

Article

Determinism, Counterfactuals, and the Possibility of Time Travel

Kadri Vihvelin

School of Philosophy, University of Southern California, Los Angeles, CA 90089, USA; vihvelin@usc.edu

Abstract: The Consequence argument is an argument from plausible premises—our lack of causal power over the laws and past—to an implausible conclusion: that if determinism is true, we are equally powerless with respect to the future. What the compatibilist needs is a theory of counterfactuals that preserves the links between counterfactuals, causation, and the natural laws in a way that supports our commonsense belief that we have the power to make a causal difference to the future but no such power with respect to the past. Lewis’s critique of the Consequence argument was based on his theory of counterfactuals and his analysis of causation as a counterfactual relation between particular events. He argued that, at a world that is deterministic in the way that ours might be, counterfactuals are temporally asymmetric in a way that matches the contingent temporal asymmetry of causation. So it is not surprising, but only to be expected, that the past is causally closed while the future is causally open. If this worked, it would be just what the compatibilist needs. But it doesn’t work. There is an argument, due to Tooley and recently endorsed by Wasserman, that a fundamental feature of Lewis’s theory of counterfactuals is inconsistent with the metaphysical possibility of time travel and backwards causation. If this is right, then Lewis’s response to the Consequence argument fails. I endorse this conclusion, but argue that there is a better theory of counterfactuals—a theory that leaves open the metaphysical possibility of time travel to the past and backwards causation.

Keywords: determinism; counterfactuals; time travel; Consequence argument; laws; Lewis; causation; abilities



Citation: Vihvelin, K. Determinism, Counterfactuals, and the Possibility of Time Travel. *Philosophies* **2023**, *8*, 68. <https://doi.org/10.3390/philosophies8040068>

Academic Editor: Alasdair Richmond

Received: 14 June 2023

Revised: 13 July 2023

Accepted: 13 July 2023

Published: 25 July 2023



Copyright: © 2023 by the author. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Should a compatibilist about free will and determinism endorse, or at least not deny, the metaphysical possibility of time travel, understood the usual way as including time travel to the past and the backward causation that this entails?

This may seem an odd question. If time travel to the past and backward causation turns out to be metaphysically possible then no one, including the compatibilist, should deny this. But why should the defense of compatibilism require taking a stand, one way or the other, about the possibility of backward causation? What does the former have to do with the latter?

There is a connection. To be a compatibilist is to believe, at least¹, that determinism is compatible with the kind of control over our choices and actions that is a necessary condition of being morally responsible for those choices and actions. Control is a causal notion. So a compatibilist must believe that determinism is compatible with causation.

It is widely taken for granted, in the free will literature, that causation, including mental causation and even reasons-responsive mental causation [1], is compatible with determinism: the debate is about whether this kind of causation provides us with the kind of control that justifies or grounds or at least makes permissible our practice of praise, blame, and holding one another responsible. But the kind of determinism taken seriously in contemporary physics permits actual and counterfactual inferences in both temporal directions [2,3]². Given enough information about the present state of the world, we can, at least in principle, use a deterministic system of laws to deduce the future and also the past. If the present were different, in any way, then, given the laws, the past as well as the future would also be different—*would have to be different*—at every other time. Despite this, we

think that causation goes in one temporal direction, but not the other. We think that the past causes the future (by way of the present), but we do not think that the future causes the past (by way of the present).

Why is this? The compatibilist needs an answer because compatibilism is true only if our ordinary beliefs about causation are more or less true, and our ordinary beliefs include the belief that our choices often make a causal difference to the future, but they *never* make a causal difference to the past.

One answer is that this is because it is the way *it has to be*. Causation is, as a matter of *metaphysical necessity*, temporally asymmetric. Causes must, *always*, happen before their effects. Time travel to the past and backward causation are metaphysically impossible.

A compatibilist might make this claim. However, in doing so, the compatibilist is making a claim that is empirically and metaphysically risky. It is empirically risky because the nomological possibility of time travel (and thus backward causation) is now taken seriously by some physicists and philosophers of physics [4]. It is metaphysically risky not only because making claims of metaphysical impossibility is always risky but also because, after many years of debate [5]³ the consensus in the literature is that time travel worlds, while strange in many ways, are metaphysically possible worlds.

The better way to be a compatibilist, in my view, is to not deny the possibility of time travel or backward causation and to find a way to show that the kind of backward counterfactuals that are true in deterministic worlds are not causal counterfactuals. In order to do this, you will need a theory of counterfactuals.

I am that kind of compatibilist and so was David Lewis. In “Are We Free to Break the Laws?” [6] he defended compatibilism against the best argument for incompatibilism, the Consequence argument [7]. In “The Paradoxes of Time Travel” [5], he defended the conceptual and metaphysical possibility of time travel.

Both papers have been highly influential but in different ways. The time travel paper succeeded in convincing nearly everyone who thinks about the topic that time travel is conceptually and metaphysically possible. Lewis’s reply to the Consequence argument has not had a similar effect; arguably it has had a negative effect insofar as there is now a perception, in parts of the free will literature, that to be a certain kind of compatibilist (a “leeway” compatibilist or, as I prefer, a commonsense compatibilist) one must be a Lewisian “local miracles” compatibilist. And since the philosophers who think seriously about the free will/determinism problem are, for the most part, not inclined to accept Lewisian “local miracles compatibilism”, the net effect is that there are fewer commonsense compatibilists than there were fifty years ago.

In other work, I have defended commonsense compatibilism [8] and argued that the fact that the freedom of time travelers is constrained in certain kinds of ways does not count against the possibility of time travel [9,10]. In this paper, I will focus on a connection between the possibility of time travel and Lewis’s way of responding to the Consequence argument. I will be arguing that if you want—as you should, given the vast literature defending the possibility of time travel—to endorse or at least not reject, the claim that time travel and backward causation are metaphysically possible, you will have reason to reject Lewis’s theory of counterfactuals in favor of an alternative theory—the Simple theory—which bans miracles and preserve the laws. And if you are a compatibilist, you should use this theory, rather than Lewis’s, in your response to the Consequence argument.

2. Counterfactuals, Causal Powers, and the Consequence Argument

The Consequence argument, defended, most prominently, by Peter van Inwagen [7], changed the minds of a generation of philosophers. This is surprising, for the conclusion of the argument is that the truth of determinism is incompatible with our ever having a choice about *anything*. What we actually do or cause is, always, the *only* thing we are able to do or cause. This goes directly against our commonsense belief that the future, unlike the past, is at least partly up to us. If the Consequence argument is sound, then determinism

entails the kind of modal collapse that the logical fatalist [11]⁴ says follows from truth and logic alone.

How did that happen? The Consequence argument, in its best-known rule Beta version [7]⁵ is a modal argument that uses a necessity operator (informally expressed in terms of ‘no choice about’ or ‘no power over’ or something similar). Its conclusion is that if determinism is true, the entire future is necessary *in this sense*. However, we always knew that there is a sense in which the future, in a deterministic world, is necessary: there is only one future that is compossible with the laws and the total state of the world at any earlier time.

This sense of necessity is not the one we use in real life when we make judgments about what human beings and other animals can and cannot do. As G.E. Moore [12] observed, long ago, determinism does not have the consequence that there is no difference between cats and dogs, so far as tree-climbing is concerned. Cats can climb trees; dogs cannot. Climbing a tree is sometimes an option for a cat in a way that it never is for a dog. The future of a cat is open, with respect to tree-climbing, in a way that the future of a dog is not. Similar things can be said about the more complex powers of deliberation, choice, and intentional action possessed by human agents [12]⁶.

Moore’s observation should not be confused with the view that is usually attributed to him: that to ascribe an ability to an agent is to assert some conditional or counterfactual about what that agent would do if she chose. Moore did suggest that “we very often mean by ‘could’ merely ‘would, if so and so had chosen’”. However, he was less confident of this than he was of his belief that there are real differences in abilities and that the truth of determinism would not have the consequence that we are mistaken about this.

The key to understanding Moore’s common sense compatibilism lies in two facts: the distinction between the *existence* of an ability or power and the *exercise* of that power, and the fact that, regardless of the truth or falsity of determinism, the successful exercise of human abilities always depends on circumstances outside our control. If I had voted ‘no’, at that department meeting—I actually voted ‘yes’—it would have been because I had come to believe that there were good reasons for voting ‘no’, reasons I did not in fact see at the time. And this would have been because something about the recent past (prior to my decision) was a bit different in certain kinds of ways—one of my colleagues might have made a better argument, or I might have remembered something I did not actually remember or thought harder about the possible consequences of the proposal, or something else. Past history, whether recent or remote, is not in my control. However, this fact—the fact that I would have voted “no” only if something not in my control had been different—doesn’t mean that it was not in my power to vote ‘no’. It does not mean that I was not able to vote ‘no’, that it wasn’t up to me whether I voted ‘yes’ or ‘no’, that I had no choice about how I voted.

So says commonsense. Moreover, so said compatibilists, back in the past, before the Consequence argument. Why did this argument change so many minds? Why did it change the minds of philosophers who might, had they grown up in a different philosophical era, have been compatibilists?

Some of the reasons were historical. The Consequence argument first came on the scene in the mid-seventies of the last century [7,13–15], hard on the heels of the last days of attempts by ordinary language philosophers to defend the claim that all we ever mean when we say things such as ‘he could have done otherwise’ or ‘he was able to do A’ can be analyzed in terms of a conditional along the lines of ‘if he had chosen (or wanted, decided, tried) to do otherwise/to do A, he would have done otherwise/done A’⁷. A series of apparent counterexamples had left everyone convinced that no analysis of this kind could succeed [8,16,17]⁸. Arguments by Strawson [18] and Frankfurt [19] had offered compatibilists an appealing alternative: that so far as blame and moral responsibility are concerned, it does not matter whether we are ever able to do other than what we actually do; all that matters is the “actual sequence” [1,19,20]. So by the time the earliest versions of the Consequence argument came along [13,14]⁹, the ground had been prepared for an

argument that the kind of commonsense compatibilism that Moore took for granted might be fundamentally mistaken.

The timing was right in another way. Earlier arguments about free will and determinism had understood determinism as the thesis that every event has a cause. This left incompatibilists in the uncomfortable position of having to explain what the libertarian alternative is. How can an uncaused event be a free choice? The Consequence argument defines determinism in the way that physicists and most philosophers now understand it: as a thesis about the laws: that they are deterministic, rather than probabilistic and that they apply to everything, without exception¹⁰, and van Inwagen went out of his way to make it clear that he was not making any claims about causation¹¹. This left room for incompatibilists to argue that determinism might be false in a way that is compatible with the ability to *cause ourselves* to make undetermined yet free choices [21,22].

However, it was not just a matter of timing. The Consequence argument is a simple and elegant argument with just two premises that seem undeniable: the “fixity” or necessity of the laws and the “fixity” or necessity of the remote past. However we understand this, it seems uncontroversially true that there is some sense in which facts about the laws and the past (remote or recent) are “fixed” or necessary¹². It is not in *our* power to “render them false”, as van Inwagen put it. Given these two premises, van Inwagen needed only two inference rules for his necessity operator: Rule Alpha, which says that if *p* is a “broadly logically necessary” truth, we may infer *Np* [23,24]¹³, and Rule Beta, which says that from *Np* and *N(p ⊃ q)*, we may infer *Nq*.

In the formal argument below, ‘*P*’ abbreviates any sentence expressing a true proposition.

1. *N* (*L*)
2. *N* (*H*)
3. $\Box (L \ \& \ H. \ \supset \ P)$
4. $\Box (L \ \supset \ (H \ \supset \ P))$
5. *N*(*L* $\supset$ (*H* $\supset$ *P*))
6. *N*(*H* $\supset$ *P*)
7. Therefore, *N* (*P*)

Premise (1) asserts the *N* necessity of the true proposition that is the conjunction of all the laws of nature. (2) asserts the *N* necessity of a true proposition about the total state of the world at some time in the remote past. (3) is entailed by determinism and (4) follows from (3) by modal and sentential logic. (5) follows from (4) by rule Alpha. (6) follows from (1) and (5) by rule Beta. (7) follows from (2) and (6) by rule Beta. (7) says that every true proposition whatsoever, including propositions about the future choices and actions of agents, is *N* necessary.

If you have not thought about determinism, it might seem obvious that if we accept the premises we must also accept the conclusion, and so we must agree that if determinism is true then every true proposition is *N* necessary. However, this is a mistake, for the reasons I have already given. As a matter of fact, I did not vote ‘no’ (by raising my hand, at the right time, at that department meeting). If you were willing to say that I was able to vote ‘no’, despite the fact that I would have exercised my ability only if facts *not in my control*—facts about the past just prior to my decision—*had been a bit different*, the truth of determinism and the Consequence argument should not make you change your mind. The Consequence has the intuitive force that it does because it points out, in a particularly dramatic way, something we already knew—that the successful exercise of our abilities always depends, at least partly, on facts not in our control.

If something like this is right—and it seems to me that it is (that is, that the only sensible story about human abilities, whether in a deterministic or non-deterministic world, must say something similar to this)—[8]¹⁴ why are we still talking about the Consequence argument?

There is a short answer and a much longer answer. The short answer is that I blame Lewis. The longer answer is that I still blame Lewis, but for deeper and more interesting reasons, and I offer a non-Lewisian alternative. I will explain.

3. Counterfactuals, Abilities, and Lewis's Reply to the Consequence Argument

In his highly influential “Are We Free to Break the Laws?” [6], Lewis did get the most important thing right. The Consequence argument fails because it equivocates. We can understand the premises in a way that makes them clearly and uncontroversially true. However, if we understand them this way, the argument is invalid. Alternatively, we can understand the premises in a way that makes the argument valid, but in that case, the premises are no longer clearly and uncontroversially true. The sense in which the premises are clearly and uncontroversially true is causal. No human being is ever able to do anything¹⁵ that makes or would make, a *causal difference* to the laws, the past (recent or remote), or any logical truths. But it does not follow—even on the supposition of determinism—that no human being is ever able to do anything that makes, or would make, a causal difference to the future. The argument is valid if we understand the premises in a stronger way: as asserting a claim about a *counterfactual constraint* on our abilities or causal powers: that we are never—even on the supposition of determinism—able to do anything which is such that *if we did it*, either the laws or the past (or a logical truth) *would be other than they actually are*. However, if we understand the premises in this stronger way, then they are no longer clearly and uncontroversially true. If determinism is true, and common sense is right, then I was able to raise my hand (at that department meeting, to vote ‘no’ on that proposal), and had I done so, at least some of past history would have been at least somewhat different. But it doesn’t follow, Lewis says, that my choice or my act would have *caused* this difference. “If anything, the causation would have been the other way around” [6,8,25,26]¹⁶.

However—this is the short answer—I still blame Lewis, and for two reasons. First, he did not explain the key distinction between the *existence* of an ability or power—its possession, by an agent, at a time—and its *exercise*—by that agent, at that time—nor did he say anything to suggest that it is always the case that the successful exercise of our abilities depends on factors not in our control [5,6,27]¹⁷. Second, while he briefly acknowledged the possibility that a compatibilist might respond to the Consequence argument in a different way, on the basis of a theory of counterfactuals that, unlike Lewis’s, holds the laws fixed, most of “Are We Free to Break the Laws?” consisted in an elaboration of how Lewis’s theory of counterfactuals can be used to show that the argument fails.

Moreover, because he was Lewis, and because he did eventually succeed in persuading many people [28,29]¹⁸ that his theory of counterfactuals (or something in the neighborhood) [30,31]¹⁹ is the right theory, Lewis came to own the reply to the Consequence argument. This led to the widespread belief that to be a (commonsense) compatibilist you had to be a Lewisian Humean “local miracles compatibilist”: you had to accept Lewis’s claim that the past is not only causally but also *counterfactually fixed*²⁰, while the laws are only causally fixed. There was also confusion about what it is to be a Lewisian compatibilist, due to the misleading language that Lewis used to draw the distinction between the ability to *cause* a difference to the laws and the ability to act so that the laws *would have been different, but not in any way caused by the agent*. Instead of making it clear that the distinction is between what the agent would have *caused* in the non-actual world and what *would (merely) have been the case at that world*, Lewis talked about a “weak” and a “strong” ability to “break” the laws. However, ‘break’ is a causal verb and this, together with Lewis’s insistence that no particular law-breaking event would have occurred, gave some people the false impression that the weak ability to break the laws is similar to a weak ability to break windows: the ability to break some window (law) or other, but not the ability to break the particular window (law) one wants to break.

Lewis did not in fact intend anything of the kind, and since he was, in his own words “a commonsensical kind of fellow except where possible worlds are concerned”, he knew, even if he failed to clearly state, that a successful reply to the Consequence argument must deny that free deterministic agents have what he called an “incredible ability”—the ability to causally affect either the past or laws. Since Lewis believed that causation is a relation between events, he thought that it sufficed, for the denial of the “incredible ability”, to

show that there is no *particular past event*—whether the divergence miracle or some other event—that counterfactually depends on the free act of a deterministic agent.

To sum up: The short answer for why I blame Lewis is that he replied on the basis of his theory of counterfactuals and causation, and in an overly cryptic way. Those familiar with the rest of Lewis's work understood what he was saying. However, many of the participants in the free will debates did not, which led to a lot of misunderstanding. Some even seemed to think that Lewis had invented his theory of counterfactuals for the sole purpose of defending a peculiar view: "local miracles compatibilism".

Nevertheless, there was a good reason why Lewis continued to own the reply to the Consequence argument. He understood that the "incredible ability" that the compatibilist must deny is a causal power with respect to either the laws or the past. And he alone had the metaphysical resources—a theory of counterfactuals and a counterfactual analysis of causation [32]—to explain the difference between a *mere* counterfactual dependence and a counterfactual dependence that is also a *causal dependence*. This put him in a position to say, without making it true by stipulation, that causation and our causal powers are *temporally asymmetric* in more or less exactly the way we believe them to be. The past *could* be different, but it is not in *our power* to *cause it to be different*.

If all this worked, the compatibilist would be on very solid ground. Unfortunately, it does not work. (This is the beginning of the longer answer for why I blame Lewis.) What Lewis needed was a theory of counterfactuals with a similarity metric that is temporally neutral but which combines with a contingent feature of the world to yield counterfactuals that are temporally asymmetric in just the right kind of way to match the contingent temporal asymmetry of causation. In other words, Lewis needed a theory of counterfactuals that does not rule out, by stipulation, the possibility that a cause happens after its effect. Moreover, Lewis thought, not unreasonably, that he had such a theory. In his "The Paradoxes of Time Travel" he had defended the claim that time travel to the past and the backward causation it entails is metaphysically possible [5].

Lewis defended his claim that counterfactuals are *contingently* temporally asymmetric by the argument he gave in "Counterfactual Dependence and Time's Arrow" [2]. There have been many criticisms of that argument [33–39]. However, this paper has nevertheless had a lasting impact insofar as Lewis convinced many, including those who were not persuaded by his argument, that it is *in fact* the case that we standardly evaluate counterfactuals in a way that corresponds, *more or less*, to the temporally asymmetric similarity metric that Lewis called 'Analysis 1' [30,31,40].

Analysis 1 says that where the antecedent of a counterfactual is entirely about affairs in a stretch of time t_a , we standardly evaluate the counterfactual 'if it were that A, it would be that C' by considering all those worlds w where:

- (i) A is true;
- (ii) w is exactly like our actual world at all times before a transition period beginning *shortly before* t_a ;
- (iii) w conforms to the actual laws of nature at all times after t_a ;
- (iv) during t_a and the preceding transition period, w differs no more from our actual world than it must to permit A to hold.

And it says that the counterfactual 'if A, it would be that C' is true just in case C holds at every such world w [2].

The argument that Lewis gave in "Counterfactual Dependence and Time's Arrow" was based on his belief that Analysis 1 gets it right about a central class of counterfactuals—the ones that we evaluate by using what he called "the standard resolution for counterfactual vagueness"²¹. Lewis thought this because he made a mistaken diagnosis of our belief in time's arrow. He thought that our belief that time has an "arrow" or "direction" is the belief that the past is (given the standard resolution) *counterfactually* fixed. Lewis used examples that made this seem plausible but he was wrong. His mistake was at least partly due to his failure to consider an alternative theory of counterfactuals, one that holds the laws fixed.

This mistake was due to an earlier mistake, a mistake at the heart of the possible world semantics for counterfactuals that he presented in his 1973 book *Counterfactuals*.

This brings me to the Tooley/Wasserman argument.

4. Counterfactuals, Time Travel, and the Tooley/Wasserman Argument

In his recent book, *Paradoxes of Time Travel*, Ryan Wasserman defends an argument for a surprising claim [41], pp. 174–182: Either time travel to the past (and, more generally, backward causation) is metaphysically impossible or Lewis’s theory of counterfactuals is false and so too is Lewis’s, or anyone’s, counterfactual analysis of causation.

If Wasserman is right, this is potentially a problem for all of us and not just for Lewis. On the one hand, it seems that time travel to the past is conceptually coherent and metaphysically possible [5]²². Wasserman’s detailed examination and critique of the many attempts to argue otherwise supports this claim. However, if time travel to the past is metaphysically possible, then backward causation is also metaphysically possible. So any adequate philosophical account of causation must be one that accommodates the possibility that the causal relation holds in *both temporal directions*. To date, the only theories of causation that have any promise of doing this are counterfactual theories.

Wasserman’s argument is based on a case that Michael Tooley [42] once gave as a counterexample, to a very general feature of a possible worlds semantics for counterfactuals. More specifically, it is to the logic that Lewis defended in his 1973 book, *Counterfactuals* [43].

Tooley argued that if backward causation is possible, then no theory of counterfactuals that imposes an absolute ordering on worlds with respect to overall similarity is correct. Since this is a key feature of Lewis’s theory of counterfactuals, Lewis’s theory is false. Wasserman draws the further conclusion that no counterfactual analysis of causation can succeed.

Tooley’s case is a description of a world where it is stipulated that the laws permit time travel to the past as well as the future. It is intended as a counterexample to someone who believes both that time travel to the past is possible and that counterfactuals are evaluated the Lewis–Stalnaker way. This world has two time machines, one on Earth and one on Mars. Each machine is designed to serve a single purpose: travel to the other planet, disable the other time machine (by removing its fuel), and flip its blastoff switch. The Earth machine works by traveling one year to the future; the Mars machine by traveling one year into the past. During the time relevant to Tooley’s story—the time from t_1 to t_2 —none of this happens. The two machines remain on their respective planets, fueled and ready to go, but with blastoff switches not flipped.

Given the stipulated laws and facts of this world, these counterfactuals are both true:

Earth: If the Earth machine switch had been flipped to On at t_1 , the Mars machine would have been unfueled at t_2 .

Mars: If the Mars machine switch had been flipped to On at t_2 , the Earth machine would have been unfueled at t_1 .

Tooley notes that a possible worlds semantics for counterfactuals gets the right results for **Earth** and **Mars** only if its similarity metric tells us that the closest worlds (to the base world, W_0) where **Earth** is true are the W_1 worlds and the closest worlds where **Mars** is true are the W_2 worlds.

W₀ Earth switch off (Earth fueled) from t_1 to t_2 ; Mars switch off, (Mars fueled) from t_1 to t_2

W₁ Earth switch On (Earth fueled) at t_1 **Mars unfueled**, Mars switch on at t_2

W₂ Mars switch On (Mars fueled) at t_2 ... **Earth unfueled**, Earth switch On at t_1

It might seem that there is no problem, so far as the Lewis–Stalnaker semantics is concerned. The Lewis–Stalnaker semantics tells us that we evaluate a counterfactual by considering the closest (most similar) world or set of worlds where the *antecedent* is true, and this is compatible with the antecedents of some counterfactuals being true at worlds more distant than others. To evaluate **Earth**, we look at worlds where the Earth machine

switch is flipped to On at t_1 ; to evaluate **Mars**, we look at a different set of worlds: worlds where the Mars machine switch is flipped to On at t_2 .

However, there are two kinds of worlds where the antecedent of **Earth** is true and two kinds of worlds where the antecedent of **Mars** is true. The worlds where the antecedent of **Earth** is true include the W1 worlds. But they also include worlds where the antecedent is true because the Mars machine has traveled to Earth and flipped the switch of the Earth machine to On—the W2 worlds. Likewise, the worlds where the antecedent of **Mars** is true are not limited to the W2 worlds. They also include the W1 worlds where the Mars machine's switch is flipped to On by the Earth machine. Since the Lewis–Stalnaker semantics imposes an absolute *time-invariant* ranking on worlds, it cannot say what we intuitively want to say: that when we evaluate **Earth**, the closest worlds are the W1 worlds, and when we evaluate **Mars**, the closest worlds are the W2 worlds. We must choose. Either the closest worlds are the W1 worlds or they are the W2 worlds. So the Lewis–Stalnaker semantics must say that one of these two counterfactuals is false.

Wasserman's argument:

1. If backward causation and time travel to the past are possible, there is a world, W_0 , where **Earth** and **Mars** are both true.
2. **Earth** and **Mars** are both true only if counterfactuals are *not* evaluated according to overall similarity in an absolute ranking.
3. Lewis's theory of counterfactuals is committed to an absolute ranking of worlds according to overall similarity.
4. Therefore, if backward causation and time travel to the past are possible, Lewis's theory of counterfactuals is false.
5. If Lewis's theory of counterfactuals is false, no counterfactual theory of causation is true.

I'm persuaded that the Tooley case shows that if time travel to the past and backward causation are possible, Lewis's theory of counterfactuals is false. More precisely, any account of the truth conditions of counterfactuals that imposes an absolute *and time-invariant* ordering on worlds with respect to some metric of world similarity is false [44,45]²³. Because of this, Lewis's various attempts to defend a counterfactual analysis of causation all fail (because they assume the truth of Lewis's theory of counterfactuals). However, I reject the conclusion of Wasserman's argument. 5 does not follow from 1–4. Wasserman assumes that 5 follows because he assumes that the only alternative to a theory that says that worlds are ordered with respect to their overall similarity to ours is a causal theory that makes similarity depend on the content of the antecedent [30,41]²⁴ and we cannot, on pain of circularity, use a causal theory of counterfactuals to provide a counterfactual analysis of causation. However, there is another option. A similarity metric might rank worlds at least partly on the basis of their similarity with respect to particular facts *at the time of the antecedent*. There is a theory of counterfactuals which ranks worlds in this way. It is the Simple theory, defended by Jonathan Bennett in 1984 [33,36,38]²⁵.

The Simple theory says, roughly²⁶: The counterfactual *if A, it would be that C* is true iff C is true at all the legal A-worlds which most resemble the actual world at or during the time of A. A legal world is a world that has all and only the laws of the actual world. An A-world is a world where A is true.

5. What Is the Simple Theory and How Does It Differ from Lewis's Theory?

First, the Simple theory is a Law theory. It tells us that the closest counterfactual worlds are worlds with the same laws of nature as the actual world. No miracles, not even small divergence miracles.

Second, the Simple theory says that the closest worlds are those legal worlds that are most similar to the actual world not with respect to *all time*, as Lewis told us, but with respect to how things are during the time of the antecedent. This is consistent with an *absolute* (as opposed to *antecedent-relative*) ranking of worlds, and it is consistent with worlds

being ranked in the same ordering at different times. However, unlike Lewis's theory, the Simple theory does not insist that there is a single *time-invariant* ranking of worlds. That's why it's not refuted by the Tooley case.

Third, the Simple theory is not a causal theory; it does not require us to assume that laws are causal laws—it does not rule this out, but it also does not assume it. It does not tell us to maximize match of causally independent fact [30] and it does not prohibit counterfactual or causal reasoning in either temporal direction. It is therefore the kind of theory that is eligible for the purpose of a counterfactual analysis of causation.

Even if you think that no counterfactual analysis of causation can succeed²⁷, you still need a theory of counterfactuals, and you especially need a theory of counterfactuals if you want to think seriously about the free will/determinism problem [9,10]²⁸. So I am going to conclude by giving some reasons to prefer the Simple theory.

6. Why We Should Prefer the Simple Theory (and How Lewis Led Us Astray)

First, because it is simple and therefore can explain how we actually evaluate counterfactuals. No tradeoffs between miracles and the size of the spatiotemporal region of exact match of particular fact are required. All that is needed is some rough and ready knowledge of the laws and matters of particular fact at the time of the antecedent. In practice how things are locally will often suffice. Even a child can use the Simple theory.

Second, because it gets the right result in the Tooley case, and gets it easily. Tooley's story is complicated, but it comes with an easy set of directions for evaluating the relevant counterfactuals. The directions correspond to the Simple theory. We evaluate **Earth** by looking at worlds with the stipulated laws where the facts most closely match the facts of the base world at *t*₁. We evaluate **Mars** by looking at worlds with the stipulated laws where the facts most closely match the facts of the base world at *t*₂. That is all there is to it [46]²⁹.

Last, but not least, because it holds the laws fixed. This should be our default assumption (and always was, until Lewis came along). How were we ever persuaded otherwise? How did we come to agree with Lewis that we standardly evaluate counterfactuals by considering worlds with small miracles?

The answer is that we followed Lewis, step by step from his seminal *Counterfactuals* [43] to his boldly ambitious "Counterfactual Dependence and Time's Arrow" [2]. In *Counterfactuals* Lewis made two claims, both revolutionary.

He said, first, that we should understand counterfactual assertions as claims about what is true in a world or set of worlds that are similar, *in relevant respects*, to the actual world. This much I think we should all accept.

However, he also made a stronger claim. He said that we can make "overall comparisons" of the similarity of worlds in the way that we make overall comparisons of the similarity of cities and, moreover, that *there is an absolute and time-invariant ranking* of worlds with respect to their overall similarity to the actual world. He worked out a logic for counterfactuals that was based on this stronger assumption.

The Nixon counterfactual, offered by Kit Fine [47] and others, was a counterexample to this second claim. We can stipulate that determinism is true and the facts are such that if Nixon had pressed the button, there would have been a nuclear holocaust. Worlds with a nuclear holocaust are very different from our world (or so we hope). So Lewis must say that this counterfactual is false.

Lewis's response was a tour de force.

First, he swiftly backpedaled, telling us that he never meant to claim that we can compare worlds in the way we can compare cities. Rather, we should start with the counterfactuals that we believe are true and extract from these a theory of similarity, "not necessarily the first that comes to mind," a theory that validates our verdicts about clearly true counterfactuals.

He then proposed Analysis 1 and showed that it gets the right result for the Nixon counterfactual³⁰.

However, Analysis 1 is temporally asymmetric and Lewis wanted a theory of counterfactuals that would serve the purposes of a counterfactual analysis of causation without making it an analytic truth that counterfactual dependence is temporally (and thus causally) asymmetric.

He therefore set himself the task of showing how the temporally asymmetric Analysis 1 is the product of a *temporally neutral* similarity metric, a metric that preserved his initial thought that worlds can be ranked with respect to their similarity to the actual world not just at particular times, but *across all times*.

He argued that, at least in the case of the Nixon counterfactual, there is a temporally neutral similarity metric³¹ that together with a contingent but pervasive fact about our world—the so-called “asymmetry of overdetermination”—yields a contingent temporal asymmetry of miracles, which in turn yields the contingently temporally asymmetric Analysis 1.

This argument has been criticized [32,35,40]³² but the critics typically do not challenge Lewis’s starting point—that Analysis 1, or something similar, is at least contingently correct [48]³³. To concede this is to agree that we evaluate counterfactuals by a metric that permits events that are miraculous by the standards of our laws.

Did Lewis have any argument for Analysis 1 apart from his need to avoid the Nixon counterexample to his original account?

Remember: Before Lewis came along, everyone assumed that the laws are always cotenable with the antecedent of a counterfactual. In possible worlds language, they assumed that the most similar worlds are worlds where what happens always conforms to the actual laws.

Lewis had two arguments.

The first argument was the time’s arrow argument.

Lewis claimed that our belief in time’s arrow (or, at least, that part of it that consists in the belief that the past is at least contingently “fixed” in some way that contrasts with the way that the future is “open”) is neither a belief about time itself nor a belief about the falsity of determinism. He proposed that it is a belief about a *temporal asymmetry of counterfactuals*. It is often true, he claimed, that if the present were different, the future would be different, in definite and detailed ways, but it is “seldom, if ever” true that if the present were different, the past would have been different in any definite and detailed way [2]³⁴. From this interesting and plausibly true claim [49]³⁵ he swiftly (and without argument) moved to a very different claim: that it is often true that if the present were different, the past (or most of the past), *would still have been exactly the same*³⁶. He then argued that this shows that our de facto theory of counterfactuals is the temporally asymmetric Analysis 1.

However, (as Jonathan Bennett has pointed out) [33] our belief in time’s arrow is not our belief in Analysis 1. Analysis 1 has a transition period [2]³⁷, and for some counterfactuals, this period extends a long time into the past. If Kennedy had run for a second term as President, he would have to have not been killed in Dallas in 1963. Our time’s arrow belief—whatever it is—does not make an exception for any part of the past. So Analysis 1 cannot be what explains our belief in time’s arrow.

The second argument was an epistemic argument. Lewis briefly considered the possibility that a theory of counterfactuals that holds the laws fixed might also get the right verdict about the Nixon counterfactual. He dismissed this on the grounds that if determinism is true, this means that the closest worlds are different all the way back to the remote past and this kind of “backtracking unlimited” makes it difficult (if not impossible) to figure out which counterfactuals are true.

In making this argument, Lewis was ruling out the Simple theory. The Simple theory says that the closest worlds are those worlds with our laws which are *most like our world at the time of the antecedent*. Determinism means that any such world would have to have a different history from ours all the way back to the Big Bang. However, the Simple theory does not require us to know or guess what those alternate histories might be. All that it requires is that we are able to judge—as we surely are able to judge—ways the *present*

might be different that are nomologically possible. The Simple theory is as well equipped as Lewis's Analysis 1 (and arguably better equipped than his temporally neutral similarity metric) to give us epistemic guidance in our evaluation of counterfactuals [28,40,48]³⁸.

There was a reason that Lewis did not consider the Simple theory. He assumed that the similarity metric for counterfactuals must yield a *time-invariant* ranking of worlds (as he had proposed in the 1973 book, and continued to defend in 1979). But the Tooley case shows that such a ranking is hostage to the possibility of time travel to the past and/or backward causation. If backward causation is metaphysically or nomologically possible, then it is at best only contingently true that there is such a ranking of worlds. So this argument also fails.

7. Conclusions

The Consequence argument is an argument from plausible premises—our lack of causal power with respect to the laws and past—to an implausible conclusion: that if determinism is true, we are *equally* powerless with respect to the future. To give a fully satisfactory reply to the argument, the compatibilist must explain *how* the premises can be true while the conclusion is false. The right answer is that we lack causal power over the past and the laws. Whatever it is that we are able to do, it would not include *causing* a difference—of any kind, big or small, in our control or not—to the past or to the laws. However, the fact that we lack causal power over the past and laws does not mean that we lack causal power over the present and future.

The problem, for the compatibilist, is that causation and counterfactual dependence typically go hand in hand. Overdetermination and causal pre-emption are the exception rather than the rule. It is a consequence of determinism that either the past or the laws (or both) *counterfactually depend* on our present actions. If determinism is true and I were doing anything other than what I'm actually doing right now, then either the past, the laws, or both, would have been different (would have to have been different). The compatibilist must explain why *this* counterfactual dependence is not a causal dependence, and to do this she needs a theory of counterfactuals that does not make it true, *by stipulation*, that the only true causal counterfactuals are ones whose antecedent is about a time *earlier* than the consequent. Lewis offered such a theory. This theory is vulnerable to many criticisms. I have argued that it goes wrong in an even more fundamental way than most of the literature has noticed, a way that makes it impossible to use Lewis's theory of counterfactuals (or any theory that shares that feature) as part of an answer to the Consequence argument. What we need, and what Lewis did not deliver, is a theory of counterfactuals that preserves the links, not just between counterfactuals and causation, but also between counterfactuals and *laws*, in a way that allows us to understand how causation is possible at a deterministic world that might, for all we know, be our world. I believe that the Simple theory has the resources to do the job. But that will have to be a task for another day [36,38,50]³⁹.

Funding: This research received no external funding.

Acknowledgments: Thanks to the three anonymous referees for this journal. This paper began life as part of a Book Symposium on Ryan Wasserman's *The Paradoxes of Time Travel* at the 2019 meetings of the Pacific APA. I presented a version closer to the present one at the Free Will and Physics Zoom workshop during the first pandemic summer, in 2020. My thanks to all who participated, and especially to Christian Loew, Barry Loewer, and Terrance Tomkow.

Conflicts of Interest: The author declares no conflict of interest.

Notes

- ¹ "At least" because some compatibilists—I call them 'commonsense compatibilists'—also claim that determinism is compatible with our ability to choose, do, and cause other than what we actually choose, do, and cause. I'm one of those compatibilists.
- ² Lewis 1979, p. 32 [2] "A deterministic system of laws is one such that, whenever two possible worlds both obey the laws perfectly, then either they are exactly alike throughout all of time, or else they are not exactly alike through any stretch of time. They are alike always, or never." See also Earman 1986, pp. 12–14 [3].

Beginning with Lewis 1976 [5].

Taylor, 1992: “A fatalist . . . is someone who believes that whatever happens is and always was unavoidable . . . A fatalist thus thinks of the future in the way we all think of the past” [11].

van Inwagen, 1983. See below for a statement of this version of the argument [7].

“I *could* have walked a mile in twenty minutes this morning, but I certainly could *not* have run two miles in five minutes. I did not, *in fact*, do either of these two things; but it is pure nonsense to say that the mere fact that I *did* not, does away with the distinction between them, which I express by saying that the one *was* within my powers, whereas the other was *not*.” (Moore 1912) [12].

Unfortunately, these attempts took place at a time when the relevant conditionals were poorly understood, just before the development of the Lewis-Stalnaker possible worlds semantics for counterfactuals.

Chisholm 1964, Lehrer 1968. In Vihvelin 2013, I argue that this verdict was a mistake, due in large part to a failure to understand counterfactual conditionals [8,16,17].

See, for instance, van Inwagen 1974, Lamb 1977 [13,14].

The standard definition of determinism is in terms of entailment: A complete description of the state of the world at any time and a complete specification of the laws of nature together entail every truth about what happens at every other time. Something not usually noticed in the free will literature is that this entailment relation goes in both temporal directions.

“Causation is a morass in which I for one refuse to set foot. Or not unless I am pushed.” (van Inwagen 1983, p. 65 [7]).

Or, at any rate, it seems uncontroversially true given the assumption that there is *in fact* no time travel to the past or other variety of backwards causation. It’s worth noting that the question of the possibility of backwards causation rarely comes up in the literature on the free will/determinism problem; I suspect that this is because it is (in that literature at least) widely taken for granted that causation is some kind of nonHumean relation that is necessarily temporally asymmetric.

Rule Alpha is almost universally accepted, but see Spencer 2017 for an argument that we should reject it. Nguyen 2020 criticizes Spencer’s argument [23,24].

I argue this in Vihvelin 2013 [8].

At least not given the facts as we currently believe them to be. If time travel and backwards causation are metaphysically or nomologically possible, then things are more complicated. I will discuss this later.

Lewis 1981. Lewis’s response, published two years before van Inwagen’s book, was not directed at the rule Beta version of the argument. But the essence of his criticism, which was also made by Fischer 1983, Watson 1987, Vihvelin 2013, and others, is that the argument equivocates in this way [6,8,25,26].

I think it was because Lewis did not acknowledge this that he was vulnerable to the critique in Beebe 2003. A referee noted that in the time travel paper Lewis did address this point. (“Success at some tasks requires not only ability but also luck and lack of luck is not a temporary lack of ability”, Lewis 1976, p. 76.) That’s right, but this doesn’t affect my point because while everyone in the free will literature knows Lewis 1981, not many are familiar with Lewis 1976 [5,6,27].

But not everyone. Notable exceptions include Roberts 2008 and Lange 2009 [28,29].

For instance, the theory defended by Schaffer 2004 or Bennett 2003 [30,31].

We will see that Lewis didn’t—and couldn’t—claim that the past is counterfactually fixed. But this is the impression that he gave, and that many believe.

Lewis introduced the contrast between the standard and non-standard resolution in response to the objection that we don’t always evaluate counterfactuals according to Analysis 1.

Lewis 1976 remains the classic defense of the metaphysical possibility of time travel [5].

Noordhof 2003 argues that Tooley’s case can be reconciled with a Stalnaker-Lewis account of the truth-conditions of counterfactuals if the similarity relation is allowed to be antecedent-relative. Cross 2008 argues that taking similarity to be antecedent-relative results in a formal semantics which is a comparative world similarity semantics in name only [44,45].

For instance, a theory that says that the closest worlds where the antecedent is true are worlds which are most like the actual world with respect to all facts that are *causally independent* of the antecedent. (Schaffer 2004) [30]. Wasserman defends this kind of theory on pp. 176–182 [41].

Bennett 1984. Bennett later gave up this theory for a theory more like Lewis’s Analysis 1, but I don’t think he had any good reasons for doing so. For defense of the Simple theory, including responses to the objections that persuaded Bennett to give up the Simple theory, see Tomkow 2013 and Tomkow and Vihvelin 2017 [33,36,38].

Roughly, because there is a modification to accommodate counterlegals. Since these are not relevant for my purposes, I’ll use the simpler version.

My view is that it’s a mistake to think that there is a conflict between counterfactual and law-entailment theories of causation: a theory of causation must make sense of the links between causation and law as well as causation and counterfactuals.

And, I would argue, about related freedom problems that arise in time travel scenarios. I discuss some of these problems in Vihvelin 1996 and 2020 [9,10].

- In this way, Tooley's case resembles Schaffer's famous "trumping" counterexample to Lewis's analysis of causation as the ancestral of counterfactual dependence. What made that case work was that Schaffer told us a law of magic that we can easily understand and apply despite the fact that the world he described is a world very different from our own. (Schaffer 2000) [46].
- It's worth noting that the Nixon counterfactual is a counterfactual about a person who makes what we assume is a free choice. Lewis made no attempt, in that paper or elsewhere, to defend the claim that Analysis 1 gets the right results for other counterfactuals.
- Roughly: the closest antecedent-worlds are worlds that most resemble the actual world with respect to exact match of particular fact—counting past and future resemblances equally—provided that they differ from the actual world by the occurrence of the smallest miracle required for the truth of the antecedent.
- As I indicated earlier, cf. Horwich 1987, Elga 2000, Loewer 2018. [34,35,40].
- Wilson 2014 is an exception. Her critique is directed against the miracle-permitting temporally asymmetric Analysis 1 rather than Lewis's argument for the claim that there is a temporally neutral similarity metric that combines with a contingent fact about the world to yield Analysis 1 [48].
- "Today I am typing words on a page. Suppose today were different. Suppose I were typing different words. Then plainly tomorrow would be different also . . . Would yesterday also be different? If so, how? Invited to answer, you will perhaps come up with something. But I do not think that there is anything you can say about how yesterday would be that is clearly and uncontroversially true." (Lewis 1979, p. 32) [2].
- In Tomkow and Vihvelin 2022 we argue that a qualified form of this counterfactual temporal asymmetry is what accounts for the contingent temporal asymmetry of causation [49].
- Lewis *ibid.*, p. 33.
- This point is worth stressing since it isn't always noticed. Lewis is committed to the claim that there are true backtracking counterfactuals given what he calls "the standard resolution" of counterfactual vagueness. For instance, Lewis says, of a match that is a foot away from a striking surface at *t*, that if the match had been struck at *t*, the past shortly before *t* would have been different in some way that avoids "abrupt discontinuities." He needs to say this to avoid being committed to the truth of "if the match had been struck at *t*, it would have traveled a foot in no time at all." (Lewis 1979, p. 40) [2].
- For other arguments against Lewis's claim that holding the laws fixed would make it impossible for us to have the counterfactual knowledge we actually have, see Bennett 2003, 202–209, Roberts 2008, 238–241, and Wilson 2014 [28,40,47].
- For some steps in that direction, see Tomkow 2013, Tomkow and Vihvelin 2017, and Tomkow and Vihvelin 2020 [36,38,50].

References

- Fischer, J.; Ravizza, M. *Responsibility and Control: A Theory of Moral Responsibility*; Cambridge University Press: Cambridge, UK, 1998.
- Lewis, D. Counterfactual Dependence and Time's Arrow. *Nous* **1979**, *13*, 455–476. [CrossRef]
- Earman, J. *A Primer on Determinism*; Reidel Publishing Company: Dordrecht, The Netherlands, 1986.
- Arntzenius, F. Time Travel: Double Your Fun. *Philos. Compass* **2006**, *1*, 599–616. [CrossRef]
- Lewis, D. The Paradoxes of Time Travel. *Am. Philos. Q.* **1976**, *13*, 145–152.
- Lewis, D. Are We Free to Break the Laws? *Theoria* **1981**, *47*, 113–121. [CrossRef]
- van Inwagen, P. *An Essay on Free Will*; Clarendon Press: Oxford, UK, 1983.
- Vihvelin, K. *Causes, Laws, and Free Will: Why Determinism Doesn't Matter*; Oxford University Press: New York, NY, USA, 2013.
- Vihvelin, K. What Time Travelers Cannot Do. *Philos. Stud.* **1996**, *81*, 315–330. [CrossRef]
- Vihvelin, K. Killing Time Again. *Monist* **2020**, *103*, 312–327. [CrossRef]
- Taylor, R. Fate. In *Metaphysics*; Prentice-Hall: Englewood Cliff, NJ, USA, 1992.
- Moore, G.E. Free Will. In *Ethics*; Oxford University Press: Oxford, UK, 1912.
- van Inwagen, P. A Formal Approach to the Problem of Free Will and Determinism. *Theoria* **1974**, *40*, 9–22. [CrossRef]
- Lamb, J. On a Proof of Incompatibilism. *Philos. Rev.* **1977**, *86*, 20–35. [CrossRef]
- Ginet, C. *On Action*; Cambridge University Press: Cambridge, UK, 1990.
- Chisholm, R. Human Freedom and the Self. In *The Lindley Lecture*; University of Kansas: Lawrence, KS, USA, 1964.
- Lehrer, K. Cans without Ifs. *Analysis* **1968**, *29*, 29–32. [CrossRef]
- Strawson, P. Freedom and Resentment. *Proc. Br. Acad.* **1962**, *68*, 1–25.
- Frankfurt, H.G. Alternate Possibilities and Moral Responsibility. *J. Philos.* **1969**, *66*, 820–839. [CrossRef]
- Fischer, J.M. *The Metaphysics of Free Will*; Blackwell: Cambridge, MA, USA, 1994.
- Kane, R. *The Significance of Free Will*; Oxford University Press: Oxford, UK, 1996.
- O'Connor, T. *Persons and Causes: The Metaphysics of Free Will*; Oxford University Press: Oxford, UK, 2000.
- Spencer, J. Able to Do the Impossible. *Mind* **2017**, *126*, 466–497. [CrossRef]
- Nguyen, A. Unable to do the Impossible. *Mind* **2020**, *129*, 585–602. [CrossRef]
- Fischer, J. Incompatibilism. *Philos. Stud.* **1983**, *43*, 121–137. [CrossRef]
- Watson, G. Free Action and Free Will. *Mind* **1987**, *96*, 145–172. [CrossRef]

27. Beebe, H. Local Miracle Compatibilism. *Nous* **2003**, *37*, 258–277. [[CrossRef](#)]
28. Roberts, J.T. *The Law-Governed Universe*; Oxford University Press: Oxford, UK, 2008.
29. Lange, M. *Laws and Lawmakers: Science, Metaphysics, and the Laws of Nature*; Oxford University Press: New York, NY, USA, 2009.
30. Schaffer, J. Counterfactuals, Causal Independence, and Conceptual Circularity. *Analysis* **2004**, *64*, 299–309. [[CrossRef](#)]
31. Bennett, J. *A Philosophical Guide to Conditionals*; Clarendon Press: Oxford, UK, 2003.
32. Lewis, D. Causation. *J. Philos.* **1973**, *70*, 556–567. [[CrossRef](#)]
33. Bennett, J. Counterfactuals and Temporal Direction. *Philos. Rev.* **1984**, *93*, 57–91. [[CrossRef](#)]
34. Horwich, P. *Asymmetries in Time*; MIT Press: Cambridge, MA, USA, 1987.
35. Elga, A. Statistical Mechanics and the Asymmetry of Counterfactual Dependence. *Philos. Sci.* **2000**, *68*, 313–324. [[CrossRef](#)]
36. Tomkow, T. The Simple Theory of Counterfactuals. 2013. Available online: www.tomkow.com (accessed on 12 July 2023).
37. Dorr, C. Against Counterfactual Miracles. *Philos. Rev.* **2016**, *125*, 241–286. [[CrossRef](#)]
38. Tomkow, T.; Vihvelin, K. The Temporal Asymmetry of Counterfactuals. 2017. Available online: www.vihvelin.com (accessed on 12 July 2023).
39. Vihvelin, K. How the Laws Constrain: Causation, Counterfactuals, and Free Will. In *Free Will and Causation, Philosophie de la Connaissance*; Tiercelin, C., Ed.; College de France Press: Paris, France, 2017.
40. Loewer, B. The Mentaculus: A Probability Map of the Universe. In *The Probability Map of the Universe: Essays on David Albert's Time and Chance*; Loewer, B., Weslake, B., Winsberg, E., Eds.; Harvard University Press: Dordrecht, The Netherlands, 2023.
41. Wasserman, R. *The Paradoxes of Time Travel*; Oxford University Press: Oxford, UK, 2018.
42. Tooley, M. Backwards Causation and the Stalnaker-Lewis Approach to Counterfactuals. *Analysis* **2002**, *62*, 191–197. [[CrossRef](#)]
43. Lewis, D. *Counterfactuals*; Harvard University Press: Cambridge, MA, USA, 1973.
44. Noordhof, P. Tooley on Backward Causation. *Analysis* **2003**, *63*, 157–162. [[CrossRef](#)]
45. Cross, C. Antecedent Relative Comparative World Similarity. *J. Philos. Log.* **2008**, *37*, 101–120. [[CrossRef](#)]
46. Schaffer, J. Trumping Preemption. *J. Philos.* **2000**, *97*, 165–181. [[CrossRef](#)]
47. Fine, K. Review of Lewis's Counterfactuals. *Mind* **1975**, *84*, 451–458. [[CrossRef](#)]
48. Wilson, J. Hume's Dictum and the Asymmetry of Counterfactual Dependence. In *Chance and Temporal Asymmetry*; Wilson, A., Ed.; Oxford University Press: Oxford, UK, 2014.
49. Tomkow, T.; Vihvelin, K. Why We Cannot Change the Past. 2022. Available online: www.vihvelin.com (accessed on 12 July 2023).
50. Tomkow, T.; Vihvelin, K. Causation. 2020. Available online: www.vihvelin.com (accessed on 12 July 2023).

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.