

Radical indeterminism and top-down causation

Helen Beebe, University of Manchester

Please do not cite this version. The published version is: ‘Radical Indeterminism and Top-down Causation’, *Res Philosophica*, 91 (2014): 537-45 (part of a symposium on Helen Steward’s *A Metaphysics for Freedom*).

1. Introduction

The bulk of Steward’s rich and subtle account of agency, and at least part of her grounds for claiming that agency is incompatible with determinism (and indeed with a standard version of indeterminism), appears in Chapter 8, “Agency and causation”. In this short essay, I provide some reasons for thinking that a stronger case needs to be made for this incompatibility claim. In §2, I briefly sketch those elements of Steward’s overall metaphysical position that are relevant to the present discussion. In §3, I raise some objections to her claim that ‘top-down’ causation is incompatible with determinism, and in §4, I make some conciliatory remarks about the centrality of the debate about laws of nature to the question of where the burden of proof with respect to the discussion of §3 lies.

2. Indeterminism, agency and supervenience

The picture of the world that has come to dominate discussions of free will – and indeed metaphysics more generally – is one according to which either determinism or what I shall call ‘pseudo-determinism’ is true. According to determinism, the complete set of facts about the Universe at a given time together with a statement of the laws of nature uniquely determines the complete set of facts about the Universe at any later time. According to pseudo-determinism, the facts at a time plus the laws uniquely determine a probability distribution for the facts at a later time (where some outcomes may be assigned probability 1: pseudo-determinism does not entail that *no* future outcomes are fully determined by the past plus the laws). Pseudo-determinism is, of course, what most philosophers have in mind when they use the term ‘indeterminism’; I am abandoning common usage here in order to contrast pseudo-determinism with the more radical kind of indeterminism (hereafter labelled ‘radical indeterminism’) that Steward believes both to be required for agency and to actually obtain. In what follows, except where separating the two theses is required, ‘DPD’ is shorthand for ‘determinism or pseudo-determinism’ (or, sometimes, ‘deterministic or pseudo-deterministic’).

According to radical indeterminism, while the laws place *constraints* on what agents are able to do – they limit the range of available possibilities – they determine neither what agents will do nor some specific likelihood of their acting in a particular way. In articulating the kind of picture she has in mind, Steward appeals to Nancy Cartwright’s conception of laws (Cartwright 1999), according to which “such physical laws as there are, are narrow in their scope, strictly applicable only to the relatively small number of situations in that correspond closely to the physical models

that supply those laws with their concrete interpretations, and applying only *ceteris paribus*, even when they do apply” (Steward 2012, 120).

On Cartwright’s view, then, the idea that the mythical ultimate end-point of scientific enquiry will deliver a set of laws that covers absolutely everything that happens in the Universe is just a piece of philosophical dogma. And this is Steward’s view too. At least part of the reason for this, I take it, is precisely that the *ceteris paribus* clauses that we assume to apply to the ‘laws’ include, *inter alia*, lack of interference by agents (see in particular 230-1). A human hand can pluck a billiard ball off the table, a meandering goat might blunder into one’s carefully constructed experimental set-up, or (as happened at CERN in 2009) a passing bird might drop a bit of baguette into one’s large hadron collider. Of course, one might want to insist that the occurrence, or not, of such interventions is itself covered by deterministic or pseudo-deterministic laws; but Steward’s point is that, at least as things currently stand, we have no scientific or philosophical reasons to assume that this is so.

Of course, being an incompatibilist about DPD and agency, Steward is committed to thinking not only that the laws *do* not determine whether or not, for example, on a given occasion a human hand (mine, say) will pick up the billiard ball just as it was about to drop into the pocket, but that they *cannot*, if I am genuinely to *act* with respect to the ball at all. It’s important to be clear about Steward’s view here. The view is *not* that, were the laws to determine the movement of my hand after all, everything would be just as it is except that the movement would not legitimately be seen as an *action* – so that, writ large, if DPD were true, the world would evolve just as it does, except that it could only legitimately be described as a world in which human beings (and, for that matter, goats) are causally involved in what happens rather than a world in which humans and goats *do* things. After all, if the world is *not*, in fact, a DPD-world – and indeed, if it is not such a world partly because of the behaviour of humans and goats and other agents – then a world where DPD is true will be very different to the actual world at the level of what *happens*, and not merely at the level of what is and is not correctly describable as an action or an agent. The view, then, is rather that such a world is the kind of world where there simply are no beings possessing a sufficient degree of organisational integrity to behave in the way that we – and goats – do behave.

Nor is Steward’s view susceptible (or at least not obviously so) to the standard charge levelled at agent-causalists that the view suffers from irredeemable ontological obscurity. Admittedly, Steward maintains a pluralist view about causation – encompassing what she calls ‘movers’ (roughly: objects), ‘matterers’ (roughly: facts) and ‘makers-happen’ (events) – that might be considered somewhat ontologically profligate (though the case for both matterers and makers-happen is made in detail in her 1997). Agents, for Steward, are a species of mover. While agents, such as people and goats, exhibit a kind of control over their parts that mere movers, such as washing machines, do not – and, moreover, a kind of control that requires radical indeterminism – the overall picture is supposed to be consistent with the supervenience of facts about agents on facts about the physical. That is: no difference in facts, at a given time, at the level of agents without a difference in facts at a physical level of description. (Call this supervenience ‘at-a-time determination’.) So there is no whiff of substance dualism here: the claim is not that goats have souls.

The need, as Steward sees it, for radical indeterminism stems from what it takes for agents to be capable of the kind of ‘top-down’ causation that she takes to be required

for them to be in control of their parts in a way that a washing machine is not. That control, on her view, requires a rejection of *over-time* determination. Consider a picture familiar from discussions of the problem of mental causation (see e.g. Kim 1998). Our supervenience base – the neurological level, say – evolves over time, let's assume, according to deterministic laws. That being so, neurological goings-on at one time are causally sufficient for what happens at later times. Hence the 'exclusion problem': there now seems to be no causal role left to play for the mental, assuming that mental properties supervene on (but are not identical to) physical properties. In effect, Steward endorses a version of the exclusion problem: on her view, there would be no room for *agents* to cause their bodily movements if those movements had sufficient physical causes. As she puts it:

Perhaps if there truly were *necessitating* microphysical conditions in existence for the next global state of the universe at every stage, it really would be true that there was nothing more to be said about the causal metaphysics of any matter once the proximal sufficient conditions were given, they being in their turn secured by *prior* and once again entirely necessitating conditions. Here one can indeed see no gap into which a phenomenon like top-down causation might be fitted. But to suppose that this is so is precisely to insist upon the thesis of determinism. (240)¹

I shall argue in the next section, however, that it is not clear why we should accept that the existence of necessitating microphysical conditions excludes top-down causation.

3. Top-down causation, the exclusion problem, and indeterminism

The assumption manifested in the passage quoted above is that lower-level (e.g. microphysical) *lawful sufficiency* entails lower-level *causal completeness*: it entails that the full causal story can be told at the lower level (disregarding cases of overdetermination, at least). Hence the need to deny lower-level sufficiency, thereby finding a 'gap' into which top-down causation might be fitted.

One line of response to Steward, then, is to reject this assumption, and in principle there are two ways in which we might do this. First, we can accept that there is not only lawful but genuinely causal sufficiency at the lower level, but deny that this entails causal completeness. Or, second, we can deny that lawful sufficiency amounts to causal sufficiency, so that even if causal sufficiency entails causal completeness, the existence of (lawfully) necessitating microphysical conditions does not entail causal completeness.

Two (amongst several) recent attempts to solve the exclusion problem – in the case of the problem of mental/physical causation as it arises for non-reductive physicalism – manifest these two ways to proceed. James Woodward (2008), for example, employs an interventionist, counterfactual-based account of causation in order to argue that just the right kind of patterns of counterfactual dependence can exist between an effect *e* and, say, prior mental feature *M* for *M* to count as a *bona fide* and non-overdetermining cause of *e*, even if *M*'s supervenience base, some physical feature *P*,

¹ The 'perhaps' here can, I think, safely be ignored, since the core of Steward's view seems to be that the existence of truly 'necessitating microphysical conditions' really is incompatible with top-down causation.

is itself a sufficient cause of *e*. Christian List and Peter Menzies (2009) take the second approach, claiming that in some special cases there is counterfactual dependence of the right kind at the higher level but *not* at the lower level, so that it is *M*, and *not P*, that causes *e*. Crucially, this claim does not require any ‘gap’ at the lower level; in effect, List and Menzies deny that lawful sufficiency at the lower level suffices for *causal* sufficiency at that level.

Of course, these attempts to solve the exclusion problem as it applies to mental/physical causation do not obviously generalise to cover the full range of cases of ‘top-down’ causation that Steward recognises – and in particular, restricted as they are to mental/physical properties or facts, the attempts do not speak to the causal efficacy of agents. But the cases of ‘top-down’ causation that Steward considers include the behaviour of a wheel (as opposed to the particles that constitute it – pp.233-7) and a whirlpool (as opposed to the individual water molecules that constitute it – pp.240-3), as well as the causation by agents of their bodily movements (pp.243-7). And it seems clear that the kinds of solution to the exclusion problem just mentioned – solutions that require no ‘gaps’ in the lower-level laws – will fairly straightforwardly apply to at least *some* cases of top-down causation. So Steward needs a principled reason to think that such solutions fail.

I suspect that Steward’s reason is simply that she cannot see how lower-level necessitation can possibly be compatible with top-down causation. Thus she says:

... there is absolutely nothing incoherent about the idea that there could be such top-down causation, provided we reject (i) the assumption that there is necessitation at the lower level of each momentary supervenience base by the next and (ii) the idea that a succession, $C_1 \dots C_n$ of momentary supervenience bases for a higher-level entity *M* could do the causal work done by the [higher-level] persisting entity *M* itself. If could not do so, I maintain, because the coincidence that is represented by $C_1 \dots C_n$ cannot be resolved without appeal to forces, principles, or regularities that apply only *because* $C_1 \dots C_n$ constitute *M*. The existence of *M* therefore answers causal questions that the existence of its successive supervenience bases does not. (242-3)

The problem is that it is unclear what the connection is supposed to be between (i) and (ii). According to standard ‘compatibilist’ solutions to the exclusion problem, we can indeed reject (ii). We cannot *fully* causally explain why, say, Jane just took a beer from the fridge without appealing to Jane’s mental states (or perhaps to Jane herself, given a suitable account of agent causation) because appeal merely to the lower-level facts-plus-laws will not exhaust the *causal* facts – since the causal facts include facts about the efficacy of Jane’s mental states (or Jane herself). But why should we also have to reject (i)? The only reason I can think of is that Steward takes ‘necessitation at the lower level’ to amount to causal completeness at the lower level. But, as we have seen, there are reasons to think that we should reject that assumption.

Indeed, I think even if we were to assume that some version of the exclusion problem is still very much in the running, it is unclear why radical indeterminism should be thought to provide the solution, for there are other ways in which we might find a causal role for top-down causation – at least in the case of movers –given Steward’s overall picture of the metaphysics of causation. Grant that we have lawful – indeed, causal – necessitation at the lower level. I take it that on Steward’s view, this would be necessitation at the level of ‘matterers’ or *facts* (213-4); as she says, “[m]uch of what has been written about the relation of causation to such other concepts as law,

counterfactuals, and probability really relates to these mattering causes” (214). So matterers would seem to be the natural home for the notion of lawful necessitation. Now, when it comes to movers (roughly: substance causes), Steward recognises a standard objection: that the object is not ‘really’ doing any causing; rather it is the relevant matterer that is doing all the work. (Thus, according to this objection, “it was not the car per se that demolished the wall, since a car parked touching the wall has no effect and moreover any object of a similar size moving at similar velocity would have done as well as the car” (214).) Steward’s response is that “this is a non-sequitur; ‘it simply does not follow ... that the car did not demolish the wall” (*ibid.*)

Now, we might think of the car case as raising a kind of exclusion problem: surely if the car *itself* is not itself amongst the conditions that necessitated the demolition of the wall, it cannot be a *cause* of the demolition. And Steward’s response, in effect, is to claim that this is not really a case of causal *competition*. The mover (the car) and the matterers (the wall’s being hit with such-and-such a velocity, etc.) are not in competition for causal status, so that the causal sufficiency of the latter does not exclude the causal status of the former. But with this response on the table in the case of the car, it is unclear why Steward can’t make the same move with respect to agents. Grant that there is some set of lower-level *matterers* that are collectively sufficient for the agent’s bodily movement. That in no way undermines the claim of the *agent* to be a bona fide cause as well. Thus no ‘gap’ is required, at the level of matterers, in order to make room for top-down causation, when the causation in question involves causation by a mover (and hence, in particular, by an agent).

It is clear that Steward does not want to embrace this line of thought. Indeed, before her discussion of top-down causation, she says: “It seems to be right that the agent causationist ought to believe that actions do not have prior *necessitating conditions*, and indeed that follows from the conception of actions as settlings of matters by agents that has been developed during the course of this book” (217). But this leaves her in what I think is an uncomfortable position. If what is supposed to justify the claim that necessitating conditions are incompatible with agency is the fact that agents are *settlers* of matters, then – since the notion of settling does not apply in non-agential cases such as whirlpools and wheels – we have no reason to think that necessitating conditions are incompatible with top-down causation *in general*. We have no reason to think that, for example, supervenient matterers can only play a causal role if their supervenience-base matterers do not necessitate, since questions of ‘settling’ are simply not on the agenda here. Nor, as we have just seen, do we have any reason to think that *wheels* and *whirlpools* (*qua* persistent entities and hence *movers*) cannot cause if there are also necessitating matterers in the vicinity. But Steward *does* apparently think that *all* top-down causation requires the absence of lower-level necessitation – and so the appeal to the notion of settling, since it applies only to the case of agency, will not do the job.²

Finally, here is a second kind of worry for the claim that top-down causation excludes lower-level necessitating conditions. The requirement of indeterminism at the lower level in order to make room for top-down causation is one that places very strong

² I have ignored Steward’s claim that top-down causation is required in order to make sense of the fact that only top-down causation can resolve what are, from a lower-level perspective, mere coincidences; see especially 236-9 and 242-3. While this claim forms a central plank in the argument for the need for top-down causation, I do not see any distinctive argument here for the claim that top-down causation requires lower-level indeterminism.

demands on the laws of nature. Even if something like Cartwright's account of laws is correct, it doesn't rule out the possibility that the patchwork of laws might uncover deterministic laws that apply to the very same bit of reality at different levels of description. Perhaps the laws of classical mechanics and thermodynamics, say, only apply in very tightly constrained experimental setups, and only to closed systems. Nonetheless, *in such circumstances* we have something like an exclusion problem to address. If so, our solution to it had better not be to demand that classical mechanics is indeterministic. Granted, there are reasons to think that classical mechanics is not fully deterministic, in that there are certain kinds of specific situation where more than one outcome is compatible with the laws of classical mechanics (see Hoefer 2010, §2.1), but the range of cases is fairly small. A deterministic whirlpool, for example – in some remote corner of the world where there are no animal agents or other potential interveners in the vicinity to interfere with its temporal evolution, so that the whirlpool and its immediate environment constitute a closed system, or near enough – seems to be at the very least a physical possibility, if not an actual phenomenon. And it seems that Steward would need to concede that the causal facts about such a whirlpool are just the same as those for a whirlpool that is susceptible to intervention.

In addition, it is unclear exactly what kind of 'gap' is needed in order for top-down causation to be possible, on Steward's view. Recall that her central claim is that agency requires radical indeterminism, and not merely what I have been calling 'pseudo-determinism'. But her account of top-down causation seems to be entirely compatible with pseudo-determinism: it requires only that lower-level phenomena are not *necessitated* by other lower-level phenomena. (Indeed, the same may be said of 'settling'.) This being so, it seems that additional argument is required in order to establish the incompatibility of agency and DPD, as opposed to the incompatibility of agency and determinism.

5. Radical indeterminism reconsidered

I end on a rather more conciliatory note than I have struck so far. Suppose we start out (as most of us do) by believing that DPD is true, and (as almost all of us do) that agency exists. If we take the former belief to be non-negotiable then, from our perspective, the claim that agency requires radical indeterminism constitutes a sceptical claim: it casts doubt on something we take for granted pretty much all the time, namely the existence of agency. This being so, we will set the bar very high indeed when it comes to assessing any argument in favour of the claim that agency requires radical indeterminism – as we do with sceptical arguments in general (Lewis 1996).

But suppose we grant that Steward is right that our allegedly non-negotiable starting assumption – our belief in DPD – is a mere article of faith that merits no credence on either philosophical or scientific grounds. And suppose, moreover, that one of the grounds for rejecting that article of faith is precisely the recognition that the 'laws', such as they are, are simply not equipped to deal with the interventions of agents into the goings-on in the well-behaved closed systems where the 'laws' have their natural home. Under *this* set of starting assumptions, the claim that agency requires radical indeterminism is not a sceptical hypothesis at all; after all, it doesn't threaten *any* beliefs we're antecedently committed to. And so we can happily set the bar much lower.

From this perspective, then, attempts to argue (as I have been doing) that the case for Steward's brand of incompatibilism is not water-tight are problematic for two reasons. One is that such arguments are somewhat otiose: they are only relevant to the viability of our firmly entrenched belief in the existence of agency if we sign up to an item of philosophical dogma that we really ought to dispense with, namely the truth of DPD. The second, related problem is that they set the bar unreasonably high. Our epistemic standards are set higher when we are confronted with a sceptical hypothesis. But, again, Steward's incompatibilism only counts as a sceptical hypothesis about agency if we cling to our belief in DPD.

The point of all this is really to make clear just how central Steward's commitment to the truth of radical indeterminism is to the argument for her own distinctive brand of incompatibilism. I have long believed that the issue surrounding the correct account of the laws of nature is of crucial importance to the free will debate (see e.g. Beebe and Mele 2002; Beebe 2003), but I had previously only seen the connection in the context of the dispute between Humeans and anti-Humeans about laws. What Steward's enlightening book reveals is that a central shared assumption in most of the debate about the nature of laws – namely the truth of DPD – is one that places constraints on what counts as a viable theory of agency. For those who find Steward's theory attractive, and especially for those who wish to place agent-causation and the ability to do otherwise (in the traditional incompatibilist sense) at the heart of an account of free will, embracing radical indeterminism may deliver the best chance of success. And for those who, for whatever reason, want to show that Steward's account fails, a defence of DPD is needed.

References

- Beebe, H. 2003. "Local Miracle Compatibilism." *Nous* 37: 258-77.
- & Mele, A. R. 2002. "Humean Compatibilism." *Mind* 111: 201-23.
- Cartwright, N. 1999. *The Dappled World*. Cambridge: Cambridge University Press.
- Hoefer, C. 2010. "Causal Determinism." In *The Stanford Encyclopedia of Philosophy* (Spring 2010 Edition), edited by Edward N. Zalta. URL = <http://plato.stanford.edu/archives/spr2010/entries/determinism-causal/>.
- Kim, J. 1998. *Mind in a Physical World: An Essay on the Mind-Body Problem and Mental Causation*. Cambridge, MA: Bradford.
- Lewis, D. K. 1996. "Elusive Knowledge." *Australasian Journal of Philosophy* 74: 549-67.
- List, C. and Menzies, P. 2009. "Nonreductive Physicalism and the Limits of the Exclusion Principle." *Journal of Philosophy* 106: 475-502.
- Steward, H. 1997. *The Ontology of Mind*. Oxford: Oxford University Press.
- 2012. *A Metaphysics for Freedom*. Oxford: Oxford University Press.
- Woodward, J. 2008. "Mental Causation and Neural Mechanisms." In *Being Reduced: New Essays on Reduction, Explanation, and Causation*, edited by Jakob Hohwy and Jesper Kallestrup. Oxford: Oxford University Press, 218-62.