism* be the view that some of our torn decisions are TDW-undetermined. Given this, I think that libertarians can respond to the luck argument by arguing for the following thesis:

> *Central Libertarian Thesis (CLT)*: If our torn decisions are undetermined in the right way—i.e., if they're TDW-undetermined—then they're appropriately non-random and L-free. Or more succinctly: *If TDW-indeterminism is true, then HB-libertarianism is true.*

Notice that CLT doesn't just say that our torn decisions *could be* both undetermined and non-random; it says that if they're undetermined in the right way, then they *are* appropriately non-random and L-free, so that the question of whether our torn decisions are L-free just reduces to the question of whether they're undetermined in the right way. If this is right, then it doesn't just refute the luck argument – it turns that argument completely on its head. For whereas the luck argument suggests that indeterminacy leads to randomness, CLT says the exact opposite – that the right kind of indeterminacy leads to *non-random*ness. This is much stronger than what other libertarians (e.g., Ginet (1990) and Kane (1996)) have tried to establish in responding to the luck argument; they've tried to show that indeterminacy doesn't necessarily lead to non-randomness – i.e., that there are ways in which a decision could be undetermined and yet still be appropriately non-random. But, again, CLT says that if our torn decisions are TDW-undetermined, then they *in fact are* appropriately non-random and L-free. If this is right, then it gives us the surprising result that we can establish libertarianism – or at any rate, HB-libertarianism – by arguing for an *empirical* thesis, namely, TDW-indeterminism.

(CLT is only about *full-blown* L-freedom. I think it can also be argued that if an ordinary torn decision is partially undetermined in the manner described in Section 5.2, then it's *partially L-free*. But I won't worry about this here; I'll just focus on arguing that if our torn decisions are *wholly* undetermined – i.e., TDW-undetermined – then they're *fully* L-free.)

5.4 *An Argument for CLT*

I argued at length for CLT in my (2010). I will here just provide a brief formulation of one of the arguments. To get started, let's return to Ralph's decision to order chocolate pie instead of apple pie, and let's suppose that he was completely torn when he made this decision. Given this, we get the following result:

> (A) Ralph's choice was conscious, intentional, and purposeful, with an actish phenomenology—in short, it was a Ralph-consciously-choosing event, or a Ralph-consciously-doing event.

If we now assume that Ralph's decision was TDW-undetermined, then we also get the following two results:

Downloaded from https://onlinelibrary.wiley.com/doi/ by Mark Balaguer , Wiley Online Library on [12/09/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

(B) Ralph's choice flowed out of his conscious reasons and thought in a nondeterministically event-causal way.

(C) Nothing external to Ralph's conscious reasons and thought had any significant causal influence (at the moment of choice—i.e., after he moved into a torn state and was about to make his decision) over how he chose, so that the conscious choice *itself* was the event that settled which option was chosen.

My argument for CLT is based on the observation that (A)-(C) seem to imply that Ralph authored and controlled his decision. This is because (A)-(C) seem to give us the two-fold result that *(i) Ralph did it, and (ii) nothing made him do it*; and, intuitively, it seems that (i) and (ii) are sufficient for authorship and control. The Ralph-did-it-ness comes from (A)-(B) – the choice flowed out of Ralph's reasons, and the event that actually settled which option was chosen was a conscious Ralph-choosing event. And the fact that nothing made him do it comes from (C).

(It's important to note that I'm not engaged here in conceptual analysis. I'm not putting forward a theory of what authorship and control *are*. It might seem that when I say that (i) and (ii) are sufficient for authorship and control, I'm making a claim of conceptual analysis; but as will become clear below, all I'm committed to saying here is that in connection with ordinary torn decisions like Ralph's, if they're TDW-undetermined, then they're authored and controlled by the agent *in certain interesting and important ways*. I don't need to claim that the kinds of authorship and control that are present here are the only kinds of authorship and control that one might care about. And I don't need to say anything about the nature of authorship and control. More on this below.)

In any event, to get the result that Ralph's decision was appropriately non-random and L-free, it's not enough to argue that he authored and controlled his decision; we also need to argue that (i) his decision satisfied all of the other conditions for appropriate non-randomness, aside from authorship and control (e.g., rationality, the Kanean plurality conditions, and so on), and (ii) the fact that Ralph's decision was TDW-undetermined *procured* the result that it was appropriately non-random.

I don't have the space to say very much in support of these two claims here, but given the above argument about authorship and control, they're both relatively uncontroversial. For instance, regarding (i), it should be clear that Ralph's choice was extremely rational; if you have to choose between two options, and if you're completely unsure which option is best, and if not choosing (or continuing to deliberate) is worse than choosing arbitrarily, then the most rational thing to do is to arbitrarily pick one of your tied-for-best options. Moreover, given that Ralph was *torn*, it also seems clear that his decision satisfied the Kanean plurality conditions; for if he'd chosen to order apple pie instead of chocolate pie, we would have been able to run the same argument about authorship, control, rationality, and so on.

As for point (ii), this should already be clear from the argument about authorship and control. For the fact that Ralph's decision was TDW-undetermined played a crucial role in the argument for the claim that he authored and controlled the decision—this is what gave us the result that nothing made Ralph choose chocolate pie over apple pie. Moreover, as we'll shortly see, the fact that Ralph's decision was TDW-undetermined also gives us the result that the event that actually settled which option was chosen was the conscious decision *itself*. (The idea here isn't that TDW-indeterminacy actively *generates* authorship and control; the idea is that it *blocks a destroyer* of authorship and control. The destroyer would be a moment-of-choice causal influence from something external to Ralph's conscious reasons and thought.

Downloaded from https://onlinelibrary.wiley.com/doi/ by Mark Balaguer , Wiley Online Library on [12/09/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TDW-indeterminacy *guarantees the absence* of such a destroyer, and this is the sense in which it procures authorship and control.)

If all of this is right, then we have an argument for CLT – although, I admit, it's really just a sketch of an argument. The key to the argument, in my opinion, is that it gives us the following result:

> *Choice-Settles-It*: The event that settled which option was chosen *was* the conscious choosing event.

You might object to this in two different ways. First, you might claim that if we're assuming a materialistic view of the mind-brain, then the event that settled which option was chosen was a *neural* event. And second, you might claim that which option was chosen was settled by some *quantum* events. I agree with both of these claims, but I don't think that either of them undermines Choice-Settles-It. For (a) I think that on any reasonable version of mind-brain materialism, the conscious choosing event *is* a neural event;[20] and (b) I think that the neural event in question is *composed* of quantum events, so that if Ralph's decision was TDW-undetermined, then which option was chosen was settled by the quantum events that *made it up*. (If the decision was settled by *prior-to-choice* quantum events, then it wouldn't be correct to say that it was settled by the conscious choice *itself*; but TDW-indeterminism rules out the possibility of this prior-to-choice determination; if the decision was TDW-undetermined, then it was settled by quantum events that were *parts of* the decision, not by quantum events that preceded the decision.)

The argument I've been rehearsing here differs from other libertarian arguments in the literature in numerous ways, but perhaps the most important difference has to do with how modest its assumptions are. Other libertarians have tended to respond to the luck argument by constructing complicated libertarian models and arguing that for all we know, these elaborate models could be true. I've taken the exact opposite approach here. I've argued that by assuming *only* that our torn decisions are TDW-undetermined, we get the result that the undetermined physical events that settle the outcomes of our torn decisions *are the decisions themselves* – i.e., the conscious mental events with a me-choosing-now phenomenology. And this, I think, gives us a clear sense in which the outcomes of these decisions are settled by *us*.

5.5 *An Objection*

The argument for CLT that I just articulated was obviously very quick, and there are a number of different worries that one might have about it. In this section, I'll say a few words about one of these objections, namely, the following:

> *The Agent-Causal Objection*: The notion of control that you're working with is too weak. Something more substantive is needed for control. In particular, it seems that something like agent causation is needed. In other words, when someone makes a torn decision, in order for it to be the case that the agent in question controls which option is chosen, it needs to be that she *causes* the given option to be chosen. But if Ralph's decision was TDW-undetermined, then he didn't cause chocolate pie to be chosen; in fact, in this scenario, *nothing* caused chocolate pie to be chosen. Or at any rate, nothing deterministically caused it to be chosen. We can say that Ralph's reasons probabilistically caused chocolate pie to be chosen, but there seems to be a clear sense in which nothing caused the choice to be a *chocolate-rather-than-apple* choice. And it seems in particular that *Ralph* didn't cause this. And so we can't correctly say that he *controlled* which option was chosen. (Pereboom raises a worry like this about my view in his (2014).)

Downloaded from https://onlinelibrary.wiley.com/doi/ by Mark Balaguer , Wiley Online Library on [12/09/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

One way to respond to this objection would be to argue that agent causation is in fact *not* required for authorship and control. I say a few words to motivate this stance in my (2014b), but the point I want to make here is that in the present context, it doesn't matter whether agent causation is required for authorship and control. To bring this out, let me distinguish two different kinds of control – *causal-control* and *noncausal-control* – where the former requires agent causation (or something like it) and the latter doesn't. I won't try to give precise definitions of these two notions; all I'll say (and all that will matter here) is that when I use the term 'noncausal-control,' I'm talking about a kind of control that applies to ordinary torn decisions if they're TDW-undetermined; i.e., it applies to torn decisions like Ralph's, where the agent makes a conscious decision with an actish phenomenology, and which option is chosen isn't significantly causally influenced (at the moment of choice) by anything external to the agent's conscious reasons and thought, so that the conscious choice itself *is* the event that settles which option is chosen. Beyond this (and beyond the fact that causal-control requires agent causation and noncausal-control doesn't), it won't matter how exactly these two kinds of control are defined. But for the sake of argument, let's pretend that we've got two precisely defined kinds of control here. Given this, one question we might ask is the following:

> *The what-is-control question*: What is control? In other words, which of the various kinds of control that we find in the literature provide correct analyses of the concept of control? Does causal-control? Does noncausal-control? Do both? Do neither?

But why should libertarians care about this question? They don't need to claim that if our torn decisions are TDW-undetermined, then they're authored and controlled by us and L-free in the only senses of these terms that anyone might care about, or in the senses of these terms that philosophers have traditionally cared about. All they need is this:

> (W) If our torn decisions are TDW-undetermined, then they're authored and controlled by us and appropriately non-random and L-free in interesting and important ways that are worth wanting and worth arguing for and that libertarians can hold up and say, "This gives us a noteworthy kind of libertarian free will."

Now, don't take me to be saying more than I am here. I don't think that libertarians can just define authorship and control and L-freedom in ridiculously weak ways and then claim victory. Or to put the point differently, while I don't need to argue that the kind of L-freedom I've articulated – the kind that we get if our torn decisions are TDW-undetermined (i.e., the kind that involves noncausal-control) – is the one and only kind of L-freedom that anyone might care about, I *do* need it to be the case that this kind of L-freedom is interesting, worth wanting, worth arguing for, and so on. In other words, I need (W).

But I think the above argument for CLT does motivate (W). Let's return to Ralph's decision. If it was TDW-undetermined, then (a) the choice was conscious, intentional, and purposeful, with an actish phenomenology; and (b) it flowed out of Ralph's conscious reasons and thought in a nondeterministically event-causal way; and (c) nothing external to Ralph's conscious reasons and thought had any significant causal influence (after he moved into a torn state and was about to make his decision) over how he chose. This might not give us every kind of L-freedom you might have wanted, but it clearly gives us *one important kind* of L-freedom – a kind that libertarians can hang their hats on and that's worth wanting and arguing for and so on. After all, in this scenario, the event that settles which option is chosen *is* the conscious decision – i.e., it's the event with a me-consciously-choosing-now phenomenology.[21]

Downloaded from https://onlinelibrary.wiley.com/doi/ by Mark Balaguer , Wiley Online Library on [12/09/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Now, you might think that this misses the point of Pereboom's argument; for you might think that his claim is that the sort of L-freedom that I'm talking about here is insufficient for moral responsibility and basic desert.[22] But insofar as moral responsibility and desert are interdefinable with freedom, it should be clear that the sort of L-freedom that I've characterized here is definitely sufficient for a certain *kind* of moral responsibility and desert. We can call them L-responsibility and L-desert, where these are defined in a way that makes it analytic that the sort of freedom they require is the sort of L-freedom that I've been describing here. Now, it is of course true that these are not the only kinds of moral responsibility and desert that one might care about; in particular, we can introduce a notion of *causal-responsibility* and stipulate that it's a kind of moral responsibility that requires causal-control. Moreover, it may be that causal-responsibility is worth wanting and that we don't have it. But so what? That wouldn't change the fact that if we have the sort of L-freedom that I've been talking about, then we also have a libertarian sort of moral responsibility, namely, L-responsibility. Now, you might think that the worry here – the worry that Pereboom is getting at – is that causal-responsibility is *real* moral responsibility. In other words, you might worry that the sort of responsibility that we get from L-freedom – namely, L-responsibility – isn't *real* moral responsibility at all. But this takes us back to the point I was making in Section 2. On my view, the facts that determine which of the various kinds of responsibility count as "real" moral responsibility are facts about the *folk*; in particular, the kind(s) of responsibility that count as real moral responsibility are just the one(s) that the folk have in mind when they think and talk about moral responsibility. If this metaphilosophical view of conceptual-analysis questions is right – and as I pointed out in Section 2, I'm aware that it requires argument – then the question of whether L-responsibility is real moral responsibility is considerably less interesting and important than it might at first seem. Indeed, it's just an empirical question about what the folk mean by 'moral responsibility.'[23] Now, I admit that if L-responsibility wasn't worth wanting at all, then that would be a problem; but it seems to me that the considerations that suggest that L-freedom is worth wanting suggest that L-responsibility is worth wanting as well.

You might think that Pereboom's point is not that causal-responsibility is "real" responsibility, but rather that this is the sort of responsibility that *justifies punishment*.[24] But the same kinds of points can be made here all over again. It's obvious that (a) causal-responsibility *causal-justifies* punishment and (b) L-responsibility *L-justifies* punishment but doesn't causal-justify punishment. This leaves open whether L-responsibility *really* justifies punishment, but once again, this is just a question about what the term 'justifies' means in ordinary moral discourse. There is no way to pop out of this circle. *All* of our moral terms can be understood in ways that fit with the kind of L-freedom that I'm talking about – i.e., the kind that we get from TDW-indeterminism – and, of course, these terms can also be understood in ways that fit with agent-causal freedom. Now, my own view is that causal-responsibility and causal-justification are *not* "real" moral responsibility and justification; in other words, I don't think it's built into our ordinary moral practices that agent-causal freedom is required for moral responsibility and the moral justification of punishment. But the point I'm making here is that this doesn't matter; for insofar as L-responsibility and so on are worth wanting, the kind of L-freedom that I'm describing – the kind that we get from TDW-indeterminism – is a worthwhile kind of libertarian free will.

5.6 *An Alternate Version of the Luck Argument*

As I pointed out above, there are many different versions of the luck argument. In this section, I'll say a few words about one more of these arguments, namely, the following:

Downloaded from https://onlinelibrary.wiley.com/doi/ by Mark Balaguer , Wiley Online Library on [12/09/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

> *The Rollback Version of the Luck Argument*: Let's return to Ralph's decision, and let's imagine that God "rolls back" the universe and "replays" the decision 100 times. If the decision is TDW-undetermined, then we should expect that Ralph would choose chocolate pie about 50 times and apple pie about 50 times. But given this—given that Ralph would choose differently in different "plays" of the decision, without anything about his psychology changing—it's hard to see how we can maintain that Ralph authored and controlled the decision. It seems to be a matter of *chance* or *luck* what he chose, and to the extent that this is true, it seems that Ralph didn't author or control the choice.[25]

The main point I want to make here is that it doesn't follow from the fact that Ralph would choose differently in different "plays" of the decision that he didn't author or control the decision. There is no inconsistency in claiming that (a) Ralph chooses differently in different plays of the decision, and (b) in each of the different plays of the decision, it's *Ralph* who does the choosing, and who authors and controls the choice. Indeed, given that Ralph is making a *torn* decision, the hypothesis that it's *him* who's making the decision (and who's authoring and controlling the decision) seems to *predict* that he would choose differently in different plays of the decision – that seems to be exactly what he would do if he was torn.

6 Responding to the Argument from Naturalism

The purpose of the above discussion was to respond to the luck argument against libertarianism. But it also gives us a response to the argument from spookiness. I have argued here that if our torn decision are neural events, and if they're TDW-undetermined, then they're also L-free. This shows that libertarianism is perfectly consistent with mind-brain materialism and, in particular, with the thesis that human decisions are neural events.

7 Responding to the Argument from Determinism

I now want to discuss the argument from determinism. The idea behind this argument is that libertarianism is false because the kind of indeterminism that it requires is false. In other words, the idea is to argue against libertarianism by arguing against TDW-indeterminism. (TDW-indeterminism is the sort of indeterminism that's needed for *my* version of libertarianism. Other libertarians might want to quibble with this in certain ways, but any (centered) version of libertarianism is going to require something very much like TDW-indeterminism.)

There are numerous ways in which one might try to argue against TDW-indeterminism, but the only really plausible strategy here is to zero in on our torn decisions and attempt to argue that *those very events* aren't undetermined in the manner of TDW-indeterminism. In other words, there's no plausible way to argue for a more general kind of determinism. This is because (a) we currently have no good reason to believe that *all* events are determined (because we have no good reason to believe that all *quantum* events are determined[26]); and more specifically, (b) we currently have no good reason to believe that all *neural* events are determined (current neuroscientific theory is probabilistic, and there's no good reason to think that this is just a simplification – i.e., for all we know right now, it could be that there are genuine, objective indeterminacies in human neural processes). Given this, it seems that in order to mount a cogent argument against TDW-indeterminism, one would have to focus in on torn decisions in particular and argue that *they* aren't TDW-undetermined.

Downloaded from https://onlinelibrary.wiley.com/doi/ by Mark Balaguer , Wiley Online Library on [12/09/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Another important point to note here is that since libertarians aren't committed to the claim that *all* of our torn decisions are TDW-undetermined (since they claim only that *some* of these decisions are TDW-undetermined), we can't undermine the view by appealing to general arguments from psychology that suggest that our decisions are often influenced by subconscious factors. Libertarians can (and should) admit this. All they need to claim is that *some* of our torn decisions aren't causally influenced by such factors (at the moment of choice).

But even granting all of these points, there are still ways in which one might try to argue that our torn decisions aren't TDW-undetermined. In particular, since our question here is an empirical one, one might try to find some scientific studies that give us reason to doubt that our torn decisions are TDW-undetermined. The most promising arguments here are based on the work of Libet et al. (1983), Tegmark (2000), and Haynes (2011).[27] I've responded to all three of these arguments elsewhere (2010, 2014a), but in this paper I'll just say a few words about the one that I think is the hardest to deal with, namely, the argument based on Haynes's study.

Haynes's study can be summarized very quickly: (i) Haynes gave his subjects two buttons, one for the left hand and one for the right; (ii) he told them to make a decision at some point as to which button to push; (iii) he used a very simple method to estimate the time at which the conscious decision occurred; (iv) using fMRI, he found unconscious neural activity in two different regions of the brain that predicted whether subjects would press the left button or the right; and, stunningly, (v) he found that this activity arose *7–10 seconds before the person's conscious decision to push the given button*.

These results generate a serious objection to TDW-indeterminism and libertarianism. We can put the objection like this:

> Since neuroscientists could predict which buttons Haynes's subjects would push 7–10 seconds before they consciously chose, their decisions weren't TDW-undetermined; for TDW-indeterminacy requires decisions to be undetermined *at the moment of conscious choice*, and Haynes's findings suggest that his subjects' decisions were already determined 7 seconds before they made their conscious decisions.

One way to respond to this argument would be to claim that the subjects' decisions weren't really *torn* decisions at all – that they were more like Buridan's-ass decisions – and that, because of this, Haynes's results don't tell us anything about torn decisions. But this seems pretty uncharitable to me, and I don't want to respond in this way. Instead, I'd like to grant for the sake of argument that Haynes's subjects' decisions were close enough to torn decisions – or, perhaps better, I'd like to just pretend for the purposes of this discussion that these decisions *were* torn decisions – and I'd like to argue that, even if we grant this point, the above argument doesn't give us any good reason to doubt TDW-indeterminism or libertarianism. In particular, I'd like to argue that when we take note of some of the details of Haynes's study, the argument against TDW-indeterminism falls apart.

There are two details of Haynes's study that I want to focus on. First, the pre-choice brain activity that Haynes found was actually not very good at predicting the outcomes of his subjects' choices. Indeed, it was only 10% more accurate than blind guessing. If we just guess which button subjects are going to push, we'll be right about 50% of the time, whereas if we use information about the brain activity that Haynes found, we'll be right at best 60% of the time. Now, this is definitely statistically significant, so it shows *something*. But it's not clear *what* it shows, and as I'll presently explain, it *doesn't* show (or, indeed, give us any good reason to believe) that TDW-indeterminism is false.

Downloaded from https://onlinelibrary.wiley.com/doi/ by Mark Balaguer , Wiley Online Library on [12/09/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

The second important detail of Haynes's study has to do with the specific *regions* of the brain where the pre-conscious-choice neural activity was found; it was found in the *parietal cortex* (or the PC) and *Brodmann area 10* (or BA10). The reason this is important is that these regions aren't associated with free decisions. Rather, they're associated with *plans*, or *intentions*. In particular, they're associated with the generation and storage of plans.[28]

I'm going to argue that when we put these two facts together – i.e., the facts described in the preceding two paragraphs – they suggest an explanation of Haynes's results that's perfectly consistent with TDW-indeterminism. I'll say in a moment what this explanation is, but first I want to make a background point.

When someone asks you *not* to think about something, it suddenly becomes very difficult to obey them. E.g., if I don't want you to think about chess right now, then one of the worst things I could do is to *tell* you not to think about chess. As soon as I say, "Don't think about chess," it will become very hard for you to avoid thinking about it, even if you sincerely want to obey me.

The same goes for *decisions*. Suppose I say this to you: "In a minute, I'm going to ask you to choose Door 1 or Door 2; but don't do it yet." It's actually somewhat difficult to refrain from thinking of one of the two options in situations like this. I'm not saying you *can't* do this; of course you can; but if we give these instructions to a group of people, it seems very likely that at least *some* of them won't succeed in refraining from thinking of one of the two options. Moreover, if you do think of one of the two options before you're supposed to choose, then this could influence how you end up choosing. Even if you tell yourself not to let this happen – even if you try not to let your pre-choice thoughts influence your decision – you might not succeed, and your pre-choice thoughts might have a causal influence over how you choose.

Now, here's the really important point for us: your choice could be causally influenced by prior goings on in your mind *even if you don't realize this*. You might subconsciously think of Door 1, and you might subconsciously store the plan to pick that door when the time comes. This shouldn't be controversial at all. For here are two things that we *know* to be true of humans: first, it's somewhat difficult for us to avoid thinking about things when people tells us not to think about them; and second, we do all sorts of things unconsciously. We don't do *everything* unconsciously, but we clearly do a *lot* of things unconsciously. When we put these two points together, we get the following (highly probable) hypothesis:

> If you tell a group of human subjects that in 60 seconds they're going to be asked to pick Door 1 or Door 2, and if you tell them not to pick yet—in other words, if you tell them to wait until the 60 seconds are up before they choose—then at least *some* of these subjects will (without realizing it) subconsciously think of one of the two doors before the 60 seconds have elapsed, and they'll subconsciously store the plan to pick that door when the time comes.

Again, given what we know about ourselves, this seems extremely plausible. Indeed, it seems to me that it would be very surprising if this weren't true.

These remarks suggest an explanation of Haynes's results that's perfectly consistent with TDW-indeterminism. The explanation I have in mind can be put in the following way:

> *An explanation of Haynes's results that's perfectly consistent with TDW-indeterminism and libertarianism*: A significant percentage of the subjects in Haynes's study (say, 20% of them) unconsciously failed to make truly spontaneous decisions about whether to press the right button or the left button. They genuinely *wanted* to follow Haynes's instructions, but for whatever reason, and without realizing it, they unconsciously formed prior-to-choice plans to push one of the two buttons. They unconsciously stored this information in their brains, and then when the time came, these

Downloaded from https://onlinelibrary.wiley.com/doi/ by Mark Balaguer , Wiley Online Library on [12/09/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

plans were activated. In other words, the regions of the brain where these plans were stored were activated. And this brain activity caused the subjects to choose in the ways in which they had unconsciously planned to choose. This explains various features of Haynes's findings. First, it explains why (in some subjects) there was relevant prior-to-choice brain activity in the PC and BA10 regions of the brain—because those regions are associated with the formation and storage of plans. Second, it explains why this brain activity predicted whether subjects would choose to push the left button or the right button. Third, it gives us a nice story about why the predictive brain activity occurred so long before the conscious choice. (We know that humans can make decisions in way less than 7 seconds (see Trevena and Miller (2010) for evidence that we can make decisions in *less than half a second*), and so the fact that Haynes found predictive brain activity 7-10 seconds prior to choice cries out for explanation. But the present explanation gives us a story to tell about this: it's because the relevant brain activity isn't a part of the conscious decision at all—it's associated with something completely distinct from, and prior to, the decision.) Fourth and finally, the present explanation of Haynes's data explains why the brain activity in the PC and BA10 is only 10% better, vis-à-vis accurate prediction, than blind guessing. The reason is that *not all* subjects unconsciously formed plans about what they were going to do. Only *some* of them did. Most of the subjects managed to avoid doing this, and so most of them succeeded in making truly spontaneous decisions. (Of course, if libertarians endorse this explanation, their claim won't be that most of us are L-free but some of us aren't; their claim will be that *all* of us *sometimes* fail to be L-free. The idea here is that we're all *sometimes* driven by things like unconscious plans; but we aren't *always* driven by such things.)

The first point to note about this explanation of Haynes's results is that if it's right, then there's no problem here for TDW-indeterminism or libertarianism. All Haynes's results show, if my explanation is correct, is that some of our decisions are influenced by unconscious factors. But we already *knew* this. Libertarians don't think that *all* of our torn decisions are TDW-undetermined. Moreover, they can admit (and *should* admit) that many of our torn decisions are causally influenced by unconscious factors in freedom-undermining ways – and in ways that make it the case that they're not TDW-undetermined or L-free. What libertarians claim is that this isn't *always* the case; again, their claim is merely that *some* of our torn decisions are TDW-undetermined. But given this, if my explanation of Haynes's findings is correct, then those findings don't give us any good reason to doubt libertarianism because they don't give us any good reason to think *none* of our torn decisions is TDW-undetermined. All these findings show is that *some* of our torn decisions (or more precisely, some decisions that are similar to torn decisions in certain ways) aren't TDW-undetermined – and this is perfectly consistent with the libertarian view that some of our torn decisions *are* TDW-undetermined.

In order to block my argument here – in order to maintain that Haynes's results create a problem for TDW-indeterminism – you'd have to endorse the following view:

The anti-libertarian interpretation of Haynes's data: Haynes's data suggest that TDW-indeterminism is false, because they suggest that *none* of our torn decisions is TDW-undetermined, because they suggest that *all* of our decisions (or at least all of our torn decisions) are determined (or at least significantly causally influenced) by prior-to-conscious-choice neural activity of the kind that Haynes found.

The first point I want to make about this view is that there's simply no good reason to believe it – Haynes's data just don't support this sweeping conclusion. And from this alone, it already follows that Hayne's data don't give us any good reason to doubt TDW-indeterminism.

But I want to argue for a stronger claim. I want to argue that my explanation of Haynes's data is *superior* to the anti-libertarian interpretation of Haynes's data. The reason that my

Downloaded from https://onlinelibrary.wiley.com/doi/ by Mark Balaguer , Wiley Online Library on [12/09/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

explanation is superior is that it's more explanatory. As we saw above, my explanation explains all of the following: (i) why we found predictive brain activity in the PC and BA10 regions; (ii) why this brain activity predicted whether subjects would push the left button or the right button; (iii) why this activity occurred so long before the conscious choice; and (iv) why this activity was only 10% better, in terms of accurate prediction, than blind guessing. In contrast, the anti-libertarian interpretation doesn't explain any of these things. Now, advocates of the anti-libertarian interpretation could explain (i)–(iii) by stealing *my* explanation – i.e., by claiming that the brain activity that Haynes found is associated with the formation and storage of unconscious pre-choice plans to push one of the two buttons. Advocates of the anti-libertarian interpretation would just have to add that this happens in *all* subjects – and, indeed, *all torn decisions*. But this doesn't fit at all with (iv); (iv) fits much better with my interpretation. Moreover, when we realize that the best version of the anti-libertarian interpretation is the one that steals the core of my explanation – i.e., the part about the formation of unconscious pre-choice plans – it throws into full relief how weird and unjustified it is for advocates of the anti-libertarian interpretation to claim that this occurs in *all* cases. In other words, it lays bare the fact that there's simply no evidence for this sweeping universal conclusion.

So I don't think we have any good reason to believe the anti-libertarian interpretation of Haynes's data, and I think that my explanation of Haynes's data is superior to the anti-libertarian interpretation.

8 HB-Libertarianism as an Open Empirical Question

The arguments of this paper are suggestive of the conclusion that, as of right now, we don't have any good arguments for or against HB-libertarianism. I obviously haven't done all the work necessary to establish this result here, but I think it's true, and so I think that the question of whether HB-libertarianism is true is an open question. Moreover, I think it's an open *empirical* question. In particular, I think the question of whether HB-libertarianism is true essentially boils down to the empirical question of whether TDW-indeterminism is true. I've partially argued for this claim here because I've argued that if TDW-indeterminism is true, then HB-libertarianism is true as well. But I think it can also be argued that if TDW-indeterminism *isn't* true, then HB-libertarianism isn't true either.[29]

Notes

1 I would like to thank Chris Franklin, Kristin Mickelson, and Leigh Vicens for comments on earlier versions of this paper.

2 Some of the philosophers who have defended libertarian views of one kind or another are: Bramhall (1655), Reid (1788), Chisholm (1964), Taylor (1966), Campbell (1967), Wiggins (1973), Thorp (1980), Nozick (1981), van Inwagen (1983), Kane (1985, 1996, 1999), Rowe (1987), Donagan (1987), Ginet (1990, 2002), Clarke (1993, 1996), McCall (1994), Goetz (1997), McCann (1998), O'Connor (2000), Ekstrom (2000), myself (2004, 2010), Pink (2004), Griffith (2007), Lowe (2008), Franklin (2011), Mawson (2011), Steward (2012), and Todd (2016).

3 In case you're wondering, the 'HB' stands for 'human being' and the 'C' stands for 'conceptual'.

4 If we wanted to, we could replace C-libertarianism with the following weaker thesis: "L-freedom is a legitimate kind of free will." If we proceeded in this way, then libertarianism would leave open whether there are other legitimate kinds of free will and, in particular, whether there are

Downloaded from https://onlinelibrary.wiley.com/doi/ by Mark Balaguer , Wiley Online Library on [12/09/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

legitimate kinds of free will that are compatible with determinism. I like this way of defining the view, but I won't worry about this here; I'll just assume that libertarianism is supposed to entail incompatibilism.

5 If they wanted to, libertarians could define 'appropriately non-random' in a way that implied that virtually Hume-free decisions don't count as non-random. But in order to do this, they would need to more or less build clause (b) of the definition of L-freedom into the definition of appropriate non-randomness—and so it wouldn't change anything in a substantive way. I think it's simpler and clearer to be less strict about what's needed for appropriate non-randomness and then to say that L-freedom requires that the indeterminacy generates (or procures, or enhances, or whatever) the non-randomness.

6 I argue for both of these claims in my (2010). Claim (b) might seem surprising. The argument for it is based on an argument for the claim that in this scenario, the answer to the conceptual-analysis question would be that the concept *free will* is identical to the concept *the freedom-like ability that humans actually have, whatever that turns out to be*; the answer would *not* be the specific freedom-like ability that humans happen to have.

7 I give much more thorough arguments for claim (a) in Chapter 8 of (Balaguer 2021) and for claim (b) in Chapter 2 of (Balaguer 2010).

8 Here's an analogy. The concept *even prime greater than 2* is obviously incoherent; but this doesn't mean that there's no such thing as the concept of an even prime greater than 2, and in fact, there clearly *is* such a concept. If there were no such concept, then the sentence 'There's no such thing as an even prime greater than 2' would be *gibberish*; but it's *not* gibberish; indeed, it's *true*.

9 We can also argue against the importance of the C-question by arguing that multiple concepts of freedom are at work in ordinary discourse so that there's no unique determinate answer to the what-is-free-will question. I argued for this claim in my (Balaguer 2010), Section 2.7.

10 Arguments of this general kind trace back at least to Kant. For a more recent version, see van Inwagen (1983). And for more on this general kind of argument, see Chan (2017).

11 See, e.g., Hobbes (1651), Hume (1748), Hobart (1934), Waller (1988), Double (1991), Bernstein (1995), Clarke (1995, 2002), Fischer (1999), Mele (1999a, 1999b), Haji (1999), O'Connor (2000), G. Strawson (2000), Berofsky (2000), van Inwagen (2002), and Levy (2011).

12 Just about everyone who has defended libertarianism has responded to arguments of this general kind. To name just a few who have responded to arguments of the kind outlined in the text, see, e.g., Kane (1999), Ginet (2007), Balaguer (2010), and Franklin (2011).

13 Agent-causal views have been developed by Reid (1788), Chisholm (1964), Taylor (1966), Thorp (1980), Rowe (1987), Donagan (1987), Clarke (1993), O'Connor (2000), Griffith (2007), Mawson (2011), and Steward (2012).

14 Event-causal views have been developed by Wiggins (1973), Nozick (1981), van Inwagen (1983), Kane (1985, 1996), Ekstrom (2000), Balaguer (2004, 2010), and Franklin (2011).

15 Non-causal views have been developed by Ginet (1990, 2002), McCall (1994), Goetz (1997), McCann (1998), Pink (2004), and Lowe (2008).

16 On some ways of conceiving of event-causal libertarianism, it has non-causal elements in it. For instance, many event-causal libertarians would say that our L-free decisions are probabilistically caused by our reasons for choosing; but one might think that what this means is that our reasons for choosing deterministically cause there to be certain objective probabilities of the various options being chosen and that, beyond this, which option ends up being chosen is uncaused. In other words, you might think that probabilistic event causation is just a mix of deterministic event causation and non-causation. Event-causalists don't have to endorse this line, but it strikes me as plausible.

17 Some people interpret Ekstrom (2000) as endorsing deliberativism; I think this is a misinterpretation, and in private correspondence, she has told me that she agrees with my interpretation of her view as a centered one.

18 I should say that it's *possible* to make a torn decision while in a Buridan's-ass situation – because you could be weird enough to *care* which can of Campbell's tomato soup you get, and so you could

Downloaded from https://onlinelibrary.wiley.com/doi/ by Mark Balaguer , Wiley Online Library on [12/09/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

feel genuinely torn about it. But *most* of us don't make torn decisions in Buridan situations. For example, in the above situation, most of us would just grab a can of soup without thinking about it.

19 On some ways of defining 'determined,' expressions like 'partially determined' don't make sense. But I'm not using the term in any such way here.

20 In saying this, I don't mean to commit to a type-type identity theory according to which the mental kind *decision* is identical to some physical kind; I just mean to commit to the token-token view that any specific decision is a neural event.

21 For more on why this sort of L-freedom is worth wanting, see Balaguer (2010).

22 Chris Franklin raised this worry to me.

23 For more on these issues, see Balaguer (2020) and Chapters 8 and 19 of Balaguer (2021).

24 Leigh Vicens raised this worry to me.

25 van Inwagen constructs an argument of this kind in his (2002).

26 Some philosophers seem to think that we have good positive reason to think that some quantum events are *un*determined. This, I think, is false – we have no good reason to believe this claim. But we also have no good reason to disbelieve it. In short, the question of whether there are any quantum events that are genuinely undetermined is an open empirical question.

27 Of course, these aren't the only scientific works that one might use to argue against TDW-indeterminism. Others include Festinger (1957), Milgram (1969), Isen and Levin (1972), and Velmans (1991). For a survey of potentially relevant scientific works, see Wegner (2002).

28 For evidence that BA10 is associated with the storage of plans and intentions, see, e.g., Burgess, Quayle, and Frith (2001) and Haynes, Sakai, Rees, et al. (2007). And for evidence that the PC is associated with the generation of plans, see, e.g., Desmurget and Sirigu (2009).

29 I argued for this claim in Balaguer (2010). I did this by arguing for the following two claims: (i) if our torn decisions aren't TDW-undetermined, then they're not (fully) L-free; and (ii) if our torn decisions aren't TDW-undetermined – and, hence, aren't (fully) L-free – then it's very likely that decisions of various other kinds that we might have thought were L-free (e.g., our leaning decisions and Buridan's-ass decisions) aren't (fully) L-free either. Claim (i) follows pretty easily from some things that I've argued for in this paper, but I haven't said anything in this paper about claim (ii).

Bibliography

Balaguer, M. (2004). A coherent, naturalistic, and plausible formulation of libertarian free will. *Nous* 38: 379–406.

Balaguer, M. (2010). *Free Will As An Open Scientific Problem*. Cambridge, MA: MIT Press.

Balaguer, M. (2014a). *Free Will*. Cambridge, MA: MIT Press.

Balaguer, M. (2014b). Replies to McKenna, Pereboom, and Kane. *Philosophical Studies* 169: 71–92.

Balaguer, M. (2020). Moral folkism and the deflation of (lots of) normative and metaethics. In: *Abstract Objects: For and Against* (ed. J. Falguera Lopez and C. Martinez-Vidal), 297–318. Synthese Library.

Balaguer, M. (2021). *Metaphysics, Sophistry, and Illusion*. Oxford: Oxford University Press.

Bernstein, M. (1995). Kanean libertarianism. *Southwest Philosophy Review* 11: 151–157.

Berofsky, B. (2000). Ultimate responsibility in a determined world. *Philosophy and Phenomenological Research* 60: 135–140.

Bramhall, J. (1655). *A Defense of True Liberty*. Reprinted 1977. New York: Garland Pub.

Burgess, P.W., Quayle, A., and Frith, C.D. (2001). Brain regions involved in prospective memory as determined by positron emission tomography. *Neuropsychologia* 39: 545–555.

Campbell, C.A. (1967). *In Defense of Free Will*. London: Allen & Unwin.

Chan, R. (2017) *Making metaphysics matter: essays on reasons and person*, PhD dissertation, University of Notre Dame.

Chisholm, R. (1964). Human freedom and the self. Reprinted in: *Free Will* (ed. G. Watson), 1982, 24–35. Oxford: Oxford University Press.

Clarke, R. (1993). Toward a credible agent-causal account of free will. *Nous* 27: 191–203.

Downloaded from https://onlinelibrary.wiley.com/doi/ by Mark Balaguer , Wiley Online Library on [12/09/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Clarke, R. (1995). Indeterminism and control. *American Philosophical Quarterly* 32: 125–138.
Clarke, R. (1996). Agent causation and event causation in the production of free action. *Philosophical Topics* 24: 19–48.
Clarke, R. (2002). Libertarian views: critical survey of noncausal and event-causal accounts of free agency. In: Kane (2002), 356–385.
Dennett, D. (1978). On giving libertarians what they say they want. In: *Brainstorms*. Cambridge, MA: MIT Press.
Desmurget, M. and Sirigu, A. (2009). A parietal-premotor network for movement intention and motor awareness. *Trends in Cognitive Sciences* 13: 411–419.
Donagan, A. (1987). *Choice*. London: Routledge & Kegan Paul.
Double, R. (1991). *The Non-Reality of Free Will*. New York: Oxford University Press.
Ekstrom, L.W. (2000). *Free Will: A Philosophical Study*. Boulder, CO: Westview Press.
Festinger, L., (1957). *A Theory of Cognitive Dissonance*. Palo Alto: Stanford University Press.
Fischer, J.M. (1999). Recent work on moral responsibility. *Ethics* 110: 93–139.
Franklin, C. (2011). Farewell to the luck (and mind) argument. *Philosophical Studies* 156: 199–230.
Ginet, C. (1990). *On Action*. Cambridge: Cambridge University Press.
Ginet, C. (2002). Reasons explanations of action: causalist versus noncausalist accounts. In: Kane (2002), 386–405.
Ginet, C. (2007). An action can be both uncaused and up to the agent. In: *Intentionality, Deliberation, and Autonomy* (ed. C. Lumer and S. Nannini), 243–256. London: Routledge.
Goetz, S. (1997). Libertarian choice. *Faith and Philosophy* 14: 195–211.
Griffith, M. (2007). Freedom and trying: understanding agent-causal exertions. *Acta Analytica* 22: 16–28.
Haji, I. (1999). Indeterminism and Frankfurt-style examples. *Philosophical Explorations* 2: 42–58.
Haynes, J.D. (2011). Decoding and predicting intentions. *Annals of the New York Academy of Sciences* 1224: 9–21.
Haynes, J.D., Sakai, K., Rees, G. et al. (2007). Reading hidden intentions in the human brain. *Current Biology* 17: 323–328.
Hobart, R.E. (1934). Free will as involving determinism and inconceivable without it. *Mind* 43: 1–27.
Hobbes, T. (1651). *Leviathan*. Reprinted 1962. New York: Collier Books.
Hume, D. (1748). *An Inquiry Concerning Human Understanding*. Reprinted 1955. Indianapolis, IN: Bobbs-Merrill.
Isen, A. and Levin, P. (1972). Effect of feeling good on helping. *Journal of Personality and Social Psychology* 21: 384–388.
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.
Kane, R. (1999). Responsibility, luck, and chance: reflections on free will and indeterminism. *Journal of Philosophy* 96: 217–240.
Kane, R. (ed.) (2002). *The Oxford Handbook of Free Will*. New York: Oxford University Press.
Levy, N. (2011). *Hard Luck: How Luck Undermines Free Will and Moral Responsibility*. Oxford: Oxford University Press.
Libet, B., Gleason, C., Wright, E., and Pearl, D. (1983). Time of conscious intention to act in relation to cerebral potential. *Brain* 106: 623–642.
Lowe, E.J. (2008). *Personal Agency: The Metaphysics of Mind and Action*. Oxford: Oxford University Press.
Mawson, T.J. (2011). *Free Will: A Guide for the Perplexed*. New York: Continuum Press.
McCall, S. (1994). *A Model of the Universe*. Oxford: Oxford University Press.
McCann, H. (1998). *The Works of Agency: On Human Action, Will, and Freedom*. Ithaca, NY: Cornell University Press.
Mele, A. (1995). *Autonomous Agents*. New York: Oxford University Press.
Mele, A. (1999a). Kane, luck and the significance of free will. *Philosophical Explorations* 2: 96–104.
Mele, A. (1999b). Ultimate responsibility and dumb luck. *Social Philosophy and Policy* 16: 274–293.
Milgram, S. (1969). *Obedience to Authority*. New York: Harper and Row.

Downloaded from https://onlinelibrary.wiley.com/doi/ by Mark Balaguer , Wiley Online Library on [12/09/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Nozick, R. (1981). *Philosophical Explanations*. Cambridge, MA: Harvard University Press.
O'Connor, T. (2000). *Persons and Causes*. New York: Oxford University Press.
Pereboom, D. (2014). The disappearing agent objection to event-causal libertarianism. *Philosophical Studies* 169: 59–69.
Pink, T. (2004). *Free Will: A Very Short Introduction*. Oxford: Oxford University Press.
Reid, T. (1788). *Essays on the Active Powers of the Human Mind*. Reprinted 1969. Cambridge, MA: MIT Press.
Rowe, W. (1987). Two concepts of freedom. *Proceedings of the American Philosophical Association* 62: 43–64.
Steward, H. (2012). *A Metaphysics for Freedom*. Oxford: Oxford University Press.
Strawson, G. (2000). The unhelpfulness of indeterminism. *Philosophy and Phenomenological Research* 60: 149–155.
Taylor, R. (1966). *Action and Purpose*. Englewood Cliffs, NJ: Prentice-Hall.
Tegmark, M. (2000). The importance of quantum decoherence in brain processes. *Physical Review E* 61: 4194.
Thorp, J. (1980). *Free Will*. London: Routledge & Kegan Paul.
Todd, P. (2016). Strawson, moral responsibility, and the "order of explanation": an intervention. *Ethics* 127: 208–240.
Trevena, J. and Miller, J. (2010). Brain preparation before a voluntary action: evidence against unconscious movement initiation. *Consciousness and Cognition* 19: 447–456.
van Inwagen, P. (1983). *An Essay on Free Will*. Oxford: Clarendon Press.
van Inwagen, P. (2002). Free will remains a mystery. In: Kane, (2002), 158–177.
Velmans, M. (1991). Is human information processing conscious? *Behavioral and Brain Sciences* 14: 651–669.
Waller, B. (1988). Free will gone out of control: a review of R. Kane's *Free Will and Values*. *Behaviorism* 16: 149–162.
Wegner, D.M. (2002). *The Illusion of Conscious Will*. Cambridge, MA: MIT Press.
Wiggins, D. (1973). Towards a reasonable libertarianism. Reprinted in: *Essays on Freedom of Action* (ed. T. Honderich), 31–61. London: Routledge & Kegan Paul.

Downloaded from https://onlinelibrary.wiley.com/doi/ by Mark Balaguer , Wiley Online Library on [12/09/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License