**Agency and Incompatibilism**

Randolph Clarke

Helen Steward’s *A Metaphysics for Freedom* argues for what she calls Agency Incompatibilism: agency itself is incompatible with determinism. The agency in question includes such things as dancing, making a decision, and absent-mindedly scratching one’s head; it includes the activities of infant and adult human beings as well as those of cows and spiders (and perhaps earthworms). On Steward’s view, it is metaphysically impossible for there to be agents of any kind in a deterministic world. Indeed, it is impossible for any exercise of agency to be determined by what precedes it, and necessarily any instance of activity evolves indeterministically from start to finish. Free will requires indeterminism because it requires agency, which requires indeterminism.

I’ll focus here on Steward’s main argument for this view. It seems to me to suffer from a serious shortcoming.

1. *The Settling Argument*

An agent, Steward maintains, is a self-mover, a being with a power to make its body, or parts of its body, move. The making-one’s-body-move that we have when this power is exercised is a matter of causation by the agent, and an action is a causing by the agent of some movement of its body, or some part thereof.[[1]](#footnote-1) When I raise my arm, I cause a movement of my arm, namely, its rising. When I make a decision, I cause some movement within me, presumably somewhere within my brain.

At one point, Steward suggests that once we’re clear that actions are movings rather than movements—causings by agents rather than movements caused by antecedents—the tension with determinism becomes “readily palpable” (35).[[2]](#footnote-2) But the notion of a moving, or of a causing of a movement, doesn’t itself generate the tension. When the wind ripples the surface of a pond, we have a causing of a movement. But I see no reason to think—and I don’t see that Steward offers any reason to think—that a moving of this kind can’t be determined by what precedes it.

Now, Steward takes the causation by an agent that we have in any instance of action to be irreducible substance causation. And one might think that it is with substance causation that she takes the conflict with determinism to arise. But this isn’t so. Steward takes irreducible substance causation to be ubiquitous: ions, molecules, stones, and masses of air and water are all causes. And nowhere does she suggest that there can’t be causation by these things in a deterministic world.

Rather, Steward’s principal argument for Agency Incompatibilism stems from her claim that we have the self-movement of agency only when some matters are *up to* the agent when she acts (24). And for some matter to be up to an agent when she acts, Steward maintains, it must be the case the agent *settles* some matter then.[[3]](#footnote-3) Further, one can settle some matter at a time only if that matter isn’t already settled (26, 39, and 40). But if determinism were true, every matter would have been settled from the beginning of time (or forever, if there’s no beginning). Hence, there can be agency only if determinism is false.[[4]](#footnote-4)

Steward takes this line of thought—I’ll call it the Settling Argument—to explicate our ordinary concept of agency, and thereby to show what must be the case given that, as we firmly believe, there are beings who are agents, and there are exercises of agency. Our concept of agency, she holds, includes the idea of matters being up to an agent, and the idea of an agent’s settling some matters when it acts. And it is these ideas that generate the conflict with determinism.

1. *Steward on Settling*

I’m happy to count cows and spiders as agents. As to whether matters are up to a cow or a spider when it acts, I’m not quite sure what to say. But let’s suppose that in some context one might truly say such a thing. What thought would one thereby express?

As Steward sees it, the things that can be up to an agent aren’t particular events. Rather, what’s up to me are “the answers to a whole range of questions that are settled by my action when I act” (37). The questions concern *whether*, *when*, *where*, or *how* one will *A* (where *A*-ing is some act type). Since when, where, or how one will *A* can be expressed as whether one will *A* then, there, or that way, it seems that the questions can all be expressed in terms of *whether* something or other.

Whenever an agent acts, then, some question is up to that agent, and the agent settles some question then—or, as Steward sometimes puts it, the agent’s action settles some question then.

What is it for an action to settle a question when one acts? The core idea, Steward says, “is that of a question that is capable of being resolved in different ways at all times up until a certain moment—the moment of settling—at which point something that happens causes it to become resolved in one particular way” (39).[[5]](#footnote-5) Further, “we can think of each matter that is subject to settling as expressible by means of a question [of] the form [*p* v ¬ *p*], whose answer is (metaphorically) decided at a certain point in time, when it becomes impossible either that *p* should obtain or that ¬ *p* should” (39). (She sometimes expresses the matters that can be settled as having the simpler form *whether p*; I’ll do the same.)

If we want to put the point non-metaphorically, how might we spell out this idea of an action’s settling some matter when one acts? I don’t see that Steward is very explicit on this point, but perhaps she would accept something like the following:

(S) An action *a* that is performed at time *t* settles at *t* whether *p* iff (i) either it is impossible that *a* be performed then and the actual laws of nature hold and *p*, or it is impossible that *a* be performed then and the actual laws hold and not-*p*, and (ii) there is nothing existing at any time *t’* prior to *t* such that either it is impossible that that thing exist at *t’* and the actual laws hold and *p*, or it is impossible that that thing exist at *t’* and the actual laws hold and not-*p*.[[6]](#footnote-6)

As it might be put, an action performed at time *t* settles at *t* whether *p* just in case that action either suffices for its being the case that *p* or suffices for its being the case that not-*p*, and nothing prior to *t* suffices for either of these.

If this is on the right track, then when I raise my arm at a certain time, *t*, my action of raising my arm might settle at *t* whether I raise my arm then, provided that nothing prior to that time suffices for its being the case that I raise my arm then. And given that my action of raising my arm settles at *t* whether I raise my arm then, it might be said that *I* settle at *t* whether I raise my arm then, and that I settle this matter at *t* by raising my arm then.

The conflict with determinism arises if we consider matters concerning any moment in history (except a first moment, if there is one). If determinism is true, then for any such moment, something prior to that moment suffices for all matters concerning what happens then and thereafter. No action, then, can settle when it occurs any matters concerning what happens then or later. No agent, then, can ever settle any matter when she acts. Nothing, then, can be up to any agent. There can’t be any agency. This, I take it, is the gist of the Settling Argument.

I grant that ‘up to you’ and ‘settle whether’ can be used to express ideas like the ones that Steward spells out in advancing this argument. And I’ll accept for the sake of discussion that some idea of matters being up to agents and some idea of agents settling matters when they act are implicated in our concept of agency. Still, unless it can be shown that the ideas tied up with the concept of agency are the ideas that Steward has spelled out—or ideas sufficiently similar to them—we don’t have the support that is needed for Agency Incompatibilism.

1. *‘Settle at a Time’ and ‘Up to You’*

Steward acknowledges that ‘settle whether’ can be used to express an idea *different* from the one to which the Settling Argument appeals. She allows that “it might be reasonable to speak of some person or event’s having settled a matter which was nevertheless determined to occur well in advance of that person’s intervention or that event’s occurrence” (41). For example, if four dominoes are lined up and the first one is knocked over, the fall of the third might settle whether the fourth one falls, even if it was determined, once the first one was knocked over, that the fourth one would fall. But what we *can’t* have in this case, Steward insists, is the fall of the third domino *settling some matter at a time*. “An event can only settle a matter at the time at which it occurs, if that matter is not already settled before that time” (40). And, she adds, “it is this feature of the concept of settling on which my argument against universal determinism depends” (40).

Is it true that an *agent* can’t settle (in some way) at a time some matter that has already been settled (in some way)? Consider: I set out to determine whether Oswald acted alone in killing Kennedy. I read various things, talk with knowledgeable folks, and then make up my mind. When I do the last of these things, of course I don’t do anything that suffices either for its being the case that Oswald acted alone or for its being the case that he didn’t. But I *do* do something that suffices for its being the case that I believe a certain thing—I come to believe that thing. I settle whether I believe that Oswald acted alone, and I settle this matter when I make up my mind, by making it up.

Saying this, I submit, is saying something true, and it is using ‘settle whether’ in a perfectly ordinary way. Indeed, it might be fair to say that it’s up to me whether I come to believe one thing or another on this matter. As does every rational being, I have a kind of freedom of belief, and I might exercise that freedom on this occasion.[[7]](#footnote-7)

Do these observations give us reason to think that making up one’s mind in this way can’t be determined by anything that precedes it? Of course, as Steward sees things, it can’t be that any of my *actions* are determined by antecedents. But what about my coming to hold a certain belief on the question at hand? I settle whether I do, and I settle this matter at a certain time. If, for this to be so, the event of my coming to hold this belief must be undetermined by what precedes it, then the Settling Argument supports a considerably stronger conclusion than Steward draws in the book. But if, as it seems to me, we needn’t have indeterminism here, then we have a use of ‘settle at a time’ that doesn’t express an idea in conflict with determinism. And then, even if some notion of settling matters when one acts is wrapped up on our concept of agency, this fact doesn’t suffice to establish Agency Incompatibilism.

There need be no conflict in the idea of one’s settling at a time something that has already been settled if in fact different kinds of settling are involved. When I come to have the belief, I settle some matter at that time by making up my mind. My mind wasn’t made up on this matter at any previous time. The matter was not previously settled in this way. But if determinism is true, then from the beginning of the world there have been conditions that sufficed for its being the case that at this moment I come to hold this belief. The matter has forever been settled in this second way. But we have here two different kinds of settling. There is no contradiction in saying that what was forever settled in one of these ways came only at this moment to be settled by me in the other way.

There is also a use of ‘settle at a time’ expressing something that doesn’t imply that settled matters must *remain* settled. (The point here doesn’t directly challenge the Settling Argument, but it is worth noting nonetheless, for it further indicates the variety of things that can be expressed using this phrase.) I might settle now the question whether I’ll attend a conference next year, even if the question can be reopened should I receive new information. I settle the question when I make a firm decision; but firm decisions are revocable.[[8]](#footnote-8)

There are, as well, uses of ‘up to you’ that don’t express what Steward has characterized. Generally when my wife and I take a walk through the neighborhood, by mutual consent it’s up to her whether we take this route or that one. Now, *we* don’t take a certain route unless she and I *both* do; and we don’t *both* take that route unless *I* do. Of course, even when it’s up to her whether we take a given route, I remain free not to go along. Perhaps there’s nothing that she can do that will suffice for my taking that route. Still, by mutual consent, the matter is up to her.

Sometimes a graduate student comes to my office to ask whether she might arrange her studies in this way or that. If the Graduate Handbook neither requires nor forbids the things in question, I might respond, “It’s up to you.” Sometimes someone wants you to make a decision for her; she doesn’t want the bother or the responsibility of making it herself. You might tell her that you can’t or won’t; “It’s up to you,” you might say. It’s far from clear, and I think rather implausible, that these uses of ‘up to you’ express a commitment to indeterminism.

‘Settle at a time’ and ‘up to you’ express modal notions. But which modal notions do they express? My conjecture is that we use them to express a variety of different ideas, many of them presenting no conflict with determinism, even if some of them do.

Suppose, then, that it is part of our concept of agency that whenever an agent acts, certain matters are up to her, and she settles certain matters when she acts, by acting. More will need to be said to show that there is anything here that is incompatible with determinism.

1. *Power, Ability, and Dependence*

Many uses of ‘up to you’ express something concerning power, ability, and dependence. In saying that some matter *p* is up to you, I might say that you have certain cognitive or agential powers that you’re able to exercise, and able not to exercise, and whether *p* depends on whether you exercise them. (When the fact that *p* is the fact that you exercise those powers, the dependence is a simple matter. Sometimes the dependence might obtain only given certain further facts, such as others’ commitments to go along with what you decide.)

The observation here falls short of what is needed to establish Agency Incompatibilism. For there are different kinds of thing that we can attribute in attributing powers and abilities—there are different kinds of power and ability. Some are such that we can possess them unexercised even on occasions when prior things determine that they remain unexercised.

It seems to me that Steward agrees with me on this point. For she agrees with compatibilists about what we express when we say certain things about agents’ powers. For example, Steward says, often in saying that a certain agent could have decided otherwise, we simply say that no one forced the agent to decide as she did, or “that [she] possessed a general decision-making capacity at the time that was in perfect working order and which [she] could then have exercised so as to have made the opposite decision, if [she] had seen good reason to do so” (140). All this can be so, Steward insists, even if it isn’t compossible with the prior course of events and the laws of nature that the agent have made the opposite decision.

But now, similar things can be said with respect to ‘power to refrain’, ‘open alternative’, ‘spontaneity’, and other expressions that Steward insists must be applicable whenever any agent acts. And similar things can be said about ‘up to you’ and ‘settle at a time’.

1. *Separating Issues*

Steward does devote much of a chapter to countering an objection to the Argument from Settling. However, the objection that she considers combines two claims: that settling some matter at the time one acts need not involve indeterminism, and that settling some matter at a time can be analyzed in terms of causation by mental states or events, as a proponent of the Causal Theory of Action might contend. Her responses are responses to the latter claim. But the former claim is independent of the latter.

Determinism says nothing about what causes are or what actions are.[[9]](#footnote-9) It doesn’t say that only events or states are causes. It doesn’t expressly rule out the view that animals cause movements of their bodies (or parts thereof) when they act. Thus, one might challenge Agency Incompatibilism without appealing to the Causal Theory of Action, and indeed without rejecting the view that actions are agent causings. Steward’s objections to the Causal Theory of Action, then, fail to meet the challenge that I’ve raised, for they aren’t addressed to that challenge.

1. *Further Argument?*

I’ve examined what Steward presents in the book as her main argument for Agency Incompatibilism.[[10]](#footnote-10) At the meeting at which these comments were presented,[[11]](#footnote-11) Steward replied that further argument later in the book addresses the problem I’ve raised. But she didn’t on that occasion say what the further argument is, and I’ve been unable to identify it. I look forward to learning more about the matter from her response here

1. The movement that is caused when an action is performed is what Hornsby (*Actions* [London: Routledge & Kegan Paul, 1980], p. 2) calls a movementI. It is, for example, the motion of my arm, rather than my act of moving my arm. [↑](#footnote-ref-1)
2. All page references are to Steward’s *A Metaphysics of Freedom* (Oxford: Oxford University Press, 2012). [↑](#footnote-ref-2)
3. At one point, Steward says that “in order that some question or matter be up to me…what I have to be able to do…is to *settle* that matter” (26). However, she also says, “When an agent acts…various matters are thereby *settled* by that agent at the time of action” (31). The argument appears to need this stronger claim, on which it isn’t just the *ability* to settle matters but the *actual settling* of matters that is required. [↑](#footnote-ref-3)
4. The argument that I’ve summarized here is presented in chapter two of Steward’s book, with the heart of it appearing in the sections titled “The argument: agency is inconsistent with universal determinism” and “Settling.” [↑](#footnote-ref-4)
5. Suppose that what happens is that an event of type *E* occurs at *t*. If the question that is resolved by the occurrence of this event is whether an event of type *E* occurs at *t*, it seems odd to say that the event *causes* the resolution of this question. In the interpretation of Steward that I proceed to offer, I don’t put any weight on her use of ‘causes’ here. [↑](#footnote-ref-5)
6. Steward says that undetermined events, such as the spontaneous emission of a particle by an unstable atom, can settle matters at the time of their occurrence, “both matters whose settling is *constituted* by the occurrence of those very events…or matters which are deterministically connected to the occurrence of those indeterministic events” (40). I don’t see that she ever says explicitly what it is for the settling of some matter to be constituted by the occurrence of some event. In S I’ve suggested what I think she might mean by this, for the case of action.

   Perhaps, in order to capture Steward’s idea, S would need the further proviso that nothing at *t* that is distinct from the action in question settles whether *p*. Since nothing in what I say hinges on this issue, I’ll set it aside. [↑](#footnote-ref-6)
7. It is for reasons such as these that some philosophers take forming a belief in this way to be an exercise of a kind of agency, even if it isn’t something that one does intentionally or purposively. I don’t myself claim such a thing, but I also don’t think that the issue is one of great substance. [↑](#footnote-ref-7)
8. One might insist that in this case, when I make the decision, I don’t settle *whether I’ll go* but only *whether I’ll intend to go*. However, if after I’ve decided not to go someone tries to persuade me to go, I might cut them off by saying, “I’ve already settled the matter.” [↑](#footnote-ref-8)
9. It is common to state the thesis in terms of entailments, but we can get an ontic thesis as follows. Let *W* be the set of all possible worlds with the same laws of nature as our world. The thesis that our world is deterministic is the claim that any world in *W* that is exactly like our world at any moment in time is exactly like our world at every moment in time. [↑](#footnote-ref-9)
10. She says at several points in chapter two (e.g., 26 and 31) that her argument for Agency Incompatibilism is located in the final sections of that chapter, which are the ones on which I’ve focused. [↑](#footnote-ref-10)
11. The 2014 Pacific Division Meeting of the American Philosophical Association. [↑](#footnote-ref-11)