frontiers in Physics

PERSPECTIVE
published: 14 December 2017
doi: 10.3389/fphy.2017.00064

Can Physics Make Us Free?

Tuomas K. Pernu [1,2]*

[1] *Department of Philosophy, King's College London, London, United Kingdom,* [2] *Department of Biosciences, University of Helsinki, Helsinki, Finland*

A thoroughly physical view on reality and our common sense view on agency and free will seem to be in a direct conflict with each other: if everything that happens is determined by prior physical events, so too are all our actions and conscious decisions; you have no choice but to do what you are destined to do. Although this way of thinking has intuitive appeal, and a long history, it has recently began to gain critical attention. A number of arguments have been raised in defense of the idea that our will could be genuinely free even if the universe is governed by deterministic laws of physics. Determinism and free will have been argued to be compatible before, of course, but these recent arguments seem to take a new step in that they are relying on a more profound and concrete view on the central elements of the issue, the fundamental laws of physics and the nature of causal explanation in particular. The basic idea of this approach is reviewed in here, and it is shown how a careful analysis of physics and causal explanation can indeed enhance our understanding of the issue. Although it cannot be concluded that the problem of free will would now be completely solved (or dissolved), it is clear that these recent developments can bring significant advancement to the debate.

Keywords: Block Universe, compatibilism, consequence argument, determinism, difference-making, free will, incompatibilism, interventionism

Edited by:
Alex Hansen,
Norwegian University of Science and Technology, Norway

Reviewed by:
Lamberto Rondoni,
Politecnico di Torino, Italy
Eirik Grude Flekkøy,
University of Oslo, Norway

***Correspondence:**
Tuomas K. Pernu
tuomas.pernu@kcl.ac.uk

Specialty section:
This article was submitted to Interdisciplinary Physics, a section of the journal Frontiers in Physics

Received: *06 June 2017*
Accepted: *20 November 2017*
Published: *14 December 2017*

Citation:
Pernu TK (2017) Can Physics Make Us Free? Front. Phys. 5:64. doi: 10.3389/fphy.2017.00064

PRELUDE

Physics is typically thought to paint a bleak picture of human existence. All there is, is a myriad of soulless particles, aimlessly bouncing on each other in an infinite darkness. We, humans, and all our actions, are nothing but products of such mechanistic interactions: everything we are, and everything we do, is dictated by the behavior of simple material particles and the blind laws that govern it. What we think, what we feel, and what we choose does not have an effect on the world; we don't really matter.

Recently, however, this view has began to receive growing critical attention [1–12]. Ismael [6] in particular, a prominent philosopher of physics, has composed a book-length argument to challenge our intuitive bleak view. Her aims in *How Physics Makes Us Free* are ambitious: what she sets out to do, in a nutshell, is to explain, not only how a proper understanding of physics does not force such a bleak view of reality and ourselves on us, but how the thoroughly physical world and the laws that govern its behavior actually enable us to be autonomous agents who are free to act according to their conscious decisions.

Compatibilism—the idea that physical determinism and free will are ultimately compatible—is not a new thesis, of course, but these recent arguments seem to take a new step in that they are relying on a more profound and detailed view on the central elements of the issue: the fundamental laws of physics and the nature of causal explanation. The following reviews the basic idea of this approach. It is shown how a careful analysis of causal explanation in the context of our current knowledge of physics can indeed enhance our understanding of the issue. Although it cannot be concluded that the problem of free will would now be completely solved (or dissolved), it is clear that these recent developments can bring significant advancement to the debate.

FREE WILL AND AGENCY

Free will is a notoriously multifaceted and elusive notion. Most treatments of the issue, even purely philosophical ones, rush directly into deciphering the intricacies of determinism and indeterminism, and their respective connections to the idea of autonomous agency. However, it should be rather obvious that before there can be such a thing as autonomous agency, which may or may not be said to include the power to act according to free choices, there needs to be an idea of who, or what, is supposed to be the bearer of such an autonomous agency. To get clear on such an idea is not a trivial undertaking.

The main thrust of the current discussion is naturalism: the aim is to show how our common sense notions of agency and free will can be reconciled with a thoroughly scientific world view. Both dualism and eliminativism are typically rejected. This leaves a rather narrow path to tread: on the one hand the aim is to hold on to our common sense notions, or at least to the core spirit of them, but on the other hand one is committed to accepting the results of physics and other empirical sciences, without a murmur. Ismael [6] defends an account of the self that neither assumes the existence of immaterial souls nor accepts "nolipsism," an eliminativist "no-self view of the self" (cf. also [13]). What's left in-between, Ismael [6] contends, is an account of the self as "self-governor": a view that leaves higher-level conscious control intact, but gets rid of the homunculus-like inner unity of the self that our naïve common sense easily presupposes.

Filling in the details of the story of how such self-governors arise, and giving a full account of what kind of an entity it ultimately is, is a huge undertaking, that is bound to remain largely speculative. However, even if the current ideas are thus inevitably incomplete, it is not hard to find them credible, or at least pointing to the right direction. Various levels and aspects of agency are recognized and studied by biology, psychology, and other empirical sciences, and even though there is still a lot to find out and explain, there is little doubt that all of this will be fully consistent with the basic physical understanding of the world.

However, using the term "self-governor" does whet one's appetite for more specific information on what these thoroughly biological governors are supposed to govern, and, more burningly: how. To remove the air of mysticism around the notion one needs to substantiate it with a solid, and completely natural account of free will.

DETERMINISM, INDETERMINISM, AND "THE REAL PROBLEM"

The first thing to note is that this discussion is taking place in a perfectly classical, deterministic setting. Now, one could point out, and rightly so, that the most basic physical level is quantum mechanical, which in turn entails that physics is fundamentally indeterministic, and that it would therefore be a mistake to discuss the issue of free will in a deterministic setting.

There are many things to be said about this objection. First, it is not at all self-evident what precisely the determinism-indeterminism dichotomy amounts to, especially in the current context (e.g., [4, 14, 15]). It is quite well-know that even Newtonian mechanics allows the possibility of indeterministic dynamics (e.g., [4, 14, 16, 17]). Additionally, both special and general relativity leave room for indeterminism—the former case is maybe not that widely known (cf. [14]), but the latter case is generally accepted due to general relativity allowing the possibility of (naked) spacetime singularities [4, 14, 18].

But more importantly, when you move from the classical realm to quantum physics, things don't become any more straight forward. Although it is the typical, received view that quantum mechanics is fundamentally indeterministic—to the point where quantum indeterminism is thought to give us the only real basis for the idea of objective probability—the issue is far from settled. According to the traditional "Copenhagen Interpretation" of quantum mechanics elementary particles lack definite properties prior to their measurement ([19–22]; cf. [23]). What's striking about this view, as the argument presented by Einstein et al. [24] made apparent, is that this interpretation of quantum mechanics violates the locality postulate of special relativity (the idea that no signal can travel faster than light). Objectively indeterministic quantum physics would thus lead to postulating "spooky action at a distance," which is absurd, according to Einstein et al. [24], and we should therefore reject the idea that the indeterministic interpretation of quantum mechanics gives us a complete theory of reality.

However, the theoretical results of Bell [25–27], and their subsequent empirical tests [28, 29] have shown us that no local hidden-variable theory—i.e., no locally deterministic theory—will agree with the empirical implications of quantum mechanics, or, conversely, as far as any such theory agrees with the implications, it will not be local. In other words, our current theoretical and empirical understanding of quantum mechanics suggests that if there exists a deterministic formulation of quantum mechanics, it will be non-local. And therefore, as long as one is prepared to accept non-locality (something that Einstein et al., [24] thought was absurd), it is perfectly consistent, both theoretically and empirically, to give up on the idea of indeterminism at the quantum level. And in fact, Bell himself (cf. e.g., [30]) was a proponent of the Bohmian interpretation of quantum mechanics—a variant of non-local hidden-variable theory—and Bohm's [31, 32] non-local theory, which was generally neglected at the time of its introduction due to its very non-local nature, and served as an impetus for Bell [25, 26] to derive his results. So even on the traditional (Copenhagen) interpretation of quantum mechanics can be deterministic—at the price of locality.

Moreover, and more to the point, Bell [33–35] also admitted that even local hidden-variable formulations of quantum mechanics could still be consistent with his theoretical results. Such formulations would just need to assume "superdeterminism," as he called it [33–35]. Now, this is a crucial, and often neglected issue in the foundations of quantum physics. The traditional understanding of quantum mechanics assumes that the experimenter is free to make any given measurement. And given this assumption, the indeterministic dynamics of the elementary particles follows (as the recent theoretical results of Conway and Kochen [36, 37] make explicit). However, why should one be justified in making such an assumption? Why couldn't one deny that the experimenter is free to choose her measurement, and claim that that event, like everything else in the universe, was determined—or "superdetermined," as Bell [33–35] would say—to happen? And more burningly: isn't this the very issue that we're faced with in relation to the problem of determinism and free will? It is true that not many are prepared to accept such "superdeterministic" quantum physics, mainly because it seems to be a non-negotiable basic assumption of empirical science that the experimenter is free to test the reality (e.g., [38, 39]). However, at the same time it must be acknowledged that it is clear that if the whole idea of agential freedom is questioned, one cannot expect to give a satisfactory answer to the question by simply assuming such freedom. And this, ultimately, is the reason why the discussions on free will and physics are typically assuming a deterministic—i.e., a "superdeterministic"—framework.

However, given that such "superdeterminism" is typically rejected in physics, it is instructive to rehearse the main reasons for doubting the relevance of quantum physics in this context (cf. [40]). According to the naturalistic view on the mind, all our mental states and processes, our conscious decisions among them, are grounded on electro-chemical processes in neurons and neural networks. Although nerve cells seem tiny to us, they are huge on a quantum scale. True, neurons are composed of particles at smaller physical scales, and ultimately of things at quantum scale. However, the relevant level to consider here is the level of neuron, and neuron to neuron communication; the initiation and propagation of action potentials—the basis of all mental processes—are macrophysical events, perfectly explainable in classical terms [41]. This is the main physical reason to suppose that quantum physics can be safely ignored in neuroscientific and psychological contexts.

To be clear, it is true that quantum indeterminacies, supposing that they are objective in the first place, cannot be completely ignored at macroscopic scales. After all, radioactive decay can be legitimately pointed to as the cause of injury to biological organisms (which is something that Schrödinger's [42] famous cat thought-experiment dramatizes). However, it is clear that the current neurosciences do not trade on quantum indeterminacies. But more importantly, even if they would, it is all but clear how that would affect the problem of free will. The fundamental problem is that the idea of free will does not seem to be incompatible only with determinism, but also with indeterminism: random events are not willed. And this, to be clear, is yet another, purely conceptual, reason to be skeptical of the appeals to quantum physics in this context: replacing the determinism of classical physics with the indeterminacy of quantum physics would not, in itself, take us any step closer to solving the problem of free will. It is thus important to realize that the issue we are faced with the problem of free will is multifaceted in a very profound way. The core issue can be analyzed to divide itself into two separate problems, each of which creates a conflict with the purely physical view on the world in its own way [43]: the problem of the causal efficacy of the mental—"the problem of the will"—and the problem of the freedom of the will. Conflating these two problems can be seen to be a source of many confusions and false hopes surrounding the issue of free will. This, the following will suggest, is ultimately also the case with respect to the recent arguments inspired by a more concrete understanding of modern physics.

Let us now dub this "The Hard" or "Real Problem of Free Will": free will presupposes both determinism and indeterminism, but determinism and indeterminism are mutually exclusive, and hence "free will" is an oxymoron. The compatibilists are eager exploit this incoherence for their benefit: not only does it show that the contrary to determinism is equally detrimental to free will, but also how determinism is actually a friend, rather than a foe of it. This is actually the main reason why many are perfectly happy to settle with classical physics in this debate: in its determinism it provides grounds for the idea of wilful causal determination, which is something that we need to tackle the problem of the causal efficacy of the mental. However, even if one would grant that compatibilism and incompatibilism are on a par with each other—the former opting for compatibilism with respect to determinism, and the latter with respect to indeterminism—this in itself does not make the threat of determinism go away: the apparent incompatibility of free will and determinism remains, and the compatibilists will still need to give us a positive story on how the two are supposed to fit together.

CAUSATION AND FREE WILL

The conflict between determinism and free will—or between determinism and the freedom of the will, to be precise—can be made tangible by the means of the consequence argument—an explicit argument for the incompatibility of free will and determinism [44, 45]. The incompatibility arises, according to the argument, because our actions are entailed by the conjunction of the initial conditions of the universe and the deterministic laws of nature, and if we were to decide to act differently, one of the conjuncts (or both) would need to be changed (to avoid contradiction). So the idea that you would be able to act according to your free choices leads to the absurdity that you would be able to change the laws of nature or the initial conditions of the universe (or both).

Let us formulate this argument more precisely. Let A be an action performed at a time t_1, let S be a sentence that expresses the complete state of the universe in a distant point in the past t_0, let L be the set of all the fundamental laws of physics, and let determinism be the thesis that the conjunction of S and L entails every event in the universe. Now we can argue:

(1) No agent can perform $\neg S$ or $\neg L$ at time t_1. (Assumption)
(2) An agent performs A at time t_1. (Assumption)
(3) An agent could have performed $\neg A$ at time t_1. (Assumption/Free will)
(4) $(S \wedge L) \Rightarrow A$ (Assumption/Determinism)
(5) $\neg A \Rightarrow \neg(S \wedge L)$ (Contraposition from 4)
(6) An agent could have performed $\neg(S \wedge L)$ at time t_1. (Substituting 5 to 3)
(7) 6 contradicts 1. (1 and 6)
(8) 3 is false.

One way to approach this issue in the context of physics is to eliminate causation altogether—on the grounds that there is no room for such an asymmetric notion in fundamental physics—and then claim that there is no real problem of our decisions being causally effective because there is no such thing as causation in the first place. However, if one is determined to tread the narrow path of naturalism, one is also bound to hold on to the notion of causation—in particular, the idea that we are, or can be at times at least, genuine sources of our actions. The way to do this, the idea is, is to adopt a precise—particularly "scientific," as many are eager to stress—view on causation, that treats causal claims as claims about hypothetical results in counterfactual scenarios (e.g., [46–48]). The basic idea, very concisely, is that you first specify a set of variables (e.g., $\{C, E\}$) and a set of structural equations defining the relationships of these variables [e.g., if ($C = 1$, then $E = 1$) and (if $C = 0$, then $E = 0$)], and you then identify causal relations with the pattern of correlations you will find among the variables under hypothetical changes in the values of the variables; C is a cause of E, for example, just in case interventions on the value of C will result in changes in the value of E.

Although there are various specifications to be made, and details to be filled in—and problems to tackle—this is currently a widely accepted way to approach causation, and corresponds well to the way causal inference works in everyday science. However, even if this view of causation is not taken to be particularly problematic, it is a whole other issue how it should be applied to the age-old problem of free will. What Ismael [6] tries to do is, first, to define the notion of "pivotal role" in terms of interventions that serve a particularly effective—"pivotal"—role in a given domain, and second, to identify free will with such a pivotal role. This view could be seen to find support from Lagnado et al. [49], who provide an illuminating empirical account on how pivotal control connects to responsibility attributions, and moral responsibility, in turn, is arguably something that we are tracking with our free will attributions. Steward [11] offers a closely related account in which free will is connected to thermodynamically open systems that are "true producers." On this view, what free will therefore is, in effect, is "pivotal control," and a person can be said to have such "pivotal control over a certain kind of behavior just in case decision plays a pivotal role in its production" ([6], p. 239). That causal sources should be identified with such pivotal "control variables" [50], "difference-makers" [51–55], "handles" [56–58] or "necessary conditions" [12] is well-entrenched in the current discussion, and it could be argued that the seeds to this view were sown already by Mackie [59, 60] who defined causes as INUS conditions: *insufficient* but *necessary* parts of conditions which are itself *unnecessary* but *sufficient* for their effects.

So far so good. It is not too difficult to see how in certain places of the vast network of physical interactions you can have nexuses—pivotal areas—that channel some of the interactions in a particularly salient way, and it is not too big a stretch to identify conscious decision making with a certain subset of such nexuses, given that you are in the business of trying to give a perfectly naturalistic explanation of such things. This sounds like a coherent and credible picture. However, it is a further question whether this offers a new solution to the problem of free will, and most importantly, whether such a solution is ultimately successful.

The problem is that when you scratch the surface, it starts to look like the fundamental issues of the free will debate are left largely untouched. Most of the ammunition in the discussion is spent on giving a naturalistic account of decision making and action production. However, whether our decisions are truly effective, and whether their effectiveness can be explained in naturalistic terms, is not the core question of free will—the incompatibilists have no quarrel with the idea that we act according to conscious decisions, even if determinism holds. Rather, the core question is this: are we *free* to make the decisions according to which we act? In other words—and this is not a completely innocent reformulation, to be clear—are we faced with genuine alternatives when we make our decisions, and could we have chosen to act otherwise in cases where we happened to act according to our conscious decisions [premise (3) above]? It seems that the problem of the freedom of the will has become conflated with the problem of the causal efficacy of the will.

This critique can be made more precise in places. For example, it seems that Ismael [6] is simply embracing the conclusion of the consequence argument—that to her it is simply not absurd to think that we can change either the initial conditions of the universe or the laws of nature (or both): "[b]ut the question is"—she urges us to ask—"[w]hy should we hold fixed the past and the laws of nature?" (p. 190). Well, the argument is a reductio, and it assumes that we have strong intuitive reasons to think that both things in the distant past and the laws of nature are out of the scope of our influence. You can get out of the contradiction that the argument imposes by simply denying that the argument is a reductio, of course. But if this is the solution one is offering, it would be useful indeed if that would be made explicit (by connecting it to Lewis [61] classical account, for example).

In places it also feels that the argumentation comes close to mixing the worries stemming from the apparent incompatibility of free will and determinism with the worries stemming from fatalism:

> "Deliberating is like following your own footsteps through an untouched desert. The footsteps are real, and others who come after you can follow them. But you put them there in the act

> of walking. They aren't there before, and they aren't "there anyway." You don't *follow* your own footsteps. They might lead others, but they don't lead *you*." ([6], p. 168.)
>
> "Fate is what happens. It doesn't determine your actions. Your actions determine fate." ([6], p. 214.)

It is a fallacy—the fatalist fallacy—to think that if determinism holds then your decisions and actions don't matter, and everything is fixed and happens as it happens irrespective of you. What moves the incompatibilist is not the worry that our actions and decisions don't matter, but the worry that they are fully determined by matters outside of our control. Again, a clear separation of the problem of the freedom of the will from the problem of the causal efficacy of the will would be welcome indeed.

However, there does seem to be a few threads of ideas that might be interpreted to have a bite on these fundamental issues and spark some genuine advancement in the debate. For example, Ismael [6] is clear, at least in places, that she thinks that the laws of nature are not indeed constant (or at least not in the sense that the debate typically assumes), and that this is the key to solving the puzzle of free will. "What might it mean to say that the laws relate to the History of the universe in something like the way that the property of being a tragedy relates to the stories to which it applies?," she asks rather metaphorically ([6], p. 176; cf. also [4]). Now, if the laws of nature fail to be constant, and, more specifically, if we and our decisions are part of the "unfolding" of the universe—as we clearly must be, since we are definitely not outside observers either—then it is indeed the case that the consequence argument does not look like a reductio anymore since our actions do in fact change the laws of nature—simply because our actions are not separate consequences of them. However, if this is what the argument really amounts to, it could have, and should have been stated much more clearly. Note also that this move seems to entail a tacit shift from compatibilism to incompatibilism: if there is a sense in which our actions can influence the laws of nature, then surely the laws of nature are not deterministic anymore? It seems that a number of things in this view call for clarification.

AN ARGUMENTATIVE ASCENT TO FREE WILL

It seems that there are concurrently three different, but interconnected arguments at play. First, it is noted that even if the fundamental deterministic laws of physics would be quite simple, any real system will exhibit such a level of complexity that it will be theoretically impossible to predict the behavior of the system (to make an accurate prediction you would need to specify not only the values of all the variables endogenous to the system, but also the values of all the exogenous variables, and hence ultimately the values of all the variables that specify the state of the universe as a whole). So given that we can never perceive more than an approximation of the state of the universe, there will always be a set of alternative courses of events (histories and futures) consistent with the currently perceived state of the universe. A devoted incompatibilist would quickly note, however, that this follows simply from our subjective perspective and limited cognitive capacities, and in no way challenges the fact that the given initial conditions and deterministic laws of nature entail only one possible (actual), real course of events. On the other hand one could still maintain that this gives us a precisely defined set of epistemically possible courses of events, which could ground at least our sense of free will.

Second, there are also appeals to the fact how the fundamental laws of physics are assumed to be symmetric under time reversal, and how in current (relativistic) physics time is not treated as an external parameter. This is a very important point, and although this is no news to philosophy of physics, these issues have not been receiving the attention they deserve in discussions related to free will. A moment's reflection should make it clear that taking these fairly simple facts into account could change the discussion significantly. For one, it would now appear that affecting the future and affecting the past are on a par with each other. That is, if the past-future asymmetry that we perceive is illusory (from the perspective of fundamental physics), then it is no more absurd to think that we can change the initial conditions of the universe than it is to think that we can change the future events. In fact, future events would now seem to affect—or at least correlate with—the past. When we worry about the possibility of free will under determinism, the argument would now go, we tacitly posit an outside observer inspecting the events from a superreal frame of reference. In particular, the consequence argument seems to assume a mistaken view on reality where the initial conditions and laws of nature are fixed somewhere in the past, and where we appear to the scene at some later point in time, after which everything is run like in an ingeniously written film. But where exactly is this past fixing happening, and who has written this film—and who is watching it? It is clearly right to maintain that once we see that we have to let go of this image, we will start to see ourselves as integral parts of the universe—as parts of its initial conditions and laws—and the force of the consequence arguments starts to evaporate. If we, and our decisions and actions, are simply parts of the fabric of the spacetime continuum, it may not be that absurd to claim that our free choices alter either the initial conditions of the universe or the laws of nature (or both): "there is no support in physics for the idea that the past is 'fixed' in some way that the present and future are not, or that it has some ontological power to constrain our actions that the present and future do not have" [4].

Finally, there is a third argument that is based on the results of the previous two. Consider first an incompatibilist response to the second argument: surely, if the past and the future are symmetrically dependent on each other, this just proves that there is no free will—that we cannot affect the future while holding past (and laws of nature) constant. Now one could point out—and arguably this is something that this argumentation is ultimately trying to get at—that such a response simply does not make sense; that one must admit either that our choices can affect both the past and the future (which renders the consequence argument invalid) or that the whole idea of causing and affecting is ill-founded (which renders debate surrounding the notion of free will thoroughly meaningless). If one now maintains that the idea of causing and affecting is non-negotiable, but is simultaneously also reluctant to hold the past and the future symmetrical (and thus accept the first horn of the dilemma), then one has to admit that in this context we are forced to consider only subsystems of the actual universe. That is, only in such (open) subsystems does the past-future and cause-effect asymmetries hold (and the causation defined in interventionist terms make sense): "if you wish to include the whole universe in the model", as Pearl [47] notes, "causality disappears because interventions disappear—the manipulated and the manipulator lose their distinction" (p. 349–350; cf. also [62–64]). But now the first argument kicks back in: each subsystem under inspection is always a member of a set of equally possible subsystems consistent with the given initial conditions and laws of nature. In other words, dividing the universe into subsystems is necessary for the idea of causing and affecting to make sense, which in turn allows the simultaneous existence of alternative possibilities, which in turn can ground the idea of the ability to do otherwise, and, consequently, the idea of free will.

FROM THE BLOCK UNIVERSE TO THE BUBBLED BLOCK UNIVERSE

Here is an attempt to make this argumentation more precise. It seems that the consequence argument is tacitly relying on an ill-licensed asymmetry. The argument assumes that our actions are consequences of the initial conditions and laws of nature [premise (4) above]. However, this is only half of the truth, for it is equally true, it has now become apparent, to say that the initial conditions and laws of nature are consequences of our actions. The reason why this sounds absurd is that we are tacitly imposing an imbalanced interpretation on the relevant variables. All parties in the debate agree that in order for the consequence argument to have any chances of getting off the ground both S and L must describe the *complete* state of the universe and the *complete* set of fundamental laws of nature, respectively, down to the smallest detail and perfect accuracy. What seems to have gone unnoticed, however, is that what follows from such a complete description is not the coarse-grained events that we typically identify as our actions, but a complete state of the universe, down to the smallest detail. Once this discursive asymmetry has been exposed, and the content of A is set on a par with S and L, the air of absurdity between the symmetrical dependency between these variables starts to evaporate.

If A would now refer to a complete chunk of the universe at a time t_1, rather than anything we would normally call an action, then it would indeed be intuitively absurd to say that an agent could cause anything like A to occur. That is, now the premise (2), that is typically not being questioned, would become undermined. In other words, *if* we think that our actions are logical consequences of the initial conditions and the laws of nature, *then* our actions are not caused by us (simply because they are not caused by anything). If, on the other hand, we want to hold onto the causal discourse—as both compatibilists and incompatibilists typically do—and maintain that our actions are caused by us, then either we must admit that the initial conditions and the laws of nature are causally dependent on our actions, or we must break the symmetry between the two. Since the former sounds absurd, we opt for the latter. But in that case there is no reason to hold onto the premise (4) anymore. The dependency stated by (4) was supposed to be holding between symmetrically related elements of reality. Now that this symmetry is gone, there is no reason to believe in the asymmetric dependency either.

To put this more formally, instead of (4) we should have the following:

(4*)[1] $(S \wedge L) \Leftrightarrow (A \wedge L)$

This in turn entails not only (4), but also:

(9) $(A \wedge L) \Rightarrow (S \wedge L)$

There is no less reason to interpret (9) in causal terms as there is to do so with respect to (4); since (4*) is not expressing a causal dependence, neither should (4) nor (9). In other words, A can be a consequence of $(S \wedge L)$ only if $(S \wedge L)$ is also a consequence of A (and the laws of nature), and if you want to deny the latter [i.e., (9)], you must also deny the former [i.e., (4)]. If you want to hold onto the idea that A is causally—asymmetrically, that is—dependent on an agent [i.e., (2)], then "A" cannot refer to the same thing in (2) as it does in (4*), (4), and (9). So if you move into the realm of causation, and detach yourself from (4*), you *ipso facto* detach yourself from (4) and (9). There can still be determinism of a sort; (4*) and its consequences can still be true. There is no *causal* determinism, however, and what (4*), (4), and (9) express are purely logical or mathematical dependencies. What S and A are, in effect, are simply arbitrary chunks of the entire spacetime block that constitutes the universe. Once you realize that, there is no less reason to accept (9) as there is to accept (4): A could simply stand for the final conditions of the universe, and since the laws that govern the evolution of the universe are symmetric, the totality of the final conditions of the universe entail the totality of the initial conditions as much as the other way around. However, neither of these are a cause of the

[1]To be precise, this assumes that the initial conditions and the laws of nature can be separated from each other, and that the latter are constant [which grounds the symmetry between the "past" (S) and the "future" (A)]. However, neither of these assumptions can be taken for granted. First, are we empirically justified in thinking that the laws of nature are constant? Einstein [65, 66] and Dirac [67] already suggested the possibility that some of the fundamental physical constants might not be time-invariant, and a number of recent studies have indicated that that might actually be the case (cf. e.g., [68–73]). The ramifications of such a conclusion would of course be radical, and it would entail (*prima facie*) that determinism does not hold at any scale (or that the whole notion would collapse). Second, if we are relying on a Humean (as opposed to non-Humean or Aristotelian) account of the laws of nature (as those proposing this line of argumentation typically are), then we cannot, in principle, separate L from S—what we actually have is only S encompassing everything. That is because L (and all other "necessary connections") supervene on the entire "Humean mosaic" of spacetime points, and they are not something that are over and above the spacetime, separately guiding or governing the evolution of the universe."All there is to the world is a vast mosaic of local matters of particular fact, just one little thing and then another", as Lewis ([74], p. ix) puts the idea (cf. also [75]). The laws of nature are simply theorems of the best systematization of this Humean mosaic. What this would then seem to entail is that any changes anywhere in the spacetime would result in changes in the laws (since chains of events are not driven by the laws of nature but exactly the opposite). This would again radically transform the basic elements of the debate, and maybe puts Ismael's ([6], p. 176) metaphor about the laws of nature relating to the evolution of the universe in the same way as the property of being a tragedy relates to the stories to which it applies in the right metaphysical context (cf. also [4]).

other, or, if you would insist on using causal terminology, both are equal causes of each other.

The key issue here is the scope of A. If A has a wide scope—if it refers to some complete chunk of the spacetime—then it can be argued that (4) should be replaced by (4*), and that (4) holds only if (9) also holds. However, one could now insist that if A has a narrow scope—if it refers only to a tiny piece of the complete chunk of the universe—then (4*) and (9) fail to hold, but (4) still holds. In other words, the right sort of asymmetry could be reintroduced back into the argument if one just gives A a narrow interpretation. Moreover, one is fully justified in giving such a narrow interpretation because A was supposed to be referring to a specific action performed by the agent all along, not to an enormous non-localized chunk of the entire universe. More precisely, let A now refer to a finite conjunction of events $A = (a_1 \wedge \ldots \wedge a_n)$, and let a_1 be the action performed (i.e., an event produced) by an agent at t_1. Now instead of (4) and (4*) we have:

(10) $(S \wedge L) \Rightarrow a_1$

While (10) is consistent with (4*), and therefore also with (4) and (9), the following does not hold:

(11) $a_1 \Rightarrow (S \wedge L)$

This seems like something that the incompatibilist has been insisting all along. Your actions are consequences of the initial conditions and laws of nature, but not the other way around. Even if there would be symmetrical dependencies at the level of the entire universe, there are no such symmetries at the level of single localized events that constitute your actions.

However—and this is the final twist in the story—this point exactly exposes the ill-licensed asymmetrical reasoning the incompatibilist is relying on. In (10) the antecedent contains quite literally everything, whereas the consequent contains an arbitrary chosen event from a myriad of events that are taking place at the given time, each of which could equally well play the role of a_1 in (10). That is, the antecedent in (10) gives simply a sufficient condition for a_1 to occur. But the focus of the debate has been on the causes of our actions, on the issue of whether performing a particular action is "up to us," and causes in turn, we have learnt, are something else than mere sufficient conditions: they are not only sufficient, but also—and first and foremost—necessary for their effects; causes are the pivotal difference-makers of their effects, something without which the effects would not have occurred. But (10) is silent about such a pivotal role; nothing in it states what is necessary in the antecedent for a_1 to occur. Maybe some proper part of it is, but definitely not all of it. To get at the genuine causes of events we need to inspect the effects and see how they respond to changes elsewhere in the given system. In particular, we need find out what we need to remove from the system to remove the effect—we need to find out what was necessary, in the given circumstances, for the effect to occur. Once we hit that, we are on a track to nailing the cause of the effect. For any of this to make sense we need to be able to compare different systems to each other—and we need to do this from within another, more encompassing system; we need to assume a position from which we can hypothesize that when some things are held constant—when most of the universe is held constant—it is true to say of some subsystem of the universe that *if* the cause-event would not have occurred in that system, *then* the effect-event would also have been absent. The initial conditions and laws of nature are not the causes of effects, but rather the background conditions in the light of which particular causal hypotheses can be formulated and tested; they simply form the basis of the general framework that enables us to engage in a causal discourse.

Now, what we typically call actions are rather coarse-grained events. Greeting somebody, let's say, can materialize in various different ways—as various different locutions and as various different bodily movements. Fine physical details are thus irrelevant to any given greeting-event. Therefore, to get to the causes of such greeting-events, we need to find out, not what the entire detailed history of the universe was prior to the occurrence of the given event, but what a very minor part of that history was necessary for the occurrence of that event; all the rest is simply irrelevant noise. And the things that play the right sort of pivotal role to get to be pinpointed as causes of greeting-events are arguably something equally coarse-grained: they are mental states—perceptions, desires, and beliefs—coupled with cultural and social constraints. Such things too, like greetings, can materialize in various different ways, and neither are their occurrences dependent on fine physical details (cf. [76–79]). In other words, the causal system under scrutiny is such that it allows for a myriad of fine-grained physical possibilities and a wide range of variation in the initial conditions; many different physical systems are consistent with our actions and the mental states that cause them. And this physical leeway, the argument would now go, is what ultimately grounds our free will: our actions are not dependent on the initial conditions of the universe—even if they (together with the deterministic laws of nature) would be sufficient for everything that happens—but only on a relevant subset of the events antecedent to the actions, leaving it possible for alternative courses of events to take place.

What the consequence argument seems to presuppose, then, is both that the initial conditions of the universe together with the deterministic laws of nature entail our actions and that our actions are caused by us. On a closer look, however, this turns out to be an incoherent presupposition. What the initial conditions and the laws of nature entail is quite literally everything. Therefore, they entail, rather trivially, also our actions. What the initial conditions and the laws of nature do not do, however, is to cause our actions—simply because they do not cause anything. They are grounds for everything, but causes for nothing. If, then, we want to hold onto the idea that there are causal relations in the world, relations between us and our actions among them, then we are committed to limit our inspection on proper subsystems of the universe. What we need to do, we could say, is to move from the Block Universe into a bubble within: from the idea of an unchanging four dimensional, all encompassing block of spacetime to a view from within such a block, a bubble where causes and effects can take place, surrounded by a plethora of equally possible, alternative bubbles. To put this more formally, we can define "small world" as an assignment of values to all the endogenous variables in a causal model, and "large world" as an assignment of values to both all the endogenous and all the exogenous variables in such a model (cf. [47, 80]. From the perspective of the "large world," any "small world" bubble would then be an equiprobable alternative to all the other "small world" bubbles; from within our bubble, where causes and effects can take place, we are always presented with alternatives, other bubbles where things turn out differently (**Figure 1**). From within our bubble, therefore, we are presented with various options, and when we act on our free choices, we are always able to behave in more than one way. What we get, then, is, as Hoefer [3] puts it, "freedom from the inside out".

THE INCOMPATIBILIST STRIKES BACK: LESSONS FROM THE SECOND LAW

Disarming the consequence argument comes therefore in two stages. First, the asymmetry of determinism that the argument rests on is rejected: from the perspective of the fundamental laws of physics the present, or any future state of the universe entails any past state of it as much as vice versa. Second, once we turn our focus on the asymmetries we perceive, most notably to the temporal and causal asymmetries that constantly surround us, the symmetric determinism imposed by the fundamental laws of physics evaporates, and we are actually faced with overwhelming degrees of microphysical freedom. Our decisions and the actions resulting from them are thus macrophysical, emergent phenomena for the simple reason that causation is a macrophysical, emergent phenomenon. There is no causation at the fundamental physical level, but that does not make causation any less real. And similarly free will, the idea now is, "is no more ruled out by the consequence argument than the Second Law of Thermodynamics is ruled out by microscopic reversibility" ([81]; cf. also [1]). Elzein et al. [82] dub this sort of view "Supervenient Libertarianism," for it postulates genuine, libertarian free will—free will that presupposes the availability of alternative possibilities—as a real higher-level phenomenon that supervenes on, or emerges from, lower-level deterministic dynamics (Backmann [83] uses the term "Humean Libertarianism" for a similar view).[2] By acting freely we can therefore bring about—*causally determine*—macrophysical changes in the future, but not in the past, although the future *logically determines* the past (and vice versa).

So what this account does not commit us to is some sort of backward causation, anymore than—and exactly to the same extent as—the Second Law commits us to water spontaneously boiling at a room temperature. That is, the total entropy of the universe practically never decreases (to any significant extent), although there exists a small likelihood that that might actually happen. In fact, as the fluctuation theorem indicates [84–86], and the subsequent experimental evidence shows [87–89], if we are inspecting a small enough system (or in principle any isolated system in a state of maximum entropy), the total entropy

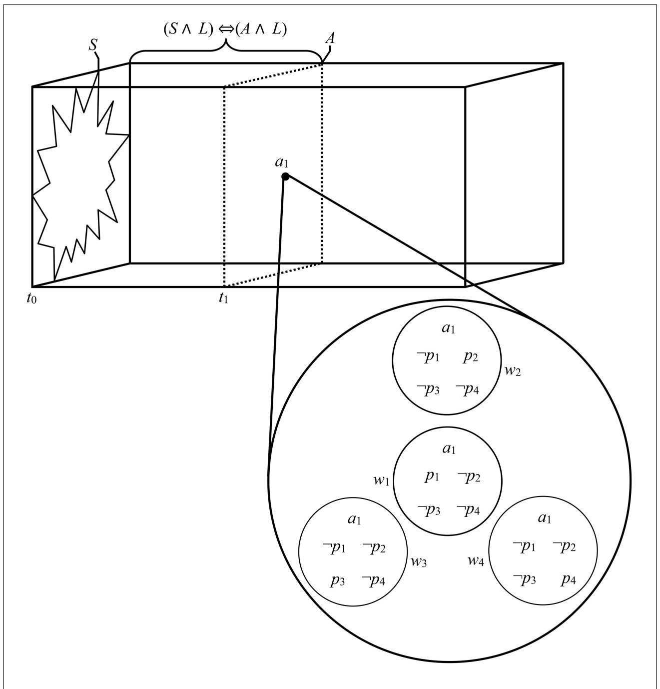

FIGURE 1 | Bubbled Block Universe, where S is the complete set of initial conditions (the low-entropy Big Bang) at the time-coordinate t_0, L is the complete set of all the fundamental laws of physics, A is a complete state of spacetime at the time-coordinate t_1, and a_1 is a specific action performed at t_1. The action a_1 is a macrophysical event, and there is therefore a set of equiprobable (mutually exclusive) microphysical states (p_1–p_4) that are consistent with the occurrence of a_1. Supposing that w_1 represents the actual circumstances (that a_1 occurs and is realized by p_1), the subsystems w_2–w_4 represent the possible, non-actual alternatives to w_1. In other words, from the perspective of a_1 each of the systems w_1–w_4 are equally possible. (Figure not to scale.)

of the system will fluctuate, and in such systems time—and causation—will actually run backwards (or back and forth). Or, to be more precise, from within such a system nothing would seem to run in reverse—exactly because the fundamental laws governing the dynamics of the system are reversible; things would run "backwards" only because we are observing the system from the outside. Local, momentary backward influences are therefore a physical fact, and although our actions will always be accompanied by such minute rippling backward influences, they will evaporate before we are able to perceive them, and they are of no practical use to us. So although we can—and do—exert backward causal influence, we posses only forward causal, pivotal control: we can only bring about changes in the future, although our actions influence also the past. If this sounds absurd, it is only because our intuition is holding onto an absolute and objective distinction between the past and the future, which is something that has no ground in current physics.

Can we now rest assured that the correct understanding of physics will deliver us from all our worries concerning free will and agency? Unfortunately not. The most burning question is this: can we really defeat the consequence argument by relying on symmetric determinism at the microphysical level while at the same time holding onto temporal and causal asymmetries

[2] It is worth noting, however, that this view is in *prima facie* conflict with the formal results of Conway and Kochen [36, 37] which suggest macroscopic freedom (the freedom of the experimenter to take measurements) entails indeterminism at the microphysical level.

at the macrophysical level? Macrophysical asymmetries are of course consistent with microphysical symmetries, and this relationship can be seen to ground the temporal and causal asymmetries that surround us, as Carroll [81] notes. However, is noting this connection really enough to undermine the consequence argument? More specifically: why wouldn't the incompatibilist be now entitled to insist that it was determination in the macrophysical, asymmetric sense all along that has been keeping her awake at nights, and that nothing that has been said above has challenged that sort of determinism—on the contrary, hasn't it now simply been given a solid footing in current physics? The initial charge was that the incompatibilist has been relying on causal determinism when in fact there is no such thing in fundamental symmetric physics. However, once an explanation is given to the perceived macrophysical asymmetries, and causal determinism is thus brought back in—albeit in terms of probability distributions—the incompatibilist could simply now restate her case in macrophysical terms. If any backward influences that might accompany our actions are negligible compared to their forward influences, as the argument went, then surely the incompatibilist could simply point out that now every decision and every action is preceded by events that exert stronger influence on those decisions and actions than vice versa? We cannot reverse the growth of entropy, and we cannot reverse the direction of causation, and we cannot change the events in the past that determine our behavior—and hence we are not free to choose the courses of our actions, the incompatibilist could now insist.

It is important to understand, however, that we are not back to square one. It is true that the consequence argument is typically phrased, explicitly or implicitly, in terms of microphysical determinism. The traditional, Laplacean, worry is that if we only knew the exact, detailed state of the universe at some moment in time, we could, if the fundamental laws of nature are deterministic, and we would posses complete knowledge of them too, calculate all the future events in perfect precision and absolute certainty. The incompatibilist argumentation based on this traditional worry has now been shown to be misguided: it is not absurd to think that you acting differently than you actually did would change the past simply because at the scale of the entire universe the past and the future are symmetrically related, and hence all future changes would, as a matter of physical fact, result in changes in the past. This is an important observation, and a real advancement in the debate. However, the compatibilists have been too quick, it would now seem, to conclude that the consequence argument is invalid and that we can therefore have such a thing as free will in a deterministic universe. The initial conditions (or the state of the universe at any arbitrary point in time in the past), S, could now be interpreted in macrophysical terms, encompassing all the microphysical variation that is compatible with the macrophysical evolution of the universe between t_0 and t_1. In other words, all the microphysical leeway that there might be is simply irrelevant, since all these microphysical states will converge into the same macrophysical states that constitute our particular decisions and actions, and you would have to be able to change the macrophysical, low-entropy state of the universe in order to have free will and be able to act otherwise. But such an entropic reversal is practically impossible at the scale of our interest, as those exploiting the microphysical symmetries to defeat the consequence argument are eager to stress.

There is another, perhaps a more tangible way of phrasing the problem we are now faced with. What the argumentation reviewed in here may have been able to achieve, is to show how we can be genuine causal agents in a thoroughly physical world—how we, by making conscious decisions, can initiate actions and exert control on future courses of events. This would be a major achievement, of course. It would show that the widespread worries about our mental states being causally inert are misguided. But that, however, as it has already been stressed, would only address a part of the issue we are facing with the problem of free will. Another, and arguably more fundamental issue has been left unaddressed: the question of whether we are truly free to make the decisions that give rise to our actions. So the core problem is not really that the rest of the universe determines our actions; the problem is that it determines our decisions. And in fact, as the preceding discussion arguably shows, the first worry can now be dispensed with: we, rather than the rest of the universe, might indeed function as genuine, pivotal difference-makers with regard to our actions—we, and our decisions and actions might actually matter. We might be, as it were, indispensable cogs in the vast clockwork of the universe, the removals of which would make things run differently. There remains the worry, however, that every movement of the cogs would still be determined by the prior movements of some of the other cogs in the clockwork; that our decisions would be fully determined by factors beyond our sphere of influence. To address that worry, and to successfully soothe it, one would need to show that our decisions are underdetermined by prior events, or, to put it in current parlance, that there are no prior events one can point to as pivotal difference-makers with regard to our decisions. But such an argument, it now becomes apparent, would seem to be missing.

There is a gap, therefore, that would need to be closed to complete the argument sketched here. It has been too quick to note that "free will is no more ruled out by the consequence argument than the Second Law of Thermodynamics is ruled out by microscopic reversibility" [81]. We can offer a consistent physical story on how the macroscopic irreversibility emerges from the microscopic reversibility, and maybe we can, in the way outlined here, ground causal relations, genuine mental causation among them, on such macroscopic asymmetries. But that would only explain how our will can make a difference. To fully solve the problem of free will—to tackle the issue of the freedom of the will—one would also need to explain how that will could be free.

CODA

It might be true that the past 100 years or so has not brought much progress in the debate on free will, as some have claimed; all involved in the debate, philosophers, and scientists, may have been talking more past each other than to—or with—each other. At the same time, it is clear that major advancements have been made in closely related issues in philosophy of science and in philosophy of mind, and, most obviously, in physics and other natural sciences. It might thus not be too audacious to claim that we might well soon be witnessing some progress in this issue as well—maybe it is just the question of putting all the pieces together the right way.

Although the recent developments reviewed here leave many questions in the air, and often the discussion could benefit from a more systematic framing of the most fundamental issues in the debate, the general framework of the argumentation is very enticing indeed, and fits quite nicely into the more encompassing story that the naturalistic philosophy and the empirical sciences have been putting together for half a century or so. Even if these developments would not turn out to constitute *the* event that made us turn the leaf of history on this matter, they most certainly will be something that the future will hold as pivotal in the chain of events that resulted in that turn.

AUTHOR CONTRIBUTIONS

The author confirms being the sole contributor of this work and approved it for publication.

ACKNOWLEDGMENTS

I would like to thank Dr. Nadine Elzein, Ms. Sophie Kikkert, Mr. Chris Oldfield, Mr. Otto Overkamp, Prof. David Papineau, Dr. Paavo Pylkkänen, Dr. Michael Ridley, Mr. Athamos Stradis, Mr. Kalle Timperi, Prof. Jürgen Vollmer and the two reviewers of *Frontiers in Physics* for criticism and comments on various earlier versions of this paper. This work has been financially supported by the Kone Foundation and the Finnish Academy of Science and Letters.

REFERENCES

1. Carroll S. *The Big Picture: On the Origins of Life, Meaning and the Universe Itself*. New York, NY: Dutton (2016).
2. Dennett DC. *Freedom Evolves*. New York, NY: Viking (2003).
3. Hoefer C. Freedom from the inside out. In Callender C, editor. *Time, Reality and Experience*. Cambridge: Cambridge University Press (2002). p. 201–22.
4. Hoefer C. Causal determinism. In Zalta EN editor. *The Stanford Encyclopedia of Philosophy*. Spring Edition (2016). Available online at: https://plato.stanford.edu/archives/spr2016/entries/determinism-causal/
5. Ismael JT. Causation, free will, and naturalism. In Ross D, Ladyman J, Kincaid H, editors, *Scientific Metaphysics*. Oxford: Oxford University Press (2013). p. 208–35.
6. Ismael JT. *How Physics Makes Us Free*. New York, NY: Oxford University Press (2016).
7. List C. Free will, determinism and the possibility of doing otherwise. *Noûs* (2014) **48**:156–78. doi: 10.1111/nous.12019
8. List C, Menzies P. My brain made me do it: the exclusion argument against free will, and what's wrong with it. In Beebee H, Hitchcock C, Price H, editors. *Making a Difference: Essays on the Philosophy of Causation*. Oxford: Oxford University Press (2017). p. 269–85.
9. Loewer B. The consequence argument meets the Mentaculus. In *Presented at the CEU-Rutgers MIND Workshop*. Department of Philosophy, Central European University (2017). Available at: https://videosquare.ceu.edu/en/recordings/details/20,Barry_Loewer_The_Consequence_Argument_Meets_the_Mentaculus
10. Menzies P. The consequence argument disarmed: an interventionist perspective. In Beebee H, Hitchcock C, Price H, editors. *Making a Difference: Essays on the Philosophy of Causation*. Oxford: Oxford University Press (2017). p. 307–30.
11. Steward H. Libertanianism as a naturalistic position. In Timpe K, Speak D, editors. *Free Will and Theism: Connections, Contingencies, and Concerns*. Oxford: Oxford University Press (2016). p. 158–71.
12. Taylor C., Dennett DC. Who's afraid of determinism? Rethinking causes and possibilities. In Kane R, editor. *The Oxford Handbook of Free Will*. Oxford: Oxford University Press (2002). p. 257–78.
13. Ismael JT, Pollock JL. Nolipsism: so you think you exist, do you? In Crisp T, Davidson M, Laan DV, editors. *Knowlege and Reality: Essays in Honor of Alvin Plantinga*. Dordrecht: Springer Verlag (2004). p. 35–62.
14. Earman J. *A Primer on Determinism*. Dordrecht: Reidel (1986).
15. Earman J. Aspects of determinism in modern physics. In Earman J, Butterfield J, editors. *Handbook of the Philosophy of Science Vol 2: Philosophy of Physics*. Amsterdam: Elsevier B.V. (2007). p. 1369–434.
16. Earman J, Norton JD. Comments on Laraudogoitia's 'Classical particle dynamics, indeterminism and a supertask'. *Br J Philos. Sci.* (1998) **49**:123–33.
17. Norton JD. Causation as folk science. *Philos. Imprint* (2003) **3**:1–22. Available online at: http://hdl.handle.net/2027/spo.3521354.0003.004
18. Earman J. *Bangs, Crunches, Whimpers, and Shrieks: Singularities and Acausalities in Relativistic Spacetimes*. New York, NY: Oxford University Press (1995).
19. Bohr NHD. The quantum postulate and the recent development of atomic theory. *Nature* (1928) **121**:580–90.
20. Heisenberg WK. Über den anschaulichen Inhalt der quantentheoretischen Kinematik und Mechanik. *Z Physik* (1927) **43**:172–98.
21. Heisenberg WK. *Physikalische Prinzipien der Quantentheorie*. Leipzig: Hirzel (1930).
22. Heisenberg WK. *Physics and Philosophy: The Revolution in Modern Science*. London: George Allen & Unwin (1958).
23. Howard D. Who invented the "Copenhagen interpretation"? A study in mythology. *Philos Sci.* (2004) **71:**669–82. doi: 10.1086/425941
24. Einstein A, Podolsky B, Rosen N. Can quantum-mechanical description of physical peality be considered complete? *Phys Rev.* (1935) **47**: 777–80.
25. Bell JS. On the Einstein-Podolsky-Rosen paradox. *Physics* (1964) **1**:195–200.
26. Bell JS. On the problem of hidden variables in quantum mechanics. *Rev Mod Phys.* (1966) **38**:447–52.
27. Bell JS. Locality in quantum mechanics: reply to critics. *Lett Épistémol.* (1975) **13**:2–6.
28. Aspect A, Dalibard J, Roger G. Experimental test of Bell's inequalities using time-varying analyzers. *Phys Rev Lett.* (1982) **49**:1804–7.
29. Aspect A, Grangier P, Roger G. Experimental realization of Einstein-Podolsky-Rosen-Bohm Gedankenexperiment: a new violation of Bell's inequalities. *Phys Rev Lett.* (1982) **49**:91–4.
30. Bell JS. *Speakable and Unspeakable in Quantum Mechanics. 2nd ed*. Cambridge: Cambridge University Press (2004).
31. Bohm D. A suggested interpretation of the quantum theory in terms of 'hidden' variables I. *Phys Rev.* (1952) **85**:166–79.
32. Bohm D. A suggested interpretation of the quantum theory in terms of 'hidden' variables II. *Phys. Rev.* (1952) **85**:180–93.
33. Bell JS, editor. Free variables and local causality. *Speakable and Unspeakable in Quantum Mechanics, 2nd Edn*. Cambridge: Cambridge University Press (2004). p. 100–104.
34. Bell JS. Atomic-cascade photons and quantum-mechanical nonlocality. *Comm. Atomic Mol. Phys.* (1980) **9**:121–6.
35. Bell JS. Interview with John Bell. In Davies PCW, Brown JR, editors. *The Ghost in the Atom: A Discussion of the Mysteries of Quantum Physics*. Cambridge: Cambridge Unniversity Press (1986). p. 45–57.
36. Conway JH, Kochen S. The free will theorem. *Found Phys.* (2006) **36**:1441–73. doi: 10.1007/s10701-006-9068-6
37. Conway JH, Kochen S. The strong free will theorem. *Notices AMS* (2009) **56**:226–32. Available online at: http://www.ams.org/notices/200902/rtx090200226p.pdf
38. Shimony A, Horne MA, Clauser JF. Comment on the theory of local beables. *Epistemol Lett.* (1976) **13**:1–8.

39. Zeilinger A. *Dance of the Photons: From Einstein to Quantum Teleportation*. New York, NY: Farrar, Straus & Giroux (2010).
40. Pernu TK. Minding matter: how not to argue for the causal efficacy of the mental. *Rev Neurosci.* (2011) **22**:483–507. doi: 10.1515/RNS.2011.043
41. Tegmark M. The importance of quantum decoherence in brain processes. *Phys. Rev. E* (2000) **61**:4194–206. doi: 10.1103/PhysRevE.61.4194
42. Schrödinger E. Die gegenwärtige Situation in der Quantenmechanik. *Naturwissenschaften* (1935) **23**:807–12.
43. Pernu TK. The five marks of the mental. *Front Psychol.* (2017) **8**:1084. doi: 10.3389/fpsyg.2017.01084
44. van Inwagen P. The incompatibility of free will and determinism. *Philos. Stud.* (1975) **27**:185–99.
45. van Inwagen P. *An Essay on Free Will*. Oxford: Clarendon Press (1983).
46. Halpern JY. *Actual Causality*. Cambridge, MA: MIT Press (2016).
47. Pearl J. *Causality: Models, Reasoning, and Inference*. Cambridge: Cambridge University Press (2000).
48. Woodward JF. *Making Things Happen: A Theory of Causal Explanation*. Oxford: Oxford University Press (2003).
49. Lagnado DA, Gerstenberg T, Zultan R. Causal responsibility and counterfactuals. *Cogn Sci.* (2013) **47**:1036–73. doi: 10.1111/cogs.12054
50. Campbell J. Control variables and mental causation. *Proc Aristot Soc* (2010) **110**:15–30. doi: 10.1111/j.1467-9264.2010.00277.x
51. Beebee H, Hitchcock C, Price H. *Making a Difference: Essays on the Philosophy of Causation*. Oxford: Oxford University Press (2017).
52. List C, Menzies P. Nonreductive physicalism and the limits of the exclusion principle. *J Philos.* (2009) **106**:475–502. doi: 10.5840/jphil2009106936
53. Sartorio C. Causes as difference-makers. *Philos Stud.* (2005) **123**:71–96. doi: 10.1007/s11098-004-5217-y
54. Sartorio C. Making a difference in a deterministic world. *Philos Rev.* (2013) **122**:189–214. doi: 10.1215/00318108-1963707
55. Waters KC. Causes that make a difference. *J Philos.* (2007) **104**:551–79. doi: 10.5840/jphil2007104111
56. Albert DZ. *Time and Chance*. Cambridge, MA: Harvard University Press (2000).
57. Albert DZ. The sharpness of the distinction between the past and the future. In Wilson A, editor. *Chance and Temporal Asymmetry*. Oxford: Oxford University Press (2014). p. 159–74.
58. Hitchcock CR. Actual causation: what's the use? In Beebee H, Hitchcock C, Price H, editors. *Making a Difference: Essays on the Philosophy of Causation*. Oxford: Oxford University Press (2017). p. 116–31.
59. Mackie JL. Causes and conditions. *Am Philos Q.* (1965) **12**:245–65.
60. Mackie JL. *The Cement of the Universe: A Study of Causation*. Oxford: Oxford University Press (1974).
61. Lewis DK. Are we free to break the laws? *Theoria* (1981) **47**:113–21.
62. Hausman DM. *Causal Asymmetries*. Cambridge: Cambridge University Press (1998).
63. Hitchcock CR. What Russell got right. In Price H, Corry R, editors. *Causation, Physics, and the Constitution of Reality: Russell's Republic Revisited*. Oxford: Oxford University Press (2007). p. 45–65.
64. Woodward JF. Causation with a human face. In Price H, Corry R, editors. *Causation, Physics, and the Constitution of Reality: Russell's Republic Revisited*. Oxford: Oxford University Press (2007). p. 66–105.
65. Einstein A. Über das Relativitätsprinzip und die aus demselben gezogenen Folgerungen. *Jahrbuch Radioaktivität Elektronik* (1907) **4**:411–62.
66. Einstein A. Über den Einfluß der Schwerkraft auf die Ausbreitung des Lichtes. *Ann Physik* (1911) **35**:898–906.
67. Dirac PAM. A new basis for cosmology. *Proc R Soc. A* (1938) **165**:199–208.
68. Barrow JD. *The Constants of Nature: From Alpha to Omega*. London: Jonathan Cape (2002).
69. Barrow JD, Lip SZW. Generalized theory of varying alpha. *Phys Rev D* (2012) **85**:023514. doi: 10.1103/PhysRevD.85.023514
70. Calmet X, Keller M. Cosmological evolution of fundamental constants: from theory to experiment. *Mod Phys Lett A* (2015) **30**:1540028. doi: 10.1142/S0217732315400283
71. Chiba T. The constancy of the constants of nature: updates. *Prog Theor Phys.* **126**:993–1019. doi: 10.1143/PTP.126.993
72. Uzan J-P. Varying constants, gravitation and cosmology. *Living Rev Relat.* (2011) **14**:2. doi: 10.12942/lrr-2011-2
73. Webb JK, Murphy MT, Flambaum VV, Dzuba VA, Barrow JD, Churchill CW, et al. Further evidence for cosmological evolution of the fine structure constant. *Phys Rev Lett.* (2001) **87**:091301. doi: 10.1103/PhysRevLett.87.091301
74. Lewis DK. *Philosophical Papers: Volume II*. Oxford: Oxford University Press (1986).
75. Loewer B. Two accounts of laws and time. *Philos Stud.* (2012) **160**:115–37. doi: 10.1007/s11098-012-9911-x
76. Pernu TK. Does the interventionist notion of causation deliver us from the fear of epiphenomenalism? *Int. Stud. Philos. Sci.* (2013) **27**:157–72. doi: 10.1080/02698595.2013.813254
77. Pernu TK. Interventions on causal exclusion. *Philos Explor.* (2014) **17**:255–63. doi: 10.1080/13869795.2013.805800
78. Pernu TK. Causal exclusion and multiple realizations. *Topoi* (2014) **33**:525–30. doi: 10.1007/s11245-013-9159-x
79. Pernu TK. Causal exclusion and downward counterfactuals. *Erkenntnis* (2016) 81:1031–49. doi: 10.1007/s10670-015-9781-7
80. Halpern, JY, Hitchcock, CR. Compact representations of extended causal models. *Cogn Sci.* (2013) **37**:986–1010. doi: 10.1111/cogs.12059
81. Carroll S. *Free Will Is as Real as Baseball* (2011). Available online at: http://www.preposterousuniverse.com/blog/2011/07/13/free-will-is-as-real-as-baseball
82. Elzein N, Pernu TK. Supervenient freedom and the free will deadlock. *Disputatio* (2017) forthcoming.
83. Backmann M. *Humean Libertarianism: Outline of a Revisionist Account of the Joint Problem of Free Will, Determinism and Laws of Nature*. Heusenstamm: Ontos Verlag (2013).
84. Evans DJ, Searles DJ, Mittag E. Fluctuation theorem for Hamiltonian systems: Le Chatelier's principle. *Phys Rev E* (2001) **63:**051105. doi: 10.1103/PhysRevE.63.051105
85. Evans DJ, Searles DJ. Equilibrium microstates which generate second law violating steady states. *Phys Rev E* (1994) **50**:1645–8.
86. Searles DJ, Evans DJ. Fluctuation theorem for stochastic systems. *Phys Rev E* (1999) **60**:159–14.
87. Carberry DM, Reid JC, Wang GM, Sevick EM, Searles DJ, Evans DJ. Fluctuations and irreversibility: an experimental demonstration of a second-law-like theorem using a colloidal particle held in an optical trap. *Phys Rev Lett.* (2004) **92**:140601. doi: 10.1103/PhysRevLett.92.140601
88. Collin D, Ritort F, Jarzynski C, Smith SB, Tinoco I, Bustamante C. Verification of the Crooks fluctuation theorem and recovery of RNA folding free energies. *Nature* (2005) **437**:231–34. doi: 10.1038/nature04061
89. Wang GM, Sevick EM, Mittag E, Searles DJ, Evans DJ. Experimental demonstration of violations of the second law of thermodynamics for small systems and short time scales. *Phys Rev Lett.* (2002) **89**:050601. doi: 10.1103/PhysRevLett.89.050601

Conflict of Interest Statement: The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.