

## Reconsidering Scientific Threats to Free Will

Manuel Vargas

University of San Francisco

In “Free Will and Substance Dualism: The Real Scientific Threat to Free Will?” Al Mele extends his groundbreaking work on scientific arguments against free will.<sup>1</sup> He replies to charges that he has missed the real threat to free will posed by experimental work, and he focuses on two issues: (1) the claim that the “real” threat of scientific work is bound up with substance dualism and (2) recent work by Soon et al. that has been taken to show that some intentions can be predicted in advance.

I will outline some points of agreement before arguing that Mele still hasn’t yet accurately diagnosed the real threat to free will posed by experimental work. I will then contend that getting the diagnosis right may be of little help stopping the underlying threat.

### Agreements

I *firmly* agree with almost everything Mele claims in “Free Will and Substance Dualism.” He is clearly right that scientific objections to free will bound up in substance dualism aren’t compelling. Nearly all serious accounts of free will make no appeal to substance dualism. Moreover, when one considers the empirical evidence about folk beliefs concerning the requirements of free will, the evidence seems to show that on ordinary usage, ‘free will’ does not require substance dualism. So, game, set, and match against those who think that science has shown we lack free will because of difficulties with substance dualism.<sup>2</sup>

Mele’s remarks on the work of Soon et al (2006), regarding the prediction of intentions are also compelling. Inclinations do not constitute intentions,. If one wants to claim that something is an empirical

---

<sup>1</sup> Alfred Mele, “Free Will and Substance Dualism: The Real Scientific Threat to Free Will?,” in *Moral Psychology, Volume 4: Free Will and Moral Responsibility*, ed. Walter Sinnott-Armstrong (Cambridge, MA: MIT Press, forthcoming). For background, see his Alfred R. Mele, *Effective Intentions: The Power of the Conscious Will* (New York: Oxford University Press, 2009).

<sup>2</sup> I should note that Mele is unfailingly generous with his interlocutors—certainly more so than I would have been. Borrowing a metaphor from Daniel Dennett, I might have said that objecting to free will on the basis of the falsity of substance dualism is like objecting to the existence of love because Cupid doesn’t exist. Daniel Dennett, *Freedom Evolves* (New York: Viking, 2003).

threat to the existence of free will, then it behooves that claimant to specify a conception of free will (whether folk or specialist) under which free will is threatened. Although Mele doesn't put it this way, we might say that claiming that some scientific result shows the falsity of "free will" under some construal accepted by neither the folk nor specialist is either bad scholarship or academic hucksterism.

In what remains, I'll focus on whether Mele's remarks genuinely answer my charge that he has missed the real threat of scientific work. Although I'm obviously sympathetic to a good deal of what he says, I'm not convinced that he has gotten right the diagnosis of the threat. I'm also unpersuaded that doing so would defeat going scientific arguments.

### Disagreement?

Mele notes that I am among those who have suggested that his book *Effective Intentions: The Power of Conscious Will* (2009) may be "missing the real threat to free will" posed by his targets (i.e., psychologists, neuroscientists, and various others claiming that experimental evidence disproves free will).<sup>3</sup> What I speculated was that some of Mele's targets were operating with a substantive view of free will that he had missed. As such, his targets could reply that he had misdiagnosed why they thought science impugned free will.

Mele and I likely agree that the kind of view I attribute to his interlocutors requires a *lot* of heavy philosophical lifting. Most philosophers who spend significant time thinking about these things tend to conclude that such heavy lifting is not worth doing. Moreover, as I noted in the original review, I take it that Mele has done a good deal to defend his preferred view(s) on free will, in contrast to his targets. What we disagree about is whether he has gotten their view right, and thus, whether he has addressed the threat I attributed to some of his interlocutors.

In my review, I speculated that

[Mele's targets] are sometimes motivated by what the philosophical literature labels as "source" intuitions—the idea that for us to act with a free will we must be the ultimate origins of strands of the causal nexus. On one way of putting things, source theorists think that free acts cannot have causal antecedents that extend

---

<sup>3</sup> Alfred Mele, "Another Scientific Threat to Free Will?," *The Monist* 95, no. 3 (2012): 423-41; Alfred Mele, "Free Will and Substance Dualism: The Real Scientific Threat to Free Will?"

back in time prior to the decisions of the agent or the agent's free formulation of the relevant characterological inputs to that decision. I do not mean to defend the view. However, I take it that this view is the engine that drives many [of Mele's targets].<sup>4</sup>

Mele takes these remarks, and other passages, to indicate that I was construing his targets as attacking free will on the basis of supposing it required substance dualism. While I think that this is a plausible interpretation of at least some of his interlocutors, and while it is consistent with some of my remarks, I had in mind a different target.

By emphasizing “source” intuitions, I meant to direct our attention to the view *on which any adequate conception of free will must be one on which we are the ultimate origins of what we do*. Although it is not always clear what such ultimacy entails, there are a variety of ways one might try to make good on the general idea. Substance dualism is one way, albeit a way usually regarded as more trouble than it is worth.<sup>5</sup> On that picture, if our intentions are suitably connected to our being thinking or spiritual substances, then it creates the possibility that intending isn't entirely governed by whatever causes constrain physical systems. So, on this picture, the promise of ultimacy is afforded by thought that our action has sources disconnected from the physical world.

My main point, however, was not that scientific free will deniers were (usually implicitly) helping themselves to the idea that free will required dualism. Rather, the idea was that scientific free will deniers were helping themselves to the idea that free will required “sourcehood,” and that their evidence impugned *that*. Objecting to the idea that scientists are implicitly helping themselves to a picture of free will which requires substance dualism is to miss the core worry, for a scientist need not think free will requires dualism for the sourcehood worry to get going.

One way to see why science might threaten sourcehood, quite apart from dualism worries, arises from reflecting on constitution, reduction, and emergence in the mind. Here's a characterization of how sci-

---

<sup>4</sup> Manuel Vargas, “Review of *Effective Intentions*,” *Notre Dame Philosophical Reviews* (2009): <http://ndpr.nd.edu/review.cfm?id=I7385>.

<sup>5</sup> Moreover it isn't obvious that it solves the basic problem. We would still need some story about how the thinking substance operates to ensure that it doesn't collapse into a different sort of causal picture.

entific results show that our sourcehood is threatened via reductionism worries, quite apart from worries about substance dualism:

What science inevitably shows is that we are ultimately constituted by the same stuff as everything else—i.e., physical stuff. However, the only physical powers we can have are those afforded to us by the physical stuff that constitutes us. So, if that is right, then when science shows us to be physically constituted, it shows us to be limited to those physical powers of our low-level constituents. However, that stuff is all just part of the inexorable causal nexus. There is simply nothing left to do the work of making us special, or ultimate causes of what we do. So, there is no free will.

It was this way of thinking about sourcehood worries that led me to claim in my review of Mele’s book that “it is the broadly reductionistic element that generates the core of the [scientific no free will] worry.”

On this interpretation of scientific free will skeptics, they understand their data to show that there are no good hooks on which to hang our sourcehood requirement for free will. Neither appeals to non-reducible constituents (e.g., substance dualism) nor appeals to emergent powers look plausible on this approach. If decisions, intentions, and willing turn out to be entirely explicable in terms of lower level neurological functions, then this counts against there being a suitable emergent substance or related causal property that can underpin the sourcehood demand. So, free will loses, dualism or not.<sup>6</sup>

My point was never to defend the sourcehood impulse or the reductionist threat.<sup>7</sup> Rather, my point was (and is) to suggest that a number of scientists seem to be motivated by a conception of free will that appeals to a sourcehood requirement. Consequently, a helpful way to construe their doubts about free will is in terms of scientific data threatening the satisfiability of *that* requirement. On the account I have offered,

---

<sup>6</sup> It was for this sort of reason that I concluded in my review of Mele’s book that “it is the broadly reductionistic element that generates the core of the [scientific no free will] worry.” Eddy Nahmias and his collaborators have done the most to clearly articulate this sort of concern in the contemporary free will debate. See Eddy Nahmias, et al., “Free Will, Moral Responsibility, and Mechanism: Experiments on Folk Intuitions,” *Midwest Studies in Philosophy* XXXI (2007): 214-41.

<sup>7</sup> To be sure, there are lots of ways to get off this train. One might reject the idea that scientific evidence shows that we are entirely constituted by low-level physical entities. One might reject the idea that scientific evidence shows anything interesting about emergent causal powers. One might reject the idea that the only way to make good on the sourcehood demand is via substance dualism or emergent causes (or emergent substances with corresponding causal powers). One might reject reductionism on some entirely distinct set of grounds.

substance dualism is merely a symptom of a more basic demand. It is the diagnosis of that demand, and a reply to it, that has been missed in Mele's substance dualism-focused discussion.

This may sound like a big objection against the adequacy of Mele's account. I don't think it is. A variety of philosophical accounts have rejected the ultimacy demand or cashed it out in ways that don't require abandonment of more basic metaphysical commitment (such as constitution exclusively by the physical, and causal closure of the physical). In other contexts, Mele has said a fair amount about arguments for and against prominent ways of understanding the sourcehood requirement.<sup>8</sup> Mele's targets have, so far as I am aware, made no serious attempt to engage with these discussion. Indeed, Mele's got a well-motivated and intricately worked out positive theory (theories, actually) of free will. His targets don't. In short, he isn't one showing up to a gunfight with only a knife.

The nub of our disagreement, such as it is, concerns the possibility of defeating scientific arguments against free will without appealing to a substantive conception of free will. Mele has said that his aspiration in *Effective Intentions* was to address scientific arguments without appeal to a substantive theory of free will (ms I).<sup>9</sup> While we can presumably make some important progress without talking about the operative notions of free will, I'm inclined to think that approach will usually end up being unsatisfying. Whether some bit of evidence counts for or against us thinking free will exists, depends on what we think free will *is*. If the shape of the present discussion is right, we cannot hope to dissuade scientists from free will skepticism without engaging them in a substantive discussion about the operative notion of free will.

I don't know whether Mele disagrees. His subsequent willingness (in this volume and elsewhere) to consider the implicit convictions of his targets suggest he might not. For my part, though, I suspect that so long as scientists feel no compulsion to defend their selection of a given conception of free will, they will continue to feel licensed to declare that it doesn't exist.

### **Incurable**

I conclude by arguing that scientific threats to free will, quite apart from their philosophical implications,

---

<sup>8</sup> Alfred Mele, *Free Will and Luck* (Oxford: New York, 2006).

<sup>9</sup> See also Alfred Mele, "Another Scientific Threat to Free Will?", pp. 425-26.

may be ineradicable for the foreseeable future.

Suppose that for most people, the bits of their psychic economy that support assenting, accepting, avowing, reasoning, and objecting turn out to be a bundle of complicated dispositions. I might now be prepared to assert that she loves me, and later conclude that she loves me not, with no new evidence. Or, the evidence might be bad, or below my conscious awareness, or at best loosely connected to what an ideal reasoner would rely upon. These effects are not always just the consequences of uncertainty. I might ordinarily be an ardent atheist but find myself a fervent a believer when the bullets are flying. If I'm a theist in a fox-hole and an atheist in church, who will complain—unless I tell them?

Call those belief-like dispositions to assent, accept, object, and so on “elements of the doxastic economy.” Such elements may be used as the basis of reasoning and attributing commitments to ourselves and others. The commitments—metaphysical, theoretical, practical—of the doxastic economy are “mottled,” i.e., patchy, ill-behaved, and as a matter of psychological and social fact variably subject to pressures for coherence and consistency. For example, I might assent that I like hoppy beers when hanging out with craft beer snobs on a Friday night, and order accordingly. I might (unknowingly) never use those considerations when reasoning about what beer to buy at the grocery store. When asked about California-style IPAs, I might even complain about their excessive hoppiness.

Am I inconsistent? Perhaps. What do I really believe about hoppy beer? I have no idea. If we must say something, perhaps it is best to say I believe in the goodness of hops on Friday nights with beer snobs, that I am agnostic at the grocery store on Saturday, and somewhat unimpressed with their virtues on Sunday. Unless one has an arguably unhealthy interest in intellectual hygiene (or beer preferences), we are typically willing to tolerate some amount of inconsistency and incoherence in ourselves and others.<sup>10</sup>

In saying all of this, my aspiration isn't to plump for one or another account of what beliefs “really” are. I only wish to assert the mottled character of belief-like dispositions, and as such, that they manifest in one way here and in another way there. Moreover, such varied manifestation is partly an effect of other beliefs and motivations. If I'm feeling combative or dissolute, and someone who irritates me wants to dis-

---

<sup>10</sup> There can be circumstances where tolerance of incoherence or inconsistency is very costly. In such circumstances, we may well tend to be more scrupulous about the hygiene of the relevant bits of our doxastic economies

miss the virtues of Cascade hops, then I might find myself enthusiastically defending hoppiness. Such an event might go on to rigidify my general disposition to acclaim the virtues of hoppy beers, and perhaps even come to structure my grocery shopping dispositions. The idea here, then, is that our “beliefs” are often a mess of tendencies, entwined with other aspects of the psychic economy.

Turn to the doxastic elements concerned with free will. As elsewhere, we should expect to find a mess. Consider those commitments concerning determinism, dualism, reductionism, and so on. Presumably, many of us are more and less disposed to accept, avow, and use as the basis for deliberation the content described by such theories. The usual complexity applies: Manifestations are varied across variegated contexts, depending on all sorts of factors.

Now consider the aspects of our moral psychology relevant to responsibility. Suppose we find a default disposition in many contexts to praise and blame harm-causing agents—unless we get some reason not to. Such reasons not to praise and blame may vary, depending how strongly we feel about the particulars, whether others are praising and blaming, whether we can get some general principle up and running that licenses praise and blame in the situation. And so on. Throughout, motivated reasoning will intertwine with variable but frequently weak rational pressures.

Here, something interesting emerges. Different ways of presenting and framing the free will problem will activate different aspects of the doxastic economy in different contexts. If I frame free will in terms of a power that makes sense of moral responsibility, this will likely pull in those elements tied to the psychology of moralized praise and blame. If I frame free will in terms of alternative possibilities, the truth of my beliefs about my powers under practical deliberation, or the power to be the ultimate source of my action, then each of these ways of framing the issue will active a distinct (if sometimes overlapping) network of doxastic dispositions. How these things reflect other commitments to dualism, physicalism, determinism and so on is likely to be complicated. For that matter, the extent to which I find dualism, physicalism, and determinism plausible may well be affected by those other elements, too. The point is that these other framings are lurking there, present but submerged.

The general point is just that we have a messy, ill-behaved, imperfectly rational, and inconsistent web of thinking that underpins any “convictions” about free will. Some of the doxastic elements will be, on reflection, bizarre. They might have implications we would never, in moments of calm, be inclined to ac-

cept. In the ordinary course of things, we are almost never *completely* committed to anything. Bizarre beliefs can live in the metaphysical foxholes of our lives, appearing when we are tired, sloppy, or have some incentive to think magically.

There are two upshots relevant here.

First, it shouldn't be surprising that when we conduct surveys we can find evidence that many people think there is free will and moral responsibility in the vignettes we test. It might be that a probe will find espousal of some commitment that doesn't run very deep, or that wouldn't show up in action, or that doesn't serve as a premise for reasoning or argument, but nevertheless shows up in the context of reactions to vignettes in a consistent way. So, if you hit me with a hoppy beer vignette, maybe I'll never vary from being willing to espouse the virtues of hops. If you take me to a bar or a grocery store, a different pattern might emerge.

This is old news to everyone conducting experiments. Still, perhaps it cuts some ice on the matter of discerning abstract philosophical commitments of the folk, such as whether free will requires dualism. On the present picture of our doxastic economy, maybe we do have some mild commitments that favor substance dualism as a requirement for free will (or agent causation, or robust alternative possibilities, or what have you). This would be compatible with what one finds in standard surveys. Depending on the question asked, such elements might be trumped or overwhelmed by other (even countervailing) doxastic dispositions or inferences following from them, at least in survey contexts.

This possibility suggests a second upshot, why scientific results may persistently threaten free will, no matter how poor the presumed metaphysics of free will. On one way of casting the matter, theories are privileged rigidifications of the doxastic economy.<sup>11</sup> A good philosophical and scientific account expunges ambiguity, accounts for countervailing reactions, and clarifies substantive commitments. Importantly, philosophical and scientific theories are typically subject to comparatively strong pressures for coherence and consistency. If my theory holds that dualism is both true and not true, it will be sure to be met with complaint. If you hold that consciousness is explicable in terms of brain states on Fridays, but not on Satur-

---

<sup>11</sup> This idea is consistent with the view, apparently widely shared among philosophers and scientists, that our theories of free will are intimately tied to the elements of our doxastic economies that are about free will. A theory that purports to be about free will, but that has no relationship to our doxastic dispositions about free will, would not be a theory of free will.

days, you are unlikely to get funding.

To the extent to which experiments and theories conflict with naive strands of our conceptual economy, science will remind us that parts of our self-conception are in error. Such pressures will be hard to put a finger on, especially when those belief-like dispositions don't rise to the level of a declared commitment or an easily accessible avowal. In such cases, the philosopher will rightly insist that no credible *theory* of free will is committed to such nonsense. Still, the scientist may well be right to insist that there has been a disturbance in the Force. Some aspect of how we think about free will—even if only a loopy, ill-conceived, largely unarticulated bit—may indeed be threatened. If so, then we can expect scientific pronouncements about the death of free will to continue until such point as the bizarre bits of our doxastic economy of free will have ceased to be widespread features of our thinking. On this view, standard philosophical diagnoses of scientific threats might amount to little more than diagnoses of mostly incurable diseases.