The Monist, 2019, 102, 478–498
doi: 10.1093/monist/onz021
Article

OXFORD

Determinism, Counterpredictive Devices, and the Impossibility of Laplacean Intelligences

Jenann Ismael*

ABSTRACT

In a famous passage drawing implications from determinism, Laplace introduced the image an intelligence who knew the positions and momenta of all of the particles of which the universe is composed, and asserted that in a deterministic universe such an intelligence would be able to predict everything that happens over its entire history. It is not, however, difficult to establish the physical possibility of a counterpredictive device, i.e., a device designed to act counter to any revealed prediction of its behavior. What would happen if a Laplacean intelligence were put into communication with such a device and forced to reveal its prediction of what the device would do on some occasion? On the one hand, it seems that the Laplacean Intelligence should be able to predict the device's behavior. On the other hand, it seems like that device should be able to act counter to the prediction. An examination of the puzzle leads to clarification of what determinism does (and does not) entail, with some insights about various other things along the way.

INTRODUCTION

In an 1814 passage from his *Essay on Probabilities*, Laplace famously asserted that an intelligence that possessed full information about the laws and the positions and momenta of each particle in a deterministic universe would be able to predict everything that was ever to happen in the universe. A puzzle arises if we ask what would happen if we asked the intelligence to reveal its prediction to a device designed to do the opposite of any revealed prediction of its behavior? It is not difficult to show that such a device can be constructed and, on the one hand, it seems that if the setting is deterministic, the Intelligence ought to be able to make a correct prediction. On the other hand, it seems that the device should be able to override any revealed prediction of what it will do. Would the Laplacean intelligence succeed in predicting the device's behavior? Or will the device undermine the intelligence's prediction? What happens here?

In what follows, I use the puzzle to explore determinism, the connection between spatiotemporal and causal structure, self-prediction, and various other things. The

*Columbia University

paper has three parts. In part I, I introduce the puzzle. It is argued that determinism seems to entail the possibility of both natural oracles and counterpredictive devices, but pairing a natural oracle with a counterpredictive device gives rise to a self-defeating causal chain. Part II shows why determinism does *not* entail the possibility of natural oracles and draws some lessons for the right way to understand determinism. Part III looks at the logic of the problem that a would-be oracle faces when confronted by a counterpredictive device. The task of making predictions in a domain in which one's predictions are connected has logical peculiarities that turn out to be revelatory for understanding our own situation in the world.

PART I: THE PUZZLE

We say a universe U is deterministic when, for any chosen time t_0, the physical laws map the state of the universe at t_0 onto the state of the universe at any arbitrarily chosen later time. We say that a universe U contains a natural oracle if a subsystem, S, embedded in U— a human, a computer, a Laplacean Intelligence—is able to predict all the future events in U, from knowledge of its past.[1]

We define a counterpredictive machine as a device that is designed to take in predictions of what it will do on any given occasion and do the opposite. Machine actions can be any kind of output, e.g., printing a symbol on a string, moving a pointer to a position on a dial, displaying an icon on a screen. Imagine a box with green and red lights, for example. When fed the prediction that it will flash the green light, it flashes red and vice versa. A machine like this is a perfectly mundane type of system that takes one of two inputs and produces one of two outputs. It just happens to be designed to act opposite to any prediction of its own behavior when it is fed that prediction in a particular form. This is a very simple device, but we could imagine much more complex devices that can process information in any of a variety of forms, and a creature that could be designed or trained (or could decide on its own) to falsify any revealed prediction of what it will do, whatever form the prediction takes.

A first thought might be that there couldn't be such a device in a deterministic universe. But any embedded system that computes a two-valued function f(x)=y can be programmed as a counterpredictive device. If you have any doubts about the consistency of determinism and the existence of counterpredictive machines, you can translate the question into that of whether the Newtonian equations of motion, for example, allow for the construction of a tabletop counterpredictive red/green light flasher and the answer should seem to be pretty obviously "yes." Or you might just ask yourself whether you could write a deterministic computer program that would take a prediction of whether it will output 0 or 1 and output 0 if predicted it will output 1 and vice versa. And the answer again should be obviously "yes." Either of these machines would have to get rather more sophisticated before it counted as 'forming counterpredictive intentions' and could recognize a wider array of inputs as 'predictions', but whatever it takes to be a system that satisfies those rich intentional descriptions, determinism by itself is no barrier to counterpredictive ability.[2]

The most sophisticated actual predictors—i.e., naturally arising information-gathering and utilizing systems like you and I—are limited in all kinds of ways; we don't know enough about the past and we make mistakes in calculation. Those limitations, however, are contingent. When all contingent limitations on knowledge of the past are removed, a simple computational device that uses information about the past to calculate predictions ought to function as a natural oracle. If natural oracles are physically possible, there will be models of the physical laws that pair a natural oracle with a counterpredictive device, and we can ask a perfectly straightforward physical question: What should we expect to happen if we set up a counterpredictive experiment? We ask our oracle to use its knowledge of the past to calculate what our counterpredictive device will do and we feed its prediction to the device.

This is not merely a philosophical thought experiment. It is a physical situation that should be in principle implementable in any deterministic setting that allows the local construction of counterpredictive devices and that can be analyzed. The reason that this presents a puzzle is that determinism entails that the initial conditions of the universe are, in conjunction with the laws, logically sufficient for determining its full temporal development, so any system that knew the initial conditions and laws ought to function as a natural oracle and ought in principle to be able to answer any question about what happens in its future, and in particular, it ought to be able to answer any question about the output of a counterpredictive device. But any prediction made by the predictor is causally linked to its own defeat. So it seems that we have an argument from determinism to a correct prediction by the oracle and a causal chain that leads through the prediction to its defeat, and something has to give.

THE HISTORY OF THE PUZZLE

The puzzle has a history. A 1965 paper by Michael Scriven called "An essential Unpredictability in Human Behavior" introduced counterpredictive behavior as a limit to predictability even in a deterministic context. An earlier paper by Karl Popper made a similar point about the limits of predictability in classical physics. A number of influential philosophers wrote papers in response to Scriven in the late 1960s and early 70s (David Lewis, I.J. Good, Alvin Goldman, Patrick Suppes, Donald Mackay, Richard Boyd, among others).[3] Daniel Dennett, who learned of the puzzle from Mackay, discussed it in *Elbow Room*.[4] I learned of it indirectly from a paper of Richard Holton's, "From Determinism to Resignation: and How to Stop it," and wrote about it in *How Physics Makes Us Free*.[5]

Connection to closed causal loops

The puzzle has been used to draw lessons, respectively, for quantum mechanics, human agency, the logic of self-prediction, and the nature of choice.[6] There are various versions of it. Scriven considers one that doesn't have the self-defeating structure in which the predictor is forced to reveal its prediction to the device. In this version, the device has access to the same information and simulates the calculation of the predictor, acting counter to the result of its simulation. This leads to discussion of calculation speed and questions about whether the predictor can get ahead of the counterpredictive mechanism. The self-defeating version of the puzzle sidelines all of that. What matters here is the causal order. The predictor can have as much information about the past, and as much time and computing power as it wants. So long as it is forced to reveal its prediction to the counterpredictive mechanism, it can't get ahead of the mechanism in the relevant sense.

There is a close connection between the self-defeating version of the Paradox and the puzzles that arise in connection with closed causal loops. Since information about the total state of the world at any past time in a deterministic context *is information about the future*, what we are doing when we pair a natural oracle with a counterpredictive device is creating a situation in which we (effectively) feed information about the future into the causal past of a system designed to undermine it. We see the same structure in this particularly clean example of a closed causal loop described by Earman:

> Consider a rocket ship which at some space-time point x can fire a probe which will travel into the past lobe of the null cone at x. Suppose that the rocket is programmed to fire the probe unless a safety switch is on and that the safety switch is turned on if and only if the 'return' of the probe is detected by a sensing device with which the rocket is equipped.[7]

If the rocket fires, then the probe travels into the past, is detected by the sensing device, the safety switch is activated, and it doesn't fire. If it doesn't fire, on the other hand, no probe is detected, the safety switch remains off and the rocket fires as programmed. So, the probe is fired, it appears, if and only if it is *not* fired.

The Grandfather Paradox is the most famous philosophical puzzles that arise in connection with closed causal loops. What happens in that case is that someone travels into the past to attempt to kill his grandfather before the conception of his father. In this case, the time traveller's existence is information that his grandfather lives long enough to conceive his father who (in turn) conceives him. And yet he embarks on a mission to keep that from happening.

Is there a contradiction?

Self-defeating causal chains don't lead immediately to contradiction for the same reason that the puzzles that arise in connection with closed timelike curves aren't genuinely paradoxical. There are consistent solutions. To avoid paradox, however, one usually has to put consistency constraints on initial conditions. We know that a perfect predictor isn't going to be able to predict something that doesn't happen, so either it won't make an accurate prediction or the device will malfunction. In a world with aspirant predictors, it seems there must be repeated failures often for motley reasons. Something has to go wrong, and whatever went wrong had to be traceable to the initial conditions, prearranged from the beginning of time.

This sort of consistency constraint is well known in the presence of closed timelike curves. There are consistent solutions to what happens to the Earman rocket, but it is not the case that any locally possible initial condition can be extended to a global solution.[8] To philosophers, this is familiar as the 'trips-on-a-banana-peel' response to the Grandfather Paradox, where some contingent accident keeps the time traveler from succeeding in killing grandfather: e.g., he slips on a banana peel, his gun jams, he gets lost. It is not a surprise that closed timelike curves introduce pathologies that go against our physical expectations. It *would*, however, be surprising to find the need for this kind of restriction in a deterministic setting on a flat space-time. And indeed, the fact that determinism is a formal consequence of local laws that place no such restrictions on initial conditions signals that there was something wrong in the set up that introduced the Laplacean demon + counterpredictive device chain.

Is this really a thing?

You might question the puzzle on the grounds that it relies on semantic notions. After all, if we look at the machine there is nothing paradoxical about the causal chain; it is only when we interpret the input to the machine (the symbols on a string, or the sounds or words coming out of the predictor's mouth) *as a prediction* that we seem to get a puzzle. There are a couple of things to say here. First, it was Laplace who introduced the predictor as a way of dramatizing what he took to be a consequence of determinism: viz. that there are events in the early history of the universe nomologically sufficient to fix its entire future. But the puzzle is really a causal one, and the reference to prediction is eliminable. The predictor is acting as a causal conduit for information from the past that is by hypothesis nomologically sufficient to determine a particular event, into the causal past of a system designed to override it.[9] Pairing natural oracles with counterpredictive devices was a simple way of constructing a self-defeating causal chain. Second, information processing and utilization is part of the natural world, and one way or another these processes have to be incorporated into it. If there are information-processing and utilizing systems in the natural world and there's no in-principle barrier to their knowledge of the past, in a deterministic setting there's no reason in principle that they shouldn't be able to calculate everything that follows from the information available to them with logical precision. So, even if the reference to prediction were ineliminable, I see no reason to think it introduces something physically illegitimate.

One might also object to the puzzle by suggesting that if we think concretely in terms of what it would take to gather the information needed to generate a prediction, we'd find something that prohibited the construction of the self-defeating causal chain: e.g., some thermodynamic subtlety about energy requirements, heat dissipation, or information loss.[10] If this sort of generic worry can be made precise and specific, it will have to be addressed. Even in generic form, however, it seems too shallow a prohibition to provide a full resolution. If the causal structure of a deterministic theory on flat space-time allowed us to feed information nomologically sufficient to determine the future into the causal past of a system designed to undermine it, higher-level constraints that kept us from implementing the process seem too weak to really address the puzzle.

PART II. WHY THERE ARE NO NATURAL ORACLES

We know that we can create a simple system in a deterministic context that can undermine any revealed prediction of its behavior. We know, that is to say, that the laws allow us to lay down a local causal chain in a globally deterministic context that can undermine any revealed prediction of what its output will be. And we also know that in a deterministic context, information about the global state of the world at any early time is nomologically sufficient to fix the future. That is why the analogy with closed causal loops is apt. A Laplacean Intelligence that knew the global state of the world would (effectively) be feeding information about the future into the causal past of a system designed to undermine it causally. And we saw that that is also the reason that additions to computational power and speed on the part of the predictor don't help. Again, so long as the prediction is fed into the counterpredictive device, it cannot "get ahead" of the device in the relevant sense. The crucial ordering here is the causal one.

This means that if information about the total state (or any other information that is nomologically sufficient to fix the total state) were available to causally regulate systems in space and time in a setting that permitted the construction of counterpredictive devices, we would have an analogue of closed causal loops. So we should ask whether there is any reason to think that information about the total state of the world at any time can't be available to an embedded subsystem, and also can't be available to causally regulate processes in the world. It turns out the answer is yes. And that is really the key here. It is because of what 'total' has to mean in order for it to entail a determinate prediction about the future.

Consider a deterministic, Newtonian universe. Consider the collection of events that happen at some early time t, and an event that occurs some time later. Indeed, consider the collection of events C that constitute the whole history of the world up until t, and an event e that occurs just a fraction of a second after t. Do the events in C nomologically determine e? That is the question of whether there are solutions to the Newtonian equations of motion that include C and not e. The answer, as a physical matter, is straightforwardly 'yes'. To obtain such a solution, we just have to add things to C that change the forces impinging on e. It is not the specific collection of events in C, but the totality that determines e. And that is quite general: for any collection of pre-t events, there are models of the Newtonian equations of motion that include those events, in which e, and also models of the laws that include those events, in which ∼e. In order for a set of past events to nomologically determine a future event, we have to specify that those events constitute the totality of what there is. That is not information that is available to any embedded subsystem of the world, because an embedded subsystem could go on collecting information forever—amassing as much knowledge as you like about what has happened at various times and places, without ever being in a position to determine the total state, because no collection of local matters of particular fact suffices to fix the total state of the world. And it is not information that is available to causally regulate events in the world.

This works in a Newtonian space-time where there is no restriction on the past events that might be relevant to some future event, and for reasons related to this

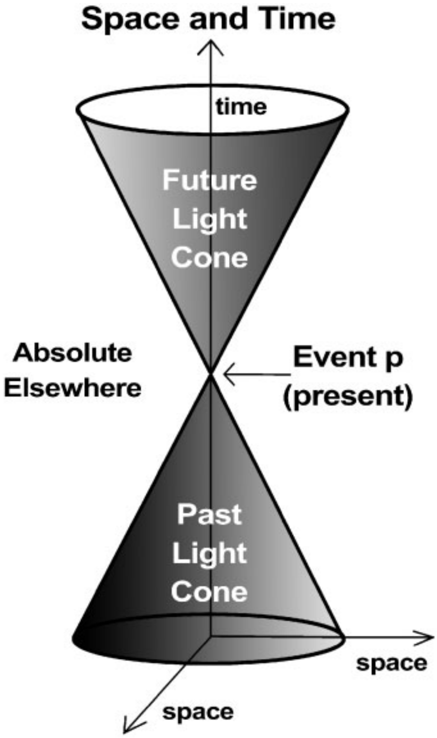

Figure 1. Light cones in Minkowski space-time

Minkowski space-time (the space-time associated with Special Relativity) is usually regarded as the most hospitable environment for determinism. In Minkowski space-time, the only spatiotemporal structure is the light cone structure, which separates space-time into three regions at every point p.

The light cone structure is defined by the speed of light. Since nothing travels faster than the speed of light, only points in the past light cone can affect what happens at p. Points in the future light cone can be affected by what happens at p. And points in the absolute elsewhere can neither affect nor be affected by what happens at p. Points in the absolute elsewhere are for all intents and purposes invisible to p.

It turns out that the situation for determinism is no better there when we consider events that lie at even a tiny finite interval in the future. In Minkowski space-time, we *can* explicitly specify a set of events that would be nomologically sufficient to determine any future event e. e's past light cone, or any temporal cross-section of e's past light cone will suffice. But if e lies at some finite time in the (absolute) future from e^{-n}, then nothing in e^{-n}'s past light cone nomologically determines e. No matter where we draw the cross section of e's past light cone, it will include events that are *not* in e^{-n}'s causal past. This means that there are always influences that are outside e^{-n}'s causal past but relevant to a future event e. Consider the past light cone of successive points along V_0 projected onto the surface A_0, below.

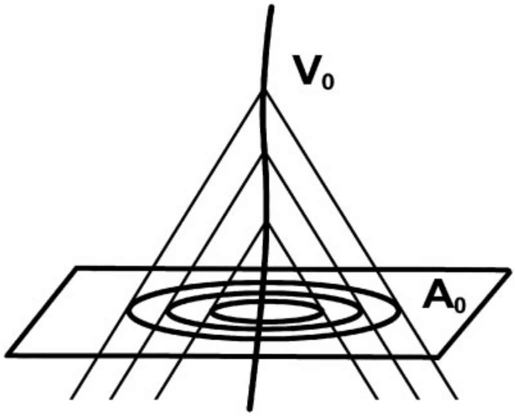

Figure 2. Past light cone of successive points projected onto a surface

It is tempting to interpret this as a purely epistemic point; a resolution of the puzzle that puts the knowledge needed to generate a prediction outside the ken of an embedded predictor. This would be a solution that ruled out natural oracles as a purely epistemic matter. The thought here is that new information becomes available at every stage of history which was not even in principle available beforehand to the inhabitants of the universe, but the information is *already there,* on its way (so to speak) to intersect our future light cones. To have that thought, however, is to impose a conception of *the past* on a Minkowski space-time that it doesn't support. The only meaningful spatiotemporal order in Minkowski space-time is the causal order embodied in the light cone structure. There's a well-defined order for events that fall in one's past causal horizon, and a well-defined order for events that fall in one's future causal horizon, but no well-defined order (relative to here-now) for events that fall in the absolute elsewhere. So there's no objective sense to be made of events that have *happened already,* but about which information isn't *here yet.*

The causal order is perfectly well behaved here. What we can't have on pain of allowing for the implementation of a self-defeating chain is a chain of events that leads from a set of events sufficient to fix the future into the causal past of a counterpredictive device. If C is a set of events nomologically sufficient to determine e, and Δt is the interval needed for a counterpredictive process, then C can't be causally fixed at $e-\Delta t$. That situation never arises in Minkowski space-time because the events in C aren't, as a class, fixed until e itself. More precisely, if p_0 is here-now, p is an event in the absolute future of p_0, and $\{e^*\}$ is the set of events that fill out the space between the light cones of p and p_0, then for any one of the events in $\{e^*\}$, there are frames in which it occurs before p_0 and frames in which it occurs after, and as a class, the events in $\{e^*\}$ aren't all fixed until p. In general in Minkowski space-time, even though the dynamics is deterministic, there is no way of fixing a set of events nomologically sufficient for e in e's causal past with enough of an interval to set up a counterpredictive chain.[11]

This gives us insight into the beautiful consistency of Minkowski space-time. It is not an accident that the worry that determinism seems to present (i.e., that our future is fixed by our past) goes away when the temporal ordering reflects the causal ordering (i.e., when 'the past' means 'the causal past', i.e., the set of events that objectively causally precede the here and now). Determinism and Minkowski space-time are a benign pairing.

Why information about the total state of the world is information from the future

By comparing the prospective and retrospective of two points in Minkowski space-time, we can get a clear diagnosis of why information about the total state of the universe at an early time imports information from the future. That, in its turn, explains why making that information available to regulate events in space and time gives us an analogue of closed causal loops.

We saw in Figure 2 that if we consider an event at any finite interval in the future, the causal past of the latter includes events that are *not* in the causal past of the earlier. This is the relativistic realization of the common-sense idea that tomorrow has a different causal past than today, because it includes all of the events that have happened in the time in between. The difference is that 'in between' has a new relativistic reading. It includes events that fall outside the past light cone of yesterday (the space between the light cones on the surface A_0 in Figure 2), but that common sense would think of as in the distant past. In Minkowski space-time, looking *pro*spectively, the laws allow indefinitely many ways of extending the causal past. Looking *retro*spectively, an event is fixed by its own causal past. The information that becomes available in that interim period between the prospective and retrospective point of view (information about what happens *in between* now and then) is not itself constrained by the causal past and is essential to nomological fixation.

There are two (nomologically interchangeable) ways of characterizing what happens in between. We can take a cross section of the past light cone of the later event and evolve it forward, or we can take a cross-section of the past light cone of the earlier event, together with exogenous variables that would cross the world line separating them in the meantime. In both cases, there are new events that aren't part of the causal past and that are nomologically necessary to fixing the future. The relativistic analogue of temporal development, from any point of view, is a kind of expansion upward and outward, expanding along the spatial dimensions as we move up the temporal, or, to put it a little less metaphorically, expanding outward along the spatial dimensions as we compare perspectives at later moments along a timelike curve. As we get more future, we also get more past.

The retrospective view from the end of time

There is a thing we can do, however, that is entailed by determinism. Choose any inextensible timelike curve, and wait until the end of time, and then take a cross section of the past light cone of that last moment shortly after the Big Bang (or at the earliest time the equations become applicable). That cross section will capture the causal past of all events that ever make an impact on points along the chosen worldline. In an infinite universe, there would be no last moment, but we can follow common practice extending the world-line to future infinity, and we take a cross section of the past light cone of that point.[12] When we describe things this way, it is almost irresistible to think that even if it comes fully into view only at the end of time, what is revealed at the end of time is, surely, what things were like *already* at the beginning.

Again, there is a basis for saying that in Newtonian space-time, but once we move to a relativistic regime, there is no basis for thinking that events in our absolute elsewhere are 'already there' and on their way to cross into our future light cone. The events in the absolute elsewhere that are nomologically necessary to fix a future event e are no more or less fixed than e itself. They don't, in any objective sense *happen before e.*

Pretheoretically, there is a large intuitive difference we draw between space and time. If you cast your eyes across a landscape, you think of the landscape as a fixed object that simply comes into view in stages. You think of time, by contrast, as coming into being as it is experienced. And the remnants of this prerelativistic way of thinking guide the imagination even for the most seasoned physicist. When I describe this picture of development as growth along both dimensions, we instinctively think of the growth along the spatial dimensions as simply revealing what was 'already there'. Determinism plays on this prerelativistic distinction between space and time. The idea was that if you thought the future was open in a manner that the past was not, determinism was supposed to undermine that conviction. Determinism was supposed to leverage the fixity of the past and the laws into the fixity of the future by showing that the future is nomologically determined by the past.

The move from Newtonian to relativistic physics eliminates the intuitive difference between space and time. There is a shift from an object-based to an event-based ontology. The basic entities are events. The geometry now encodes causal structure. There remain differences between the spacelike and timelike dimensions[13] but the basis for treating space as a substance and time as the dimension of becoming is gone. The new setting also lets us draw the distinction between what is fixed and what is open in a more refined way. Where before 'fixed' meant something like 'exists already'. In the new setting, 'fixed' means something like 'unalterable from here and now'/'lies in the causal past'/'impervious to action'. 'Not yet fixed' means 'alterable from the here and now'/'sensitive to action'. The result of these changes is the picture that I described above of growth along both dimensions, with events in the absolute elsewhere having the same status as events in the future, and both inheriting a kind of openness from their connection to action. It is often said that relativity spatializes time. It might better be said that it temporalizes space, because in a relativistic regime we can no longer think of the distant parts of space as 'there already' but coming gradually into view. The spatial past is just as infected by 'becoming' as the temporal future.

Let me repeat this in a slightly different way. The information about exogenous variables that are needed to get from p_0 to p_1 is contained in the world in two nomologically interchangeable ways: (1) if we draw a spacelike hypersurface at any past time that cross-sects both light cones, it is contained in the events along that surface between the two light cones,[14] or (2) it is distributed across the world line connecting p_0 to p_1. This goes for any two timelike related events, and it goes, in particular if we let p_0 be the initial and p_1 final points along an inextensible timelike curve. What this means is that the information that is contained along a cross section of the past light cone of the last moment, is not contained in the past of any earlier, but spread along the whole temporal dimension between the initial and final moments.[15] In this rather subtle way, information about the total state of the universe along the spatial dimensions in the past that is needed to fix its future *is information from the future*.[16]

We can, of course, introduce a global present by convention, or choose to treat events in the absolute elsewhere as fixed. But this kind of convention would only obscure the causal structure. What matters when we are assessing the implications of determinism is whether the global facts are fixed in a way that motivates treating them as constraints on action. And so what matters when we are assessing whether the global state is fixed *already* from some point in space-time is whether it is fixed by the causal past at that point. From an embedded point of view, events in the absolute elsewhere are literally *nowhere*. Events in my absolute elsewhere become locatable relative to me in the future, when information about them crosses my world-line.

Temporal vocabulary—'fixed already', 'not yet fixed'—only makes sense from an embedded point of view, and there is literally no embedded point of view in Minkowski space-time from which the global state of the world at some past time is fixed already. The only sensible way to treat global facts is as emergent from the totality of what happens on the ground, and at any given point they are only ever as fixed as the causal past. The relativistic setting gets things right here precisely *because* in a relativistic setting the temporal order reflects the causal order. The bird's eye view that leads us to suppose there is some fact now about the total state of the universe, even if it is epistemically inaccessible to us, is explicitly disallowed in a relativistic setting. In that setting information about the total state is clearly information from the future.[17]

If it seems magical that determinism *could* hold at the global level if there *isn't* some kind of global constraint on temporal development, consider the example of global conservation principles. Noether's theorem gives us a deep understanding of how global principles are related to local symmetries,[18] but even qualitatively, it is not difficult to see that so long as the laws that govern local exchanges of energy are conservative—i.e., so long as they don't allow us to raise (or lower) the energy of an open system without drawing (or depositing) energy in a corresponding amount from an external source—energy will be globally conserved. Something quite similar is true in the case of determinism. We build up wholes by respecting microscopic laws that govern the local exchange of matter and energy. So long as the laws that govern those exchanges are deterministic in the sense that they entail that there are no differences in the causal futures of points with indistinguishable causal pasts unless there is a difference in their respective absolute elsewhere, we can be assured that global solutions will be deterministic. Laws of this kind treat the absolute elsewhere as a resource for generating novelty along the temporal dimension. Global determinism is a degeneracy that results from suppressing the source of novelty in the world.

Disambiguating determinism

Recall the definition of determinism given in Part I. We said a universe U is deterministic when, for any chosen time t_0, the physical laws map the state of U at t_0 onto the state of the universe at any arbitrarily chosen later time. And we defined a natural oracle as a subsystem in U that is able to predict all events in U given knowledge of its past. This means that the existence of a natural oracle follows from determinism just in case it is possible for an embedded subsystem to have knowledge of the state of the Universe at any past time. And since a natural oracle is just a physical system, a natural oracle acts as a causal pathway or conduit for information that makes the total state of the world available to regulate events in space and time. That is why combining natural oracles with counterpredictive devices gave us self-defeating causal chains. We spent the last few sections explaining why determinism doesn't entail the existence of natural oracles. In epistemic terms, there are no natural oracles because there is no way for an embedded system to know the total state of the world at any past time. In causal terms, the loopy structure that would arise if we could use events nomologically sufficient for a chosen future event e to set up a counter-e causal chain don't arise for reasons that are clearest in Minkowski space-time, where the spatio-temporal order reflects the causal structure.[19] We are in a position now to distinguish different notions of determinism.

Local Microdeterminism. This holds just in case the laws that describe local exchanges of a conserved quantities are such that the evolution of any open system is a function of its initial state and the exogenous influences impinging on it.

Causal Determinism. This holds just in case every event is necessitated by earlier events.[20]

Laplacean Determinism. This holds just in case the future of the universe as a whole is completely determined by the state along any complete spacelike hypersurface.[21]

Local microdeterminism entails Laplacean determinism and both hold in our paradigmatic examples of deterministic theories: Newtonian Mechanics, classical mechanics, and electromagnetism on Minkowski space-time.[22] Causal determinism does not follow from local microdeterminism for the reasons that we have seen, and doesn't hold in these theories.

Although many people have drawn the implication of causal determinism from Laplace's famous image of a predicting intelligence, Laplace himself wasn't confused about this. What he says is:

> An intellect which at a certain moment would know all forces that set nature in motion, and all positions of all items of which nature is composed, if this intellect were also vast enough to submit these data to analysis, it would embrace in a single formula the movements of the greatest bodies of the universe and those of the tiniest atom; for such an intellect nothing would be uncertain and the future just like the past would be present before its eyes.[23]

He is very clear that the intellect would not just have to know the positions and momenta of each of the items of which nature is composed, but would have to know *all* of them to predict the future. Of course, he was living in a Newtonian era and the difference may not have seemed significant to him. But the habit of precise thinking served him well here.[24]

PART III. THINKING IT THROUGH FROM AN EPISTEMIC PERSPECTIVE

So far I've been working on dispelling the causal puzzle that seems to arise from counterpredictive devices in a deterministic setting. But even once worries about a conflict with determinism are dispelled, the situation remains puzzling from an epistemic point of view. It would be nice to have a clearer understanding of the logic of the inference that a would-be oracle faces in the presence of a counterpredictive device.

There are several questions that need clarification:

i. Given that there is no difficulty with predicting the behavior of a counterpredictive device when someone else feeds it a prediction, why is there a difference when it is your prediction that is given as input?
ii. Is there some connection to anomalies of self-reference?
iii. What, if anything, does this have to teach us about our own epistemic situation in the world?

Interaction effects

From a logical perspective the problem arises because even though it is missing no information, a predictor can't generate a correct prediction because the nomological determinants of the predicted event includes the prediction itself. That means that in order to make a correct prediction, the predictor has to predict *what it will predict*. As soon as it predicts what it will predict, it has to update its prediction on that information. When it does that, it will have to change its prediction . . . and so it goes. If the predictor is smart it will see that the relationship between the prediction and the event that *makes the prediction true* puts it in an impossible position.

The same sort of situation arises for you if you've written a simple, deterministic, counterpredictive computer program and challenge yourself to predict what it will do under the constraint that you reveal your prediction to it. You try to predict it by simulating the program. To do that, you need to know what the input to the device is, but the output to the simulation is input to the device. Even though the device is deterministic, you know in advance that your prediction is bound to fail. It's a mug's game because the thing is arranged so that the prediction will interfere negatively with the predicted event. You supply the input that defeats your prediction as soon as you make it.

So if we compare the case where the aspiring predictor keeps his prediction hidden and the case where he reveals it, the difference between them is the presence of interference effects. It doesn't matter what information a would-be oracle has about the past. So long as every calculational route from the initial data to the predicted event passes through the prediction, any calculation aimed at prediction will fail because the prediction will interfere negatively with the predicted event. For a system that acts in the domain that it is trying to predict, interference effects are inevitable. For many purposes they can be ignored. What the counterpredictive device does is zero in on the interference effects and create a situation in which they interfere negatively with the output.

Here's an even tighter way of getting negative interference. Suppose you had an aspiring oracle and that it had some sort of output channel that it used to display answers to questions about the universe. Ask it "is the answer to this question that's about to be displayed in the output channel 'no'?" It can't correctly answer. It can't correctly give the answer because delivering the answer undermines the prediction. The prediction *does* the opposite of what it *says*. Now we can see how the Paradox of Predictability is connected to anomalies of self-reference. The Paradox of Predictability has the same structure. It creates a situation in which the prediction undermines itself, except that there is a time lag between the prediction and the predicted event, and the link between them is causal rather than logical. In general, if a prediction (whatever form it takes) is connected in the domain being predicted, and there are no nomologically sufficient routes to the predicted event that don't pass through the prediction, interference effects will arise. In probabilistic terms, the prediction is going to screen off the information provided by any causal routes to the predicted event that pass through the prediction and will place a limit on advance certainty. It will do this even though the setting is deterministic.

So the full resolution of the puzzle has two sides. First, we need to see why determinism doesn't entail that there were events in place in the distant past that were nomologically sufficient for anything that happens after, so we can put aside the expectation of a perfect predictor. And second, we need to see how interaction effects arise. If your predictions create disturbances in the domain you are trying to predict, it is not surprising that they limit predictability.[25]

How generic and widespread are these effects? If you are trying to make cosmological predictions, guess who is going to win an Oscar, or predict how lower temperatures will affect tides in New Zealand, they can be all but ignored. They are, however, extremely disruptive in the social domain. How people act depends in all kinds of ways on what they think you expect them to do. Any person with a mind to be disagreeable can use your predictions to undermine them if they are made known. They are also pervasive in public life. Making election predictions public influences who comes out to vote. People who play the stock market are using other people's predictions to guide their own behavior.

The upside of interaction effects

If information-gathering and utilizing systems are part of the world that they are gathering information about, their own activities are going to be connected in the domain and will produce interference effects. This will go for robots and aspiring oracles and artificial chess agents and for people acting in the world they perceive.[26] It also goes for you. It doesn't matter if the situation is globally deterministic and what advance knowledge you have. Since your predictions are connected in the world, the interference effects created by your predictions are going to screen off the information provided by all antecedent causes that pass through them.

That's a bad thing when interaction effects created by your predictions are used to undermine them. But interaction effects have an upside. Notice that if you can create a self-undermining question, you can also create a self-affirming one. Go back to our would-be oracle and ask: "is the answer to the question that is about to be displayed in the output channel 'yes'?" Or, better: "what is the answer to the question that is about to be displayed in the output channel?" There is no way for the oracle to answer falsely. That means that it can choose its answer free of epistemic constraints. And if the answer has interference effects—i.e., if the answer is probabilistically connected to other elements in the domain—the oracle can choose so that the interference effects created by its answer works to its own advantage. It can choose, that is to say, the answer whose interference effects it prefers.[27]

It is built right into the on-the-ground causal order that how the past produces the future passes through information-gathering and decision-making processes of creatures like you and me. And the epistemic situation for creatures like us combines self-prediction with interference. This all has to be dealt with rather delicately, but the trick is to see decision as a kind of self-fulfilling prediction. It is not always true that you do what you decide to do, but it *is* always true that you *decide* to do what you decide to do. Your decisions are like the self-fulfilling answers to the question about what is going to be displayed on the output screen. For anybody else trying to guess what you will do, predicting your choices is difficult. For you, in epistemic terms, it is trivially easy. "What will I decide to do here?" is like "What is the answer to the question that is about to be displayed in the output channel?" Whatever you put in, there is no chance of affirming a falsehood.[28] And so long as your decisions are probabilistically connected to action, they will provide free information: information that you create, in a way that is alethically unconstrained, and information that will feed into your predictions for the future.

This means that you have the freedom to set the value of certain variables (your choices) without fear of falsehood. In practical terms, the fact that your choices interfere positively with the future is a way of saying that you have the power to affect the future.[29] In this way, for systems like us, in the business of using information to guide behavior, the difference between prediction and intention arises naturally from the combination of the self-affirming character of choice and the interference effects choices create in predicting the future.[30] Elizabeth Anscombe famously asked what the difference is between prediction and intention, e.g., between 'I'm going to be sick' and 'I'm going to go for a walk'? In epistemic terms, the signature of an intention is that it is a self-affirming act that interferes positively with what it predicts.[31]

The epistemic situation of embedded observers: managing the interaction effects

Embedded agents know instinctively that their choices create interference, and (hence) that they can't stabilize belief about their future until their choices are made.

And they make choices with an eye to bringing about a future they find desirable.[32] The degree of probabilistic dependence on choice is going to provide a measure of how sure predictions can be in advance. The strength of interference between my choices and other events will depend on how my choices are connected in the world. I don't deliberate about whether the sun should rise tomorrow because my decisions have negligible probabilistic effect on whether it will. My decisions about whether to raise my right hand a second from now, by contrast, have a high probabilistic effect on whether I will. My decisions about what time to wake up tomorrow and how much to exercise this month have a somewhat weaker effect, because I'm prone to failures of will. If you make a decision and conditionalize on the decision, the conditional expectation is a measure of (your estimation of) the effectiveness of your decision. My decisions are immediately effective only in controlling the movements of a particular chunk of matter (my body). Your decisions are effective in controlling yours. Each of us is, in this way, has a different epistemic/practical perspective on the world.

The cleanest way to take interaction effects into account in forming expectations for your own future is let yourself decide what to do, treating your decision as a free variable, then form expectations for what you *will* do by conditionalizing on what you decide. That's a start, but it's a bit too simple for various reasons.[33] I think that the right thing to say is probably that your decisions don't *trump* prediction. Rather they *feed into* prediction, without being constrained by it. So the common sense attitude to the future is roughly right: your expectations should be appropriately tempered by your confidence that you will do what you decide, but your decisions are epistemically free information.

However one manages interference effects in forming expectations, they are inevitable, perfectly compatible with determinism, and simply an artifact of the fact that our information-gathering and decision-making processes are connected in the world and part of the domain we are getting information about. If agency is acting in the domain you are representing, *intelligent* agency is realizing that you are acting in the domain you are representing, and exploiting interaction effects to create a desirable future. All of this is a fancy way of saying that some of what you have to predict is stuff that you do, and to the extent that the stuff that you do is connected in the field over which belief is defined, the stuff that you do is going to interfere with other things. And if you are smart, you will use that interference to your advantage.

Does this have anything to do with free will?

I've refrained from drawing conclusions for free will. All that we can say with definiteness from a physical perspective is that there is a perfectly consistent picture that starts with the local microscopic laws, supports the emergence of an on-the-ground causal order in which human decision is not fixed by the causal past before one has made a decision, has the kind of spontaneity that we pretheoretically suppose, and leads to global determinism. Your ability to override any revealed prediction of what you will do is entirely compatible with determinism, and the openness of the future is not only an epistemic openness that arises because nature has carefully contrived to hide things from you that would, if they were known, destroy your sense of freedom. Nature is entirely consistent here. The route by which the facts about the past determine your actions passes through your decisions. They are not fixed by anything that has occurred right up until the moment they are made, and you retain until the last second the practical ability to override anything that you already decided or that has been predicted about what you will decide.

It is worth mentioning a related argument given by Seth Lloyd to the effect that the easiest, and only sure way, of finding out what you will decide is to deliberate.[34] The argument is a computational argument. You are supposed to think of yourself as a Turing Machine trying to predict your own upcoming decision. The claim is that you cannot answer all questions about your future behavior because no general technique exists to determine whether or not you will come to a decision at all (because of the halting problem). And you can't in general know beforehand what your decision will be on some matter without (doing something at least as complex as) simulating the decision-making process. This is presented as an explanation of 'why we think that we are free'.

This is a rather peculiar argument for reasons connected to what we've been saying here. The inability to predict in advance has nothing to do with the absence of a more computationally efficient method for predicting than deciding. If you tried to simulate the decision, you could use the results of the simulation to undermine it by acting counterpredictively. You don't have to take any kind of detour through computational limitations to establish that. Nor is solving the halting problem going to change it. The output of the simulation becomes a piece of information that is available as input to the decision.

The ability to make decisions free of epistemic constraints and to confute any revealed prediction of what our decisions will be has a much better claim to explaining 'why we think we are free', and indeed a much better claim to capturing the sense in which we *are* free. Indeed, it was because determinism seemed to endanger that ability—to reveal it as illusory—that it seems so threatening.

CONCLUSION

I began with a puzzle that showed that the presence of counterpredictive devices places limits on the ability to predict the future even in a deterministic setting, and it wasn't clear how this could be since determinism seemed to allow for the possibility of natural oracles, even though it was demonstrably compatible with the existence of such devices. The puzzle was dispelled because it turned out on examination that determinism does not entail the possibility of natural oracles. The only route in space and time *from* an aspiring oracle's causal past *to* the output of a counterpredictive procedure passes through its prediction, and any attempt to predict in advance will interfere negatively with itself.

I used the puzzle to draw two lessons: a negative lesson about what determinism doesn't entail, and a positive lesson about how interaction effects structure our epistemic lives.

There is a more general lesson here, though, that I think is worth bringing into the foreground: a lesson about the difference between prediction from within (sometimes called simulation in the 'hard-mode') and prediction from without. People elide the difference between these two things because they presume that any restriction on what can be known about the past is accidental. Prediction from without is straightforward. If the setting is deterministic and you know the total state of the universe, you can calculate its full temporal development. Prediction from within is characterized by two nonaccidental restrictions: (i) it requires you to use knowledge available in the universe to calculate its future, and (ii) it is essentially limited by interference effects. These restrictions aren't unconnected. The very transcendent point of view that gives you access to the total state detaches you from the one that lets you undermine it. In physics, it makes sense to adopt a transcendent perspective for various reasons.[35] But I think that we've been too casual about the legitimacy of that transcendent perspective in drawing philosophical conclusions from physics.[36]

NOTES

1. This simple, generic definition will suit our purposes best, but see John Earman, *A Primer on Determinism* (Dordrecht, NE: Reidel, 1986) and Carl Hoefer, "Causal Determinism," *The Stanford Encyclopedia of Philosophy* (Spring 2016 Edition), E.N. Zalta, ed., https://plato.stanford.edu/archives/spr2016/entries/determinism-causal/.
2. If a simple program could do it, a more sophisticated one could as well, e.g., a machine that ran through a decision procedure that took rather more into account and that could respond to any representation of its future action counterpredictively, depending on its other desires on the occasion. A system like that might 'take it into its mind' to confute a prediction and could do so to prove a point if it came to that. It could play the 'you tell me what you think I will do, and I will do the opposite' game and win. It needn't always (however) do so. A system like that would be as unpredictable as a counterpredictive device, and again the mere deterministic setting is no barrier.
3. Michael Scriven, "An Essential Unpredictability in Human Behavior," in Benjamin B. Wolman and Ernest Nagel, eds., *Scientific Psychology: Principles and Approaches* (New York, NY: Basic Books, 1965). David K. Lewis and John S. Richardson, "Scriven on human unpredictability," *Philosophical Studies* 17:5 (1966), 69–74. Donald M. MacKay, "On the Logical Indeterminacy of a Free Choice," *Mind* 69 (1960), 31–40. Karl Popper, "Indeterminism in Quantum Physics and in Classical Physics," *British Journal for the Philosophy of Science* 1:2/3 (1950), 117–33, 173–95. L.D. Roberts, "Scriven and MacKay on Unpredictability and Free Choice," *Mind* 84 (1975), 284–88. Donald M. MacKay, "Choice in a Mechanistic Universe: A Reply to Some Critics," *British Journal for the Philosophy of Science* 22:3 (1971), 275–85. Donald M. MacKay, "The Logical Indeterminateness of Human Choices," *British Journal for the Philosophy of Science* 24:4 (1975), 405–408. I.J. Good, "Free Will and Speed of Computation," *British Journal for the Philosophy of Science* 22:1 (1971), 48–50. Alvin Goldman, "Actions, Predictions and Books of Life," *American Philosophical Quarterly* 5:3 (1968), 135–51. Richard Boyd, "Determinism, Laws, and Predictability in Principle," *Philosophy of Science* 39:4 (1972), 431–50.
4. Daniel Dennett, *Elbow Room: The Varieties of Free Will Worth Wanting* (Cambridge, MA: MIT Press, 1984).
5. Jenann Ismael, *How Physics Makes Us Free* (Oxford: Oxford University Press, 2016). http://web.mit.edu/holton/www/pubs/determinism&fatalism.pdf. See the short discussion in "The Challenge" at the very end. And there is a truly excellent recent discussion in Stefan Rummens and Stefaan Cuypers, "Determinism and the Paradox of Predictability," *Erkenntnis* 72:2 (2010), 233–49.
6. Popper was interested in playing down the novelty of the indeterminism of quantum mechanics. Rummens and Cuypers use it to demonstrate the compatibility of determinism and what they call a take-it-or-leave-it control over revealed predictions of their future behavior. Mackay is interested in the logic of self-prediction. Holton uses it to defeat the argument from determinism to the attitude of resignation advocated by fatalists.
7. John Earman, "Implications of Causal Propagation Outside the Null-Cone," *Australasian Journal of Philosophy* 50:3 (1972), 231.

8. John Wheeler and Richard Feynman, "Classical Electrodynamics in Terms of Direct Interparticle Action," *Reviews of Modern Physics* 21 (1949), 425–34. There is a qualification to this. Maudlin and Arntzenius point out that under certain conditions there will be fixed points that allow solutions that introduce no accidents or anomalies. Frank Arntzenius and Tim Maudlin, "Time Travel and Modern Physics," *The Stanford Encyclopedia of Philosophy* (Winter 2013 Edition), E.N. Zalta, ed., https://plato.stanford.edu/archives/win2013/entries/time-travel-phys/. These conditions won't in general be satisfied by counterpredictive devices.
9. I use the causal vocabulary in a minimal and neutral way. For our purposes a causal interaction is an interaction that involves the exchange of a conserved quantity. A causal process is a chain of causal interactions. When I talk about the on-the-ground causal order, I mean the collection of nomologically supported relationships between the intrinsic contents of specified volumes of space-time. There are richer notions of causation available in addition to these: e.g., causal process notions and interventionist causal pathways. See Jenann Ismael, "Causation: An Overview of our Emerging Understanding", in Samantha Kleinberg, ed., *Time and Causation in the Sciences* (Cambridge, UK: Cambridge University Press, 2019).
10. Thanks to Mark Wilson for suggesting this objection.
11. In Newtonian space-time, that is because there is no set of events that are jointly sufficient to fix e nomologically. As we saw, for any collection of events, there are models that include those events in which e happens, and ones that don't. To get nomological entailment, one always has to add that those events constitute the totality of what there is. But that latter fact isn't itself a causal agent. If we keep all of the specifically global properties out of view in describing the on-the-ground causal order, the second order properties will take care of themselves as emergent properties of completed totalities.
12. Relax philosophical scruples about the legitimacy of the common practice. They may be legitimate, but nothing here hinges on skepticism about infinities.
13. See Craig Callender, *What Makes Time Special* (Oxford, UK: Oxford University Press, 2017) for a deep look at these differences, from the fact that the metric has Lorentz signature, through the contingent differences in the way that matter is distributed.
14. More generally, if we draw a spacelike hypersurface at any time that cross-sects both light cones, it is contained in the events along that surface between the two light cones.
15. In case there *is* no final moment, as I said, we introduce a point at future infinity.
16. That and *more*. The retrospective perspective from the last point on a given world line will not in general capture the causal past of the universe as a whole, because the past light cone of spacelike related points are not coextensive. We don't get a truly *global* view unless we consider the retrospective view from that last moment along *all* inextensible timelike curves in the universe. So we consider future timelike infinity, which is the endpoints of all future-directed timelike curves. That means that (except in very special cases) there is not a point in space-time from which one can get a global view of the universe.
17. Information *from* the future is different from information *about* the future. Information *about* the future is information derived from the past by application of laws. Information from the future cannot be so derived. It is like information from crystal balls.
18. Noether's theorem establishes that every differentiable symmetry of the action of a physical system has a corresponding conservation law.
19. It doesn't arise in Newtonian space-time because there is no set of events that is nomologically sufficient for a future event (you can always defeat the nomological entailment between any set of past events and a future one by adding events to the past). It doesn't arise in Minkowski space-time, because there is a set of events that is nomologically sufficient but they are contained in e's causal past with room to insert a counterpredictive chain.
20. I've adapted the definition from the first line of Carl Hoefer's encyclopedia entry on Causal Determinism. Hoefer writes "Causal determinism [is the doctrine] . . . that every event is necessitated by antecedent events and conditions together with the laws of nature" (Hoefer, op. cit.). I've substituted 'earlier' for 'antecedent', and mean it to suggest events at some finite temporal remove in the past.
21. Or, equivalently, any complete spacelike hypersurface is a Cauchy surface.
22. Determinism has a complex status in General Relativity, and I've refrained from talking about it here. See Earman (op. cit.) and Hoefer (op. cit.).
23. Pierre Simon de Laplace, *A Philosophical Essay on Probabilities*, trans. Frederick Wilson Truscott and Frederick Lincoln Emory (New York, NY: John Wiley and Sons, 1902).

24. On the other hand, he also gives determinism a causal formulation. He precedes this passage with the sentence: "We may regard the present state of the universe as the effect of its past and the cause of its future." ("*Nous devons donc envisager l'état présent de l'univers comme l'effet de son état antérieur, et comme la cause de celui qui va suivre.*") The points here militate clearly against thinking of global states as causal agents, and against the existence of any necessary connection between events at one time and events at any temporal remove.
25. It doesn't matter what kind of domain is in question or how the elements in the domain are connected. So long as there are fixed points and probabilistic connections between the fixed points and other events in the field. For an interesting discussion of the way these issues arise in constructing AI, see https://intelligence.org/files/RealisticWorldModels.pdf and http://intelligence.org/files/ProblemsSelfReference.pdf.
26. There are two different ways of regimenting: we can think of the activity of an information-gathering and utilizing system either as the internal information-gathering and utilizing processes or as the publicly observable physical movements controlled by those activities. In either case, decisions are going to end up being connected in the domain. In the first case, they will be identical with some elements in the domain, in the second case, they will be probabilistically associated to elements in the domain (and hence, informative about them). Consider a robot with sensors, for example, that gathers and processes information that guides its behavior. Suppose that it has the task of moving through a complex domain without getting hit by other moving objects. If it is trying to predict whether it is going to get hit by an object moving towards it, it has to take into account its own movements. It knows that its movements are controlled by its choices, and that means that to predict its movements it has to predict its choices. If it doesn't want to get hit, it should make choices that minimize the probability of doing so.
27. Indeed, you might just say that preferences are what guide answers when epistemic constraints become degenerate.
28. This is also why 'wait and see' is not a strategy that works for questions about one's own upcoming decisions. Waiting won't generally produce a decision, and without a decision, there is nothing to see.
29. I'm taking for granted here that the probabilistic connections over the field of belief are intended to reflect nomological connections over the field of events, and I'm assuming for present purposes that they do so accurately. Of course they needn't. Taking account of the epistemic uncertainty would add a layer of complexity that I'm suppressing here.
30. The logic of this is something that I, and others, have written about. Jenann Ismael, *How Physics Makes Us Free*, (New York, NY: Oxford University Press, 2016). Jenann Ismael, "Decision and the Open Future," in Adrian Bardon, ed., *The Future of the Philosophy of Time* (New York, NY: Routledge, 2012), 149–68. James Joyce, "Levi on Causal Decision Theory and the Possibility of Predicting One's Own Actions," *Philosophical Studies* 110:1 (2002), 69–102, and "Are Newcomb Problems Really Decisions?" *Synthese* 156:3 (2007), 537–62. Huw Price, "Causation, Chance, and the Rational Significance of Supernatural Evidence," *Philosophical Review* 121:4 (2012), 483–538. David Velleman, "Epistemic Freedom," *Pacific Philosophical Quarterly* 70 (1989), 73–97. Wlodek Rabinowicz, "Does Practical Deliberation Crowd Out Self-Prediction?" *Erkenntnis* 57:1 (2002), 91–122. Brian Skyrms, *The Dynamics of Rational Deliberation* (Cambridge, MA: Harvard University Press, 1990). Isaac Levi, *The Covenant of Reason: Rationality and the Commitments of Thought* (Cambridge, UK: Cambridge University Press, 1997). Yang Liu and Huw Price, "Ramsey and Joyce on Deliberation and Prediction," *Synthese* https://doi.org/10.1007/s11229-018-01926-8 (2018), and "Heart of DARCness," *Australasian Journal of Philosophy*, DOI: 10.1080/00048402.2018.1427119 (2019).
31. Elizabeth Anscombe, *Intention* (Cambridge, MA: Harvard University Press, 1957). This is not a perfect example because the thought that you are going to be sick can induce the feeling. A better contrast might be 'I'm going to get old' vs. 'I'm going to go for a walk'.
32. In epistemic terms, that is what deliberation is: assessing choices for their downstream consequences, with an eye to maximizing (something like) expected utility.
33. For one thing, decisions are appropriately tempered by expectations for carrying them out. If I'm deciding on an exercise routine, although 5 days a week would be preferable, if I know I'm more likely to stick to a 3 day/week regimen, that will affect my decision. There are subtleties in sorting this all out, and wide disagreements in the formal epistemology literature. See note 30.
34. https://arxiv.org/pdf/1310.3225.pdf

35. For reasons that are, in fact, closely connected to what we've been talking about here. Things are much simpler at the global level. There are global symmetries and conservation principles that hold only at the level of physical totalities. See Neil Turok's lecture "The Astonishing Simplicity of Everything" (https://www.youtube.com/watch?v=f1x9lgX8GaE) for an especially clear expression of just how simple things become.
36. I'd like to thank Gordon Belot enormously for discussion of the Paradox of Predictability and for bringing this paper to fruition. It would not have gotten written without him. I'd also like to thank David Albert, Laura Reutsche, Ryan Smith, Brien Harvey, and audiences at the New Directions in Foundations of Physics conference in Tarquinia, University of Maryland, University of Edinburgh, and Arizona State University for very helpful discussion.