

BREAKIN' THE LAW AND FREELY WILLING THE INEVITABLE

Michael Nelson

UC Riverside

michael.nelson@ucr.edu

Abstract

Compatibilists maintain that it is possible that there are free acts in deterministic universes, whereas incompatibilists maintain that, necessarily, if determinism is true, then there are no free acts. The most influential incompatibilist argument has been some version of the consequence argument. Some compatibilists are tempted to respond to this argument by claiming that agents who act freely in deterministic universes have the ability to break laws of nature. I argue that this response is an adequate response to the consequence argument, but offers an inadequate defense against the strongest argument for thinking determinism threatens our freedom by robbing us of the control over what we do necessary for acting freely. I articulate such an argument—what I dub the *No Opportunity+* argument—for incompatibilism and show how it is immune to the above described compatibilist strategy.

BREAKIN' THE LAW AND FREELY WILLING THE INEVITABLE

§1. The Consequence Argument and Local Miracle Compatibilism

The Consequence Argument (CA) is one of the most influential arguments for incompatibilism, the thesis that, necessarily, if determinism is true, then there are no free acts.¹ The intuitive idea behind CA is that the truth of determinism would rob us of the ability to do other than we in fact do, possession of which is alleged to be necessary for acting freely. And this because, if determinism were true, being able to do other than one in fact does would require the incredible ability to effect either the laws of nature or the distant past, as determinism entails that it is impossible for the past and the laws to be as they in fact are and anything other than what in fact occurs to occur. Let '*L*' stand for the conjunction of all laws of nature, '*p*' a complete description of the intrinsic state of the universe at some point prior to my birth,² and '*q*' a true proposition to the effect that I performed some paradigmatically free act *A*. Then we can more rigorously state the argument as follows.

- (C1) If I act freely in *A*-ing at *t*, then at some time prior to *t* (call it *t**) I had the ability to do other than *A* and hence the ability to render *q* false. [*PAP*]
- (C2) If determinism is true, then, necessarily, (if *L* and *p*, then *q*). [definition]
- (C3) For any time *t*' and any propositions *p, q*, if I don't at *t*' have the ability to render *p* false and I don't at *t*' have the ability to render (if *p*, then *q*) false, then I don't at *t*' have the ability to render *q* false. [*transfer principle*]
- (C4) I don't at *t** have the ability to render (*L* and *p*) false. [*fixity of laws and past*]
- (C5) If determinism is true, then I don't at *t** have the ability to render (if *L* and *p*, then *q*) false. [from (2) and *fixity of necessary truths*, according to which, for all *p*, if *p* is necessarily true, then I never have the ability to render *p* false]
- (C) If determinism is true, then I do not act freely in *A*-ing at *t*.

There is no question of CA's validity. But there are questions of its soundness. In particular, there are questions of the truth of (C1), (C3), and (C4). ((C2) and (C5) are beyond dispute.) Although I think that all of these questionable premises are false, my focus here shall be (C4). (I shall return, in the final section, to (C1).) Following John Martin Fischer (Fischer 1983, 1986) and David Lewis (Lewis 1981), I shall argue that (C4) is false. But, I'll argue in §3, recognition of this fact leads naturally to an improved incompatibilist argument—what, following Kadri Vihvelin (Vihvelin 2000), we can call the *No Opportunity argument*—that can be made immune to the problems that face (C4) and is also independent of the transfer principle that supports (C3). This argument is thus a marked improvement over CA. In the final section I'll discuss how the best resist this argument.

Two general comments about CA. First, the idea behind thinking that possessing an ability to do otherwise is necessary for acting freely can be put thus. Acting freely involves being in control of what one does. What one does freely doesn't just happen. Rather, one makes it happen and is the ultimate source of its occurrence. But if this is so, then one should be free to do other than one in fact does. If it is truly up to me what I do, then it seems natural to insist that I could have brought about something else in its place. It seems that my being unable to keep something from happening is sufficient for its having just happened beyond my control. (I shall be returning in §4 to this line of reasoning, which, although intuitively appealing, I maintain to be faulty.)

Second, there is no issue whether or not determinism is incompatible with the ability to do other than one does *holding fixed the past and the laws*. This falls directly out of the definition of the thesis of determinism, which states that every true proposition concerning the state of the universe is entailed by the conjunction of p and L . So, any world in which the past and the deterministic laws are exactly as they in fact are is a universe in which exactly the same events occur; there is no possibility in which the deterministic laws and past are just as they in fact are and I do other than I in fact do. Yet it would be wrong to conclude from this observation that the truth of determinism thus entails that no one has the ability to do other than she in fact does. Thinking such an entailment immediate is to conflate determinism with the thesis of necessitarianism, according to which every truth is necessarily true and every falsehood necessarily false. Determinism is compatible with there being possible worlds that differ from actuality. What determinism requires is that these worlds either contain some, perhaps minute, differences of past events leading up to the point of divergence³ or not have the same laws of nature. The incompatibilist must argue, and not merely assume, that what is necessary for freedom is an ability to do otherwise holding fixed the laws and the past. And, I think, it is precisely here that CA fails. The argument to be developed in §3 will improve on this point.

Enough with the general comments. We aim to challenge premise (C4). This will allow us to claim, given that agents in deterministic universes possess the ability to render false the conjunction (L and p), deterministic agents have the ability to do other than they in fact do. There are two ways

of challenging (C4). According to the first, deterministic agents with the ability to do other than they in fact do possess the ability to render false L ; according to the second, such agents have the ability to render false p .⁴ I shall focus on the first, as it is more in keeping with the best available theory of counterfactuals—namely, the possible worlds semantics developed by Lewis (Lewis 1979).

In a characteristically brilliant yet frustrating paper (Lewis 1981), Lewis distinguishes two senses in which an agent has the ability to affect the laws. He argues that CA turns on conflating these two senses. One of these sense—the strong sense—is indeed incredible, but it is only the other sense—the weak sense—that compatibilism requires, Lewis claims. An agent has the ability to break a law in the strong sense just in case she has the ability to do something that is or causes something contrary to actual natural law. Lewis agrees that no finite being has this kind of ability to break a law⁵ and so, under this sense, premise (C4) is true. But then premise (C1) is false, as acting freely only requires the ability to do something such that were one to do it the relevant proposition would have been false. (Or so one might argue.) An agent has the ability to break a law in the weak sense, on the other hand, just in case she has the ability to do something such that, were she to do it, a law (or laws) would be different, in the sense that what is in fact a law would not have been a law. This, Lewis claims, is just the ability to perform or refrain from performing some ordinary action like raising one's hand. And that, claims Lewis, is an ability that some agents possess even in deterministic universes. So, if it is the weak sense that is operative in the argument, (C4) is false. Either way, there is no single sense of being able to render something false on which all of the premise of CA are true. Following Fischer (Fischer 1994), I shall call this *Local Miracle Compatibilism* (LMC).

§2. A Critique of Local Miracle Compatibilism

A proponent of LMC is committed to the claim that free agents in deterministic universes have the ability to break the laws of nature in the weak sense but that no agent has the ability to break the laws of nature in the strong sense. Helen Beebe (Beebe 2003) argues against LMC on the grounds that it cannot “provide us with any reason to suppose that it is *impossible* for agents at deterministic worlds to break the laws of nature” (*ibid.*, p. 268). She argues that, if true, this is

problematic as “if it cannot rule out such a possibility, then it is committed to the claim that it is merely a contingent matter whether or not agents are able to perform law-breaking acts,” whereas such a ban is not a mere contingent claim “about how our world happens to be, but a conceptual claim about the nature of laws” (*ibid.*, p. 268).

Beebee imagines herself at a real estate auction where someone else has bid above the price she has previously decided would be her highest bid. Very much wanting the house, she deliberates and decides to stick with her original conviction and does not raise her hand to become the new highest bidder. Suppose that this is a free act and hence, as she deliberates, she was able to make it so that she raises her hand. Had she raised her hand, a divergence miracle would have occurred. A proponent of LMC is committed to the claim that this divergence miracle was neither identical to nor caused by any possible act of Beebee’s. And this is what Beebee challenges. She asks why the divergence miracle cannot be her *deciding* to raise her hand. In general, and all other things being equal, a world in which the divergence miracle occurs closer to the time of the event in question rather than earlier will be more similar, as it will involve less divergence of past fact. (Of course, this can be overridden if the later divergence miracle is “too large” a miracle; a fact we shall exploit below.) So, it would seem, all other things being equal, having the divergence miracle be her deciding to raise her hand results in a world more similar to the actual world in which she doesn’t raise her hand than a world in which the divergence miracle occurs earlier. But, of course, if her decision to raise her hand is a divergence miracle, then, pace LMC, free agents do have it within their powers to do something that is or causes a law-breaking event. So LMC should be rejected.

This argument fails to persuade. Beebee herself has given a reason—a reason fully available to a proponent of LMC—why the divergence miracle is not her decision itself when she tells us that it is a conceptual truth regarding the relationship between agents and laws of nature that none of us can perform law-breaking acts. A world in which the divergence miracle is her decision, or indeed any other act on her part, is a world much less similar to actuality than a world in which some earlier event outside the agent’s control constitutes the divergence miracle as the former kind of world violates the conceptual truth Beebee reports. That is, such a world requires violating a conceptual

truth about finite beings and their relation to laws of nature; it requires too big of a miracle. If we agree with Beebee about its being a noncontingent conceptual truth that agents are not able to perform an act that is itself or causes a law to be broken—a claim her argument against LMC requires—then we have given a proponent of LMC an excellent reason for why the claim that Beebee says proponents of LMC can't justify. It is a conceptual truth!

Let's consider a case similar to one that Beebee herself uses to illustrate the difference between small and big miracles. Suppose that I am sitting in my house and that there is a bar 2 miles from my house. I could have been in the bar right now. Consider now which of the following two worlds is more similar to the actual world. In the first world everything is just as they are in the actual world up until 1 millisecond ago, at which time I disappear from my house and reappear at the bar a millisecond later. In the second world everything is just as they are in the actual world up until 10 minutes ago, at which time I walk out of my house, get in my car, and drive to the bar. It is obvious that the second world is more similar to the actual world than the first, despite the fact that the latter contains a greater departure in past facts. And this because my travel through space 10 miles in a millisecond constitutes a larger departure from actuality as it involves a big violation of law. This falls right out of Lewis's claims about similarity rankings. Similarly, the world in which Beebee's decision to raise her hand (given that in fact she doesn't) is a law-breaking event constitutes a larger departure from actuality as it involves a violation of a conceptual truth. I conclude that Beebee has not succeeded in refuting LMC. A proponent of LMC has a justification for insisting that ordinary agents can break laws of nature in the weak but not strong sense.

Many have not appreciated LMC's force. LMC provides a sound refutation of CA. Given an adequate understanding of counterfactuals, there is really no question whether or not determinism entails that no one can do other than they in fact do; they can. That being said, we should find LMC unsatisfying and indeed I think that an adequate theory of free action could hardly be constructed from the resources LMC delivers. And this is because, as I shall argue below, a weak ability to do otherwise, which is all LMC provides for, is hardly sufficient to ground the kind of control over what one does (arguably) necessary for acting freely.

§3. Lacking Opportunities

Let me build up to this by first talking about the nature of abilities generally. Suppose that I have the ability to turn a cartwheel. What does this ability consist in? Well, it's plausible that it at least involves, and arguably just consists in, the following being true: If I were to try to turn a cartwheel *in the right conditions*, then I would likely succeed. The truth of this counterfactual does not entail that I am likely to succeed in turning a cartwheel when I try with broken arms or with 200lbs weights strapped on my ankles or neck high in quick sand, as these are not the "right conditions." (Of course, these conditions must be specified independently of being the conditions in which attempt is likely to be successful.) In these circumstances, I am not at all likely to turn a cartwheel when I try. But, and here's the important point, even when I am in these unfortunate circumstances, I still have the ability to turn a cartwheel. (That is, I have the ability [to turn a cartwheel] even when I am neck high in quick sand, not that I have the ability [to turn a cartwheel even when I am neck high in quick sand], which I don't have.) One's abilities do not disappear when one is not in the right kind of circumstances for successful exercise of them. Abilities are similar to dispositions in that they have what we can call *actualizing* conditions. Actualizing conditions are something like the conditions necessary for the exercise of a given ability to be (likely to be) successful. If actualizing conditions are absent, then possession of an ability to do such and such need not imply success in such and suching is likely. Not being neck high in quick sand is an actualizing condition for my successfully exercising my ability to turn a cartwheel. If I spent all my days neck high in quick sand, my ability to turn a cartwheel would be masked, but not for that nonexistent.

LMC shows us that it is wrong-headed to think that determinism robs us of our ability to do otherwise, just as it would be wrong-headed to think that being neck high in quick sand robs me of my ability to turn a cartwheel. But what is less clear is that possessing a mere weak-ability to do otherwise suffices for capturing the intuition that in acting freely one *makes* something happen as opposed to merely being caught up in the causal flow of the universe. It seems very intuitive that the kind of ability to do otherwise relevant to freedom is the ability to *bring something about of one's own powers* and, at least one can argue, this requires its being within one's powers to bring

something else about holding the past and the laws fixed. If that's so, the incompatibilist wins. I shall now turn to an argument, inspired by the excellent discussion in (Vihvelin 2000), for just such a claim.

Vihvelin distinguishes abilities from opportunities, which corresponds to the above discussion about abilities and their actualizing conditions. Vihvelin agrees that determinism does not undermine the ability to do otherwise. But Vihvelin claims that this is a shallow victory for the compatibilist because the incompatibilist should simply recast her argument in terms of opportunities, arguing that determinism, although not robbing us the *ability* to do other than we in fact do, does rob us of the opportunity, which can be argued to be necessary for freedom. To act freely, the incompatibilist might insist, there must be nothing outside one's causal control preventing one from exercising one's abilities to do otherwise. One's ability to do other than one does is not sufficient for controlling what one does, given that there must be factors outside one's causal control that prevent one from successfully exercising that ability. Vihvelin thinks that the following principle captures this intuition. (My formulation slightly changes, and I hope improves, Vihvelin's; see *ibid.*, p. 146.)

Agent Causation Assumption (ACA): A person acts freely in X-ing only if she has the opportunity as well as the ability to do otherwise; that is, only if it is true both that she has the ability to do something other than X and also true *that if she had tried and succeeded in doing otherwise, everything except her choice, action, and the causal consequences of her choice and action would or at least might have been just the same.*

Vihvelin thinks that this principle gives voice to the following intuition: "Someone has the opportunity to exercise her ability to do X just in case there is no impediment to her doing X. If there is no impediment to her doing X, then *nothing would have to be different* in order for it to be true that her attempt to do X succeeds—except of course, her choice, action, and the causal consequences of her choice and action" (*ibid.*, p. 146). She is then assuming that the control over what one does necessary for freedom requires not just the ability to do otherwise but also the opportunity to exercise that ability, in the above sense. I agree with Vihvelin that this is what the (lee-way) incompatibilist should say.

I have the ability to turn a cartwheel. But poor me, I'm up to my neck in quick sand. Given my circumstances, there is a deep sense in which turning a cartwheel just isn't up to me—in which I can't turn a cartwheel—despite my possessing (even while I'm in the quick sand) the ability to turn a cartwheel. For me to successfully exercise the ability I possess, things outside my causal control (at the time, at least) would have to be different. I'd at least have to be out of the quick sand. Similarly, even if determinism doesn't rob us of our abilities to do otherwise, as LMC shows it does not, it still may threaten our freedom by robbing us of opportunities. Determinism may make us hapless bystanders in the production of our action, even while leaving in tact our abilities to do otherwise, precisely because it robs us of the opportunities to exercise the abilities we have. LMC leaves this worry untouched. LMC is thus unable to address the full worry—admittedly ill-expressed by standard incompatibilist arguments—that determinism threatens the control over what we do necessary for acting freely or acting with a free will.

Let's try to flesh this line of thought out in an argument. Consider the following *No opportunity argument* (*No-Op* for short), a simplification of the argument from *ibid.*, p. 147.

- (N1) I act freely in *A*-ing at *t* only if I had, at some time prior to *t* (call it *t**), both the ability and opportunity to not *A* at *t*.
- (N2) If at *t** I have the ability to not *A* at *t*, then, at *t**, I also have the opportunity to do not *A* at *t* only if, had I tried and succeeded in not *A*-ing at *t*, everything except my choice, action, and the causal consequences of my choice and action would or at least might have been just the same as they actually are.
- (N3) If I had not *A*-ed at *t*, neither my choice nor action would have caused the past (prior to my decision) to be different than it actually is.
- (N4) If determinism is true, then, if I had not *A*-ed at *t*, the past prior to my choice would have been different than it actually is.
- (N5) If determinism is true, if I have, at *t**, the ability to not *A* at *t*, I lack the opportunity to not *A* at *t*. [From (N2)-(N4)]
- (C) If determinism is true, I do not act freely in *A*-ing at *t*.

Vihvelin claims, and I agree, that this is both a marked improvement over CA and helps bring to light why so few find LMC satisfying.

Despite her helpful suggestions for supporting incompatibilism, Vihvelin is a compatibilist very much in Lewis's mold and so it is not surprising that she goes on to develop a powerful objection to No-Op. She argues that (N4) is false because it relies upon a faulty view of counterfactuals. The moves she makes are very reminiscent of Lewis's moves in responding to CA: Namely, she claims

that in evaluating counterfactuals we countenance divergence miracles, leaving the past largely in tact. So, on Lewis's theory of counterfactuals, the backtracking conditional 'Had such and such occurred, then the past would have been different' is typically false, whereas the contra-law conditional 'Had such and such occurred, then a local violation of actual law would have occurred' is true. Given this, (N4) is false, in that it (wrongly) claims that doing otherwise in a deterministic universe requires the past prior to one's choice to be different than it actually is.

Vihvelin is right about No-Op. But I think she has made things unduly easy for the compatibilist. To see why, consider the following argument—call it *No-Op+*—which is just like Vihvelin's No-Op but with the following replacements for (N3) and (N4).

(N3+) If I had not *A*-ed at *t*, neither my choice nor action would have caused the past (prior to my decision) nor the laws to be different than they actually are.

(N4+) If determinism is true, then, had I not *A*-ed at *t*, either the past prior to my choice or the laws would have been different than they actually are.

No-Op+ rests on the same basic assumptions that drive No-Op. Indeed, the stronger (N3+) seems even more in keeping with the motivation behind ACA than the weaker (N3). Everyone admits that it is incredible to claim that anyone has the ability to do something that is or causes a law breaking or that is or causes the past to be different.⁶ But then ACA would seem to not only council holding the past fixed, as Vihvelin's (N3) asserts, but also the laws, as (N3+) asserts, in assessing agency counterfactuals. After all, it isn't just the past that is outside my current causal control but also the laws; and, remember, Vihvelin's ACA tells us that we need to hold fixed *everything outside the agent's current causal control* in evaluating the relevant agency counterfactuals. (N3) is too weak to do full justice to the intuition behind ACA. (N3+) does better.

Notice that Vihvelin's objections to (N4) don't carry over to (N4+). (N4+), unlike (N4), does not rest on the (arguably) mistaken idea that in evaluating counterfactuals we vary past facts, countenancing the truth of backtrackers. Indeed, (N4+) seems impeccable, as it falls immediately out of the definition of determinism as the thesis that any world that exactly overlaps the actual world at any point of time and has the same laws as the actual laws overlaps the actual world at every point of time. But then Vihvelin's objection to No-Op is unsuccessful against No-Op+, as (N4+) survives her attack.

Vihvelin's ACA councils holding fixed not just the past but also the laws in evaluating agency counterfactuals. But if we hold both fixed, there is really no debate about whether or not determinism entails that those counterfactuals are false, as determinism (definitionally!) entails that there are no worlds just like the actual worlds in both past and laws with a divergent future. Once we move from No-Op to No-Op+, it is hard to see how Vihvelin's form of compatibilism is tenable.

No-Op+ has a long list of virtues. The argument is independent of how counterfactuals are evaluated. Everyone agrees that free agents do not have it within their powers to cause the past or the laws to be different than they actually are. Indeed, this is built right into the thesis of LMC. So, the worries proponents of LMC, and Vihvelin following their lead, raise do not cut against No-Op+. Furthermore, unlike CA, which evidently relies upon the transfer principle, No-Op+ (as well as No-Op from which it derives) is independent of the transfer principle. No step in the argument turns on distributing a power operator across a material conditional. This fact vindicates Fischer's claim, in (Fischer 1994) against van Inwagen, that there are arguments for incompatibilism that do not rely upon the transfer principle. The transfer principle and the fixity of the laws and the past—two points that have seen a lot of debate in the recent free will literature—are, I claim, not where the action is, or at least should be.

The only premise of No-Op+ up for dispute is (N1). Disputing this premise requires us to return to the conditions on control considered above in §1. This too is a point that has seen a lot of action over the years, but I think that the shift from (C1) to (N1) will help give that discussion a different shape. As we've already seen, the mere *ability* to do otherwise is cheap—almost certainly present in every case of moral responsibility. But the mere ability is far too weak to ensure agential control, Frankfurt-cases included. It is better for the incompatibilist to insist that it is the ability and opportunity to do otherwise that is necessary for acting freely in the sense relevant to moral responsibility, at least if what the incompatibilist is guided by is the idea that such action requires agential control.

In the following section I shall argue that (N1) is false, but not in a way that will be of much help to standard compatibilists. Indeed, I think that the dispute over (N1) requires accepting a rather radical suggestion of how to carve a way between traditional incompatibilist and traditional compatibilist theories of freedom, one which will allow that a free action might be inevitable even if not completely deterministic. I'll sketch such an account.

§4. Freely willing the inevitable

It is natural to think that the following two questions must be given the same answer.

Q1: Is it possible for an act to be free in the sense relevant to moral responsibility and deep attributability and yet the result of deterministic processes that stretch back prior to the birth of the agent of the act? In other words, is it possible for an act to be free in the sense relevant to moral responsibility and deep attributability and yet predetermined?

Q2: Is it possible for there to be free acts in deterministic universes, where a universe is deterministic just in case for every event e that occurs, there is a true proposition entailed by a conjunctive proposition about the laws of nature and a complete description of the intrinsic state of the universe at some time in the past to the effect that e occurs?

Naturalness does not make right and this is a case in point. The answer to Q1 is “yes,” while the answer to Q2 is “no.” Standard arguments for incompatibilism—and all three of the incompatibilist arguments considered above—are too dull to distinguish these questions. As a result, they either make our freedom unacceptably hostage to the fortunes and whims of empirical science or unacceptably allow philosophical theses to dictate empirical matters. On the flip side, standard compatibilist theories are also too dull to distinguish these two questions. As a result, they fail to provide a robust enough form of control over what we do to give us the kind of freedom moral responsibility is grounded in. To cut through this mire, we must embrace a radical seeming claim; one that will allow us to give the above answers to our questions. Or at least that's what I shall try to argue.

First, let me briefly sketch an argument that the answer to Q1 is “yes.” Let's call it the *simple-minded argument*. I shall assume that it is beyond dispute that we sometimes act freely. Standard forms of libertarianism end up entailing that our freedom is unacceptably hostage to the vicissitudes of empirical science. Paradigm acts, including free ones, are (or at least are tightly correlated with) bodily movements. But bodily movements are physical events—the stuff of scientific explanation. It

is an epistemically open possibility that the fundamental scientific explanations of all physical events—bodily movements included—are in terms of deterministic laws. That is, assuming that successful explanations are backed by causal relations, for all we know, every natural event, including every bodily movement, is the result of a deterministic process, or at the very least this is conceptually possible given everything we know immediately. But one of the things we know immediately is that we sometimes act freely and so that some of those bodily movements are free. So, my acting freely is at least conceptually compatible with those free acts being the result of deterministic processes. We can imagine being given absolutely compelling evidence that natural determinism is true. Nothing in these imagined grounds need defeat or otherwise neutralize our evidence that we sometimes act freely. This is in part because we are just more certain that we act freely than we are in any physical theory. But, more importantly, it is also because of the grounds on which we are so certain that we act freely; namely, we directly experience ourselves acting freely. If the claim that we act freely is compatible with these imagined grounds, then we must answer Q1 with a “yes,” at least if the possibility at stake is conceptual possibility.

The typical incompatibilist response to this argument—see (van Inwagen 1983, pp. 219-21)—is to claim that if we did indeed possess these imagined compelling grounds for determinism, then we would have good evidence for believing incompatibilism false. But, the response continues, it is only the actual discovery of this imagined compelling grounds, and not the mere epistemic possibility of such a discovery, that would count against incompatibilism. As things in fact are, without these good grounds for determinism in fact in place, incompatibilism is viable.

What is right about this response is that it grants that there is imaginable evidence that determinism is true that is consistent with our evidence that we sometimes act freely. Although there is no question of here doing justice to this response, I think it clear that it ultimately should not satisfy. The respondent surely takes incompatibilism to be well grounded and supported by convincing argument. But, given that the supposition that determinism is true undercuts the alleged grounds for incompatibilism, one must wonder about the strength of those grounds for incompatibilism even in the absence of any evidence in favor of determinism.

van Inwagen takes his incompatibilism to be supported by CA. I have already argued that CA offers poor support for incompatibilism, but let's set that to the side for the moment. van Inwagen suggests that, were he to know that determinism is true, he'd then be in a position to deny the transfer principle and so (C3) of CA, as his knowledge that he sometimes acts freely would then provide for compelling counterexamples to that principle. But those counterexamples are non-existent absent actual evidence that determinism is true. So, van Inwagen seems to think, his "flip-flopping" is justified.⁷ The reason the truth of determinism (rather than its truth being merely epistemically possible) undermines van Inwagen's grounds for incompatibilism is that it invalidates the transfer principle, which supports van Inwagen's incompatibilism. Although this may seem plausible, it is less clear which premise of No-Op and No-Op+ he proposes to deny given the supposition that determinism is true and on what grounds. Those grounds for denying that premise better not be capable of being employed even in the absence of compelling evidence in favor of determinism. As (N1) seems to me the only premise of No-Op+ even open to dispute, and I don't see how the ways of disputing (N1) could be sensitive to the distinction between possessing evidence that determinism is true and merely not being able to rule out that it is false, I don't see how van Inwagen's flip-flopping strategy could carry over to No-Op+. But, as I have already argued, it is No-Op+, not CA, that offers the incompatibilist the strongest argument. So, I don't think van Inwagen's flip-flopping strategy is likely tenable. The standard incompatibilist is forced to either withdraw the strongest argument in her favor—namely, No-Op+—in which case it is rather unclear what grounds her incompatibilism, or she must learn to live with the unhappy state of having whether or not we actually act freely hostage to whether or not determinism is actually true.⁸

Let us grant that the above-considered strategy for refuting the simple-minded argument fails. What does the simple-minded argument show? It would be an argument for compatibilism if we took a positive answer to Q1 to be a positive answer to Q2. This is because there is no question that the simple-minded argument, if sound, at least establishes that an act's being free is compatible with that act being the result of deterministic processes. This would imply a positive answer to Q2 as well if we did not distinguish **natural determinism**—which is the thesis that every natural event is

the result of deterministic processes—from **general determinism**—which is the thesis that every event simpliciter is the result of deterministic processes, where deterministic processes are such that prior events plus the laws necessitate the resulting effects. Now, to echo the words of Arthur Smullyan, it might be thought that making such a distinction is like being asked to distinguish tweedle-dum from tweedle-dee. But suppose that every free act has two kinds of causes, one natural and the other agential. And suppose that these agential-causings, although genuine occurrences, are non-natural, in the sense that their occurrences and non-occurrences are not governed by previous events and the laws of nature. (Randolph Clarke (Clarke 1993, 1996) suggests a similar version of agent causalism, although he insists—wrongly, in my view—that the redundant event causes must be indeterministic. This kind of view is also, I think, behind Kant’s resolution of the third antinomy (Kant (1998)), although for Kant, given his transcendental idealism, agential-causes are not occurrences.) Then there is room to say that, although every natural event is the result of a deterministic process stretching back to the distant past, it is false that every event whatsoever is the result of such processes, as the agential-causes are not. Hence, there is room to say that natural determinism is true while determinism is false. Furthermore, there is room for a view that requires that some occurrences in the causal production of a free act are independent of the natural laws—namely, the agential-causing—and hence not determined by prior events and natural law, thus requiring that determinism is false, while allowing that the act itself—which, after all, is just a natural occurrence with natural downstream effects—may well be inevitable given the past and the laws and hence allowing that acting freely is compatible with natural determinism, although not determinism.

What is this supposed to get us (besides incredulous stares)? I claim it gets us the following. We can respect the autonomy of science, not dictate from the armchair that the physical world is not governed by deterministic laws given the (alleged) conceptual argument for incompatibilism and the fact that we sometimes act freely while also not holding the fact that we sometimes act freely hostage to the final verdict on the status of natural determinism. And we do this without offering a surrogate of our concept of freedom—acceptance of which at least arguably involves accepting a

form of eliminativism, or at least radical revision, of free action—in the way standard forms of compatibilism do. This is because we can accept that acting freely, in the sense relevant to moral responsibility and authorship, does indeed require spontaneity. The key is to realize that the source spontaneity, in the sense of being an initiator of an effect without itself being caused to so initiate, need not—and indeed, should not—be located within the flow of natural events and that causal relations can be redundant, so that the sufficiency of one sort of causal chain leading to an effect does not exclude the existence of a whole distinct causal chain also leading to that same effect. This will then allow us to carve out a robust form of control that is compatible with natural determinism. That is, our distinction allows us to carve our way between traditional incompatibilist and traditional compatibilist theories of freedom in a way that will allow us to accept what seems correct about each of those views.

Our view enables us to answer standard worries with agent causalism to the effect that it is at odds with our “scientific view” of the world. For example, both David Velleman (Velleman 1992) and Michael Bratman (Bratman 2000) raise this worry. Velleman writes: “The obstacle to reconciling our conception of agency with the possible realities is that our scientific view of the world regards all events and states of affairs as caused, and hence explained, by other events and states, or by nothing at all. And this view would seem to leave no room for agents in the explanatory order” (Velleman 1992, p. 467). This drives Velleman to offer a reductivist account of agent causation, according to which there are states—for Velleman, the desire for self-knowledge, which he claims is constitutive of full-blooded action—that guarantee the agent’s participation and so, as Velleman says, “play the role of the agent.” Here I won’t consider the success of Velleman’s reductivism; I’ll just consider its motivation. He seems to be assuming that there can be irreducible, fundamental agential-causings—that is, fundamental causal relations between an agent, qua substance, and an event—only if there are events or states of affairs that are not caused by other events. Although this is no doubt true of standard forms of agential causalism—and in particular the form developed by Roderick Chisholm—see, for example, (Chisholm 1964)—it is simply false on the view that I have been articulating. The form of agent causalism I have articulated is consistent with every event

having an event cause—indeed, with every event having a deterministic cause. Of course, the view claims that some events have other causes as well, but that’s a different matter. As long as our “scientific view” doesn’t purport to be complete, telling the whole story, it can say whatever it wants about every natural event and there still be room for redundant agential-causes.

In a similar vein, Bratman writes:

One problem [with endorsing a motive being something that an agent primitively does and thus with agential causalism] is that it is difficult to know what it means to say that the agent, as distinct from relevant psychological events, processes, and states, plays such a basic role in the etiology and explanation of action. Second, and relatedly, in seeing the agent as a fundamentally separate and distinct element in the metaphysics of our action we seem to abandon the idea that our agency is as fully embedded in the event causal order as is the agency of purposive agents like dogs and cats. We are no doubt importantly different from dogs and cats; and the trio of core features of our agency that are my concern here [our reflectiveness, our planfulness, and our conception of our agency as temporally extended beings] are strong candidates for salient differences. But I would like to say what this difference is without abandoning the idea that we are all part of the same event causal order. (Bratman 2000, p. 39)

If Bratman’s worry with agential causalism is that it entails that human free action is not part the same event causal order as a cat’s purposeful activity, then he is simply mistaken; he is making the same assumption Velleman makes that a Chisholmian form of agent causalism is the only game in town for an agent causalist to play. If, on the other hand, Bratman is complaining that “our agency is **fully** embedded in the event causal order,” in the sense that event causes are all that there are in the causal production of a free action, then he is attempting to refute agent causalism by simply presenting a statement of its own thesis. Either way, his objection fails to move, once we acknowledge the possibility of a redundant form of agent causalism.

The idea that drives the simple-minded argument from above is satisfied so long as our conception of free action does not require that natural freedom be false. Empirical science is, and should be, simply silent on the existence and operation of non-natural events, if any there be. So the simple-minded argument simply could not be extended to show that our conception of freedom must be such that it does not require that general determinism be false. As long as the only reason that general determinism is false is because there are non-natural events not determined by the past and laws, then the autonomy (although not completeness, in the sense of “explaining all there is,”)

of science is respected. And, given the redundancy thesis, agential-causings do not require causal gaps in the natural events; that is, the existence of (redundant) agent-causings does not require that the physical universe not be causally closed. It's hard to see how there is going to be an argument from the alleged autonomy of science against a redundancy form of agent causalism.

In §1, I articulated a line of reasoning in support of the claim that the sort of control necessary for acting of one's own free will requires the existence of "robust alternative possibilities. Let me now briefly explain how the view I have developed can block the connection between control and the existence of "robust" alternative possibilities and, in particular, deny Vihvelin's ACA, which supports (N1). I reproduce ACA below.

Agent Causation Assumption (ACA): A person acts freely in X-ing only if she has the opportunity as well as the ability to do otherwise; that is, only if it is true both that she has the ability to do something other than X and also true *that if she had tried and succeeded in doing otherwise, everything except her choice, action, and the causal consequences of her choice and action would or at least might have been just the same.*

Our idea of agential-causings being non-natural events whose occurrences and non-occurrences are independent of natural law enables us to agree that an agent's acting freely in X-ing requires that *some aspect in the causal production of the agent's X-ing* be in her control, in the sense that she has the opportunity as well as the ability to have not done that. But it need not follow that that control must be a similar control over the bodily movement that constitutes the free action itself. If every event causally relevant to the production of the action were such that its non-occurrence would require variation in factors outside the agent's control, then, indeed, it is hard to see how the agent has a robust enough form of control over what she does to be counted as acting freely in the full-blooded sense. This is because acting freely requires spontaneity, or so I am assuming. This is exactly what is missing—and, indeed, must be missing—from standard compatibilists accounts of free action. But it does not follow that the act itself must be spontaneous. So, we can agree with the standard incompatibilist intuition that acting freely requires spontaneity in order to accord with the strictures of control without accepting that the act that is free itself must be spontaneous and hence entirely within the agent's control. Acting freely only requires that the agent had the ability and

opportunity to not participate in the production of that action; that is, acting freely only requires that the agent had the ability and opportunity to not agent-cause her act.

Needless to say, fleshing out the impressionistic view outlined in this section in an adequate way would require much more work. In closing I will briefly raise, and gesture towards an answer to, two questions that seem to me particularly pressing.

The first problem is that of saying what, exactly, agential-causings are. Here I claim we would do well to locate agential-causings in the notion that Harry Frankfurt (Frankfurt 1977, 1987) has called *identification*. On Frankfurt's view, acting freely, in the full-blooded sense, involves being moved by a motive with which one is identified. Frankfurt's idea is that one's act is truly one's own just in case one's effective will is truly one's own and that this in turn requires that the effective motive that constitutes one's will be one with which one is identified. In the original statement of the view (Frankfurt 1971), the notion of identification played no explicit role. Rather, one's will is one's own just in case one has the will one wants. So, one is identified with a motive just in case one wants that motive to be effective. As Gary Watson (Watson 1975) convincingly argued, merely desiring a given desire to be effective hardly suffices for guaranteeing that that desire "speaks for the agent;" that its guidance is the agent's guidance. Frankfurt seems to have agreed, and introduced the notion of identification, and ultimately wholeheartedness and being satisfied with one's will, where the notion of satisfaction involves the absence of pressures to change, to fill the gap.⁹ Frankfurt, and most following him, has been at pains to ensure that identifying with a motive is not a brute action of the agent; that is, most have insisted that an agent identifies with a motive just in case the agent is in a state that endorses that motive. I think that this reductivism about identification is wrong. We should take identification as brute, fundamental activity; it is something the agent does. It is, if you like, the residue of agent causalism.¹⁰ Arguing for this would require going through the candidate reductions and showing that they fail to ensure agential involvement. That is work for another day. I am here simply trying to indicate how one might go about to put flesh on the idea of there being non-reductive agential-causings.

Finally, a view like mine must come to grips with the problem of arbitrariness—a problem that faces any view that takes seriously the idea that acting freely requires spontaneity and ultimately indeterminism. The problem was first voiced, to my knowledge, by David Hume (Hume 1738, bk. II, pt. III, sec. 2). Here's how the problem affects my view. On my view, the agent's involvement—the agential-causing by identifying with an effective motive—is independent of the past and the laws of nature. But then, the worry goes, it seems arbitrary that the agent was involved in the production rather than not; it is inexplicable, that is, that she was identified with her effective motive as opposed to, say, alienated from it. After all, we explain by citing causes and those causes would have to be backed by causal, and hence natural, laws connecting the antecedents with their occurrence, which is just what my view says is lacking between the past states of the world and the agential-causings. So, agential-causings seem to be random happenings at best and hardly suitable as the source of freedom. (Donald Davidson (Davidson 1971, pp. 52-53) forcefully develops a similar argument against agent causalism.)

What we need to build a response to this argument are law-like connections between past states—and in particular, past attitudinal and deliberative states of the agent—and her agential-causings. Now, on my view, those laws cannot be natural laws. For if they were, I would have to either deny that acting freely requires spontaneity or else that the laws of nature themselves must be indeterministic if there are to be free acts. But why think that laws of nature are the only sort of laws capable of doing explanatory work? Those laws might be laws of rationality. So, to paint the picture crudely, it is explicable in terms of laws of rationality that one agent-cause such-an-such an act given so-and-so deliberative events leading up to that act because of the rational connection between the contents of those previous deliberative states and the action being agent-caused. The connection between the prior states and the agential-causings can be explicable even if they are not explicable in terms of natural law.

Derk Pereboom (Pereboom 2001, chapter 3; forthcoming) argues that a version of agent causalism like mine will face the problem of *inexplicable coincidence*. It seems a remarkable coincidence that free agents agent-cause exactly what the event causes cause. Isn't it remarkable,

that is, that the free acts (that is, the acts that are agent-caused) agree exactly with the acts that would have been produced had there been no agent-causes but the same event causes.

The same strategy considered above can be employed to solve the problem inexplicable coincidence. We need to find something that relates agent-causings to the event causings. Now agent-causings cannot be related to event causes by impersonal laws of nature. For if they were, agent-causings would not play the role they must play to allow for spontaneity. But we can still insist that there are laws that relate deliberative events to agent-causings. In particular, we can appeal to laws of rationality. It is a law of rationality that an agent agent-cause what she considers herself to have the most reason to do. But these laws—qua normative laws—admit of exceptions. It is thus *up to the agent* whether or not she agent-causes what she considers herself to have most reason to do, but still there is a law-like connection between the two causes of a free event that can make her doing so explicable. We thus have an explanation of the correlation between deliberative events and their co-causing agent-causes that does not rob agent-causes of being irreducibly “up to the agent.” The coincidence is not, contra Pereboom, inexplicable.

I have suggested a radical and, initially at least, strange sounding view of our freedom. Like standard compatibilist theories, my view sees no contradiction between one and the same act being free and yet inevitable given the past and the laws. Like standard incompatibilist theories, however, my view insists that acting freely requires spontaneity. I have aimed to rectify these two apparently incompatible claims by positing non-natural agential-causings as genuine causal, but redundant ingredients in the production of a free act. I have briefly considered some of the challenges such a view faces and gestured towards what I think are promising ways of meeting those challenges. Even if the view is rejected at the end of the day, it seems to me that its presence should be more widely acknowledged. Once it is, it will be seen that a wide host of arguments for standard forms of incompatibilism and compatibilism alike fail to make their mark.

NOTES

¹ CA has been developed independently by Carl Ginet (Ginet 1966, 1980, 1990), James Lamb (Lamb 1977), Peter van Inwagen (van Inwagen 1974, 1975, 1983), and David Wiggins (Wiggins 1973).

² What we are trying to exclude with the qualification to the universe's intrinsic state at a time is future-tensed statements to the effect that, for example, I will A at *t*. If we allow the description of the universe to contain information about future states, then the thesis of determinism is significantly weakened. (Some claim that the thesis is then trivialized, which is incorrect as some views entail that some future tensed truths are indeterminate, which is incompatible with future-tensed statements being determinately true. Allowing such "soft-facts" to be part of the entailment basis trivializes the thesis of determinism only if such future-indeterminate theses of future facts are false, which itself is a substantive matter.)

³ It is worth making explicit that, if the laws are identical between these two divergent worlds, then, if the laws are deterministic, there can be no instance in the past at which the two worlds are exactly alike; that is, there must be some (however minute) difference in at every instance.

⁴ Although this thesis is most often associated with David Lewis, John Martin Fischer (Fischer 1983, 1986) independently developed a similar response, as well as a version of the second variety in which it is claimed that free agents in deterministic universes have the power to affect the past in one sense but not in another. Fischer has since apparently altered his view of the successfulness of these responses—see (Fischer 1994, chapter 4). Fischer's reasons for abandoning his earlier view, however, are unsatisfying. He claims that both responses lead to a dialectic stalemate. (See *ibid.*, pp. 83-85.) However, he doesn't develop the dialectic far enough, it seems to me. Indeed, as I shall attempt to show below, there is no stalemate: The response is adequate with regard to CA and completely inadequate with regard to the best development of the incompatibilist intuition that if determinism is true, then we don't have the control over what we do necessary for acting freely.

⁵ We leave to the side the relationship between divine agency, if such there be, and laws of nature. On some views of divine agency—most famously, a Cartesian conception—the laws of nature, as well as the laws of logic!, are dependent on God's will and so God does have power over the laws of nature and, on some interpretations, is free to break the laws in the strong sense. Our focus here, however, is exclusively on finite agency.

⁶ First, recall that we are speaking here only of finite agents. Second, we should note that, if time travel into the past is possible, then it is possible to do something that affects the past, as it is possible for me to travel to the past and hence causally affect the past in the strong sense. (It is, of course, an entirely different issue whether or not it is within my power to *change* the past, making it other than it actually is.) But we should set aside cases of time travel as irrelevant to the discussion at hand. But it seems safe to set this to the side. Even if it is possible to causally affect the past as a time traveler, this is clearly of little help to the compatibilist who would like to claim that ordinary, non-time traveling agents can act freely.

⁷ Fischer describes van Inwagen's response as that of a "metaphysical flip-flopper." See (Fischer and Ravizza 1998, 253-54) and (Fischer 1999, 130n67).

⁸ Although he remains stubbornly unmoved by my efforts (as, alas, is his way), Keith DeRose has been extremely helpful in getting me to appreciate the strengths of van Inwagen's response and articulating my criticism.

⁹ This is true of Frankfurt's early and middle work on the hierarchical conception of acting of one's own free will. But in more recent work Frankfurt has turned to a different strategy, making use of "volitional necessities" and concerns (alleged to be) constitutive of one's identity in order to fill this gap. See, for example, (Frankfurt 1999).

¹⁰ I am here following (Velleman 1992) in seeing a connection between Frankfurt's work on identification, and the threat that identification is something an agent fundamentally does, and agent causalism.

REFERENCES

- Beebe, Helen. (2003) "Local miracle compatibilism," *Noûs* 37: 258-77.
- Chisholm, Roderick. (1964). "Human freedom and the self," The University of Kansas Lindley Lecture: 3-15. Reprinted in G. Watson (ed.) *Free Will, second edition*. 2003. (New York: Oxford University Press: 26-37.)
- Clarke, Randolph. (1993) "Toward a credible agent-causal account of free will," *Noûs* 27: 191-203.
- Clarke, Randolph. (1996) "Agent-causation and event-causation in the production of free action," *Philosophical Topics* 24: 19-48.

-
- Davidson, Donald. (1971) "Agency," in his (1980) *Essays on Action and Events* (Oxford: Clarendon Press, pp. 43-61).
- Fischer, John Martin. (1983) "Incompatibilism," *Philosophical Studies* 43: 127-37.
- Fischer, John Martin. (1986) "van Inwagen on free will," *Philosophical Quarterly* 36: 252-60.
- Fischer, John Martin. (1994) *The Metaphysics of Free Will* (Oxford: Blackwell.)
- Fischer, John Martin. (1999) "Recent work on moral responsibility," *Ethics* 110: 93-139.
- Fischer, John Martin and Mark Ravizza. (1998) *Responsibility and Control: A theory of moral responsibility* (Cambridge: Cambridge University Press.)
- Frankfurt, Harry. (1971) "Freedom of the will and the concept of a person," *Journal of Philosophy* 68: 5-20. Reprinted in his *The Importance of What We Care About*. (1988) Cambridge: Cambridge University Press, 11-25.
- Frankfurt, Harry. (1977) "Identification and externality," in A. Rorty (ed.) *The Identities of Persons*. (Los Angeles: University of California, pp. 239-51). Reprinted in his (1988) *The Importance of What We Care About*. (Cambridge: Cambridge University Press, pp. 58-68).
- Frankfurt, Harry. (1987) "Identification and wholeheartedness," *Ethics* 98: 5-20. Reprinted in his (1988) *The Importance of What We Care About*. (Cambridge: Cambridge University Press, 159-176).
- Frankfurt, Harry. (1999) "Autonomy, necessity and love in his *Necessity, Volition, and Love*. (Cambridge: Cambridge University Press, 129-41).
- Ginet, Carl. (1966) "Might we have no choice?" in K. Lehrer (ed.) *Freedom and Determinism* (New York: Random House), 87-104.
- Ginet, Carl. (1980) "The conditional analysis of freedom," in P. van Inwagen (ed.) *Time and Cause* (Dordrecht: D. Reidel Publishing Company), 171-86.
- Ginet, Carl. (1990) *Action* (Cambridge: Cambridge University Press).
- Hume, David [1738] (1978). *A Treatise of Human Nature*. (Oxford: Clarendon Press).
- Kant, Immanuel. [1781/1787] (1998) *The Cambridge Edition of the Works of Immanuel Kant, Critique of Pure Reason*. Trans. and eds. P. Guyer and A. Wood. (Cambridge: Cambridge University Press).
- Lamb, James. (1977) "On a proof of incompatibilism," *Philosophical Review* 86: 20-35
- Lewis, David. (1979) "Counterfactual dependence and time's arrow," *Noûs* 13: 455-76.
- Lewis, David. (1981) "Are we free to break the laws?" *Theoria* 47: 112-21. Reprinted in his (1986) *Philosophical Papers volume II* (New York: Oxford University Press), 291-98. [Citations to reprint.]
- Pereboom, Derk. (2001) *Living Without Free Will*. (Cambridge: Cambridge University Press).
- Pereboom, Derk. (forthcoming) "Kant on transcendental freedom" *Philosophy and Phenomenological Research*.
- van Inwagen, Peter. (1974) "A formal approach to the problem of free will and determinism," *Theoria* 24: 9-22.
- van Inwagen, Peter. (1975) "The incompatibility of free will and determinism," *Philosophical Studies* 27: 185-99.
- van Inwagen, Peter. (1983) *An Essay on Free Will*. (Oxford: Oxford University Press.)
- Velleman, David. (1992) "What happens when someone acts?" *Mind* 101: 462-81.
- Vihvelin, Kadri. (2000) "Libertarian compatibilism," *Philosophical Perspectives* 14: 139-66.
- Watson, Gary. (1975) "Free agency," *Journal of Philosophy* 72: 205-20.
- Wiggins, David. (1973) "Towards a reasonable libertarianism," in T. Honderich (ed.) *Essays on Freedom of Action* (London: Routledge & Kegan Paul), 31-61.