If Free Will Doesn't Exist, Neither Does Water

Manuel Vargas
University of San Francisco (Philosophy & Law)

Forthcoming in G. Caruso, ed. *Exploring the Illusion of Free Will and Determinism*

There are now many thoughtful people who insist that science has shown us that there is no free will. In one way, this is unsurprising. The march of science has a long history of overturning a variety of beliefs that have seemed obvious, even vital to those who have had them. In another way, though, the triumphalism among free will skeptics is surprising. As much as science has a track record of overthrowing beliefs, it also has a distinguished history of changing the details of what we believe without abandoning the general outline. Skepticism is not always the result when science shows us something troubling about how we understand the world.

Consider that at one point many educated people thought that water was one of the four basic indivisible substances of the world, that race was a strict biological kind, that marriage was a property exchange between two men, and that the normative force of laws depended entirely on the God-sanctioned powers of the sovereign. Most of us now have little trouble accepting the chemical theory of water, that race is mainly a social kind, that women aren't property to be exchanged between men, and that law is not just the divinely sanctioned threat of a sovereign. Some of these transitions went easier than others, and the role of experimental science played larger or smaller roles in each of these cases. However, it is also clear that anyone who would now deny the reality of water, race, marriage, or the law is going to have to explain themselves.

It is not obvious that our current notion of free will is an immutable feature of our conceptual framework. But many scientists seem to think that it is, and that the operative conception of free will was fixed by commitments that no physicalist should accept. For example, neuroscientist P. Read Montague (2008) asserts that:

> Free will is the idea that we make choices and have thoughts independent of anything remotely resembling a physical process [. . .] From this perspective, your choices are not caused by physical events, but instead emerge wholly formed from somewhere indescribable and outside the purview of physical descriptions. This implies that free will cannot have evolved by natural selection, as that would place it directly in a stream of causally connected physical events. Consequently, the idea of free will is not even in principle within reach of scientific description (R584).

In a similar vein, John Bargh and co-author Brian Earp (2009) have maintained that:

> Free will may be defined as an agent's ability to act on the world by its own volition, independently of purely physical (as opposed to metaphysical) causes and prior states of the world.

Most scientists who are "no free will" enthusiasts fail to consider whether free will might be something—like water, marriage, race, law, and so on—about which we change our views regarding its fundamental nature. To them, elimination of the concept of free will seems to be the only live option, and the basis for the elimination is (in part) that they think free will must be as they happen to conceive of it.

However, skepticism about free will has to be *earned,* just as it does in the case of eliminativism about water, marriage, race, and laws. The free will skeptic has to, at the very least, show that free will skepticism is a more plausible view than its alternatives. Crucially, when one of those alternatives is that we are mistaken about what free will entails, skepticism can't be demonstrated simply by observing that there is something problematic about free will if we understand it in pedestrian ways.

My task here is to show why eliminativism or nihilism about free will has not yet been earned, at least in those quarters that tend to say some form of science shows that we lack free will.[1] The target of my discussion is limited in an important way. My present target is work by contemporary scientists—Montague, Haggard, Bargh, Greene and Cohen, Pockett, and Cashmore, among others—who claim that something about the state of contemporary science shows that we lack free will. The present account is *not* intended to function as a reply to free will skeptics whose work is grounded in largely philosophical considerations. I've engaged with that work elsewhere, and while my grounds for resisting it are not unrelated to some of what I say here, the details are distinct (see Vargas 2013, esp. chapters 2, 3, and 9).

My engagement with scientific free will skeptics is not motivated by a more broadly anti-scientific, anti-physicalist, or anti-naturalist view of these matters. I accept, as they do, that science is relevant to our understanding of free will, and that our growing grasp of the physical constituents of abilities will change how we understand ourselves. Nevertheless, I maintain that our concept of free will is more resilient and flexible than most free will skeptics tend to acknowledge, and given their concerns, there is typically better reason to prefer a revisionist conception of free will over free will skepticism.

I make my case in four parts. First, I will distinguish between several distinct problems associated with free will, and identify the strand that I take to be central (or, at least, central to the issues discussed here). Second, I turn to criticism, illustrating what is confused or mistaken about some of the main paths to contemporary forms of scientific free will skepticism. There, I will argue that many of its familiar forms turn on overly-ambitious claims about determinism, characterizations of free will, and what follows from a commitment to a scientific worldview. Third, I argue that even the best forms of scientific skepticism fail to block the possibility of revisionism. Finally, I offer a revisionist alternative that captures much of what motivates skepticism about free will, without giving up the integrity of our ordinary judgments and practices.

I.

What is free will? One can, of course, define any term as one sees fit, if mere stipulation is the name of the game. However, a proposed definition is only interesting if it corresponds to either ordinary usage or some technical notion whose appeal or function can be established. So, when someone claims that free will is some or another thing, we should ask ourselves whether there is any reason to think that the proposal does capture ordinary usage, or whether it represents a specialized notion, the interests of which should be stated.

Free-will-doubting scientists are not univocal about what they think 'free will' means. For example, Wegner has characterized it as "the feeling of conscious control" (Wegner 2002). This notion is markedly distinct from "undetermined choices of action," (Bargh 2008, 130) and the idea that we choose "independent of anything remotely resembling a physical process" (Montague 2008, R584-85). So, for

any of their accounts, we should ask whether such characterizations are intended to reflect folk or technical discourses.

Although these are serious, thoughtful scientists doing important work in their primary fields, what is striking is that none of them give any evidence that their definitions of free will track either ordinary *or* standard technical discourse. In the articles from which these quotes are extracted, there is no appeal to any empirical or experimental work on how it is that ordinary folks actually think about free will, nor argument for why we should think of free will as they characterize it, nor any engagement with those whose professional occupation it is to think about these things.[2]

This might be reasonable if there were no disagreement about the meaning of free will, or whether it exists, but that is surely not the case. Indeed, looking at the specialist literature on free will, it is clear that there is disagreement about all of these things. Philosophers have variously characterized free will as the ability to do otherwise, decision-making in accord with reason, a mesh between one's actions and values or privileged desires, and so on.

These considerations suggest that we do better to characterize free will in some way that does not beg the question against the various accounts and the ensuing disagreements about its existence. One way to do this is to consider the how and the why of our concern for free will. If we start with why we are supposed to be worried about free will, we can move to consider what sorts of powers would be adequate to secure that thing.

Indeed, a cursory examination of the sprawling literature on free will reveals that free will is rarely presented as significant for its own sake.[3] More commonly, it is invoked as a key ingredient in some other, more readily apparent concern. We are usually told that free will matters because it is crucial to, for example, the truth of our beliefs about ourselves when we deliberate, our deservingness for moralized praise and blame, all of morality, the idea of human agency itself, or our general distinctiveness in the natural order.

The version of free will at stake in the present essay is this: free will is the power or capacity characteristic of agents, in virtue of which they become appropriate targets for moralized praise and blame. This is a "responsibility-centric" notion of free will. On this way of putting things, the stakes are whether we have the power required to license moralized praise and blame. The free will nihilist holds that we lack the sort of power necessary for moral responsibility, so that our inferences about responsibility are mostly false, our responsibility practices (including praising, blaming, and punishing) are unjustified, and our attitudes are unwarranted, even if inevitable.

One advantage of making explicit an understanding of free will as linked to responsibility, is that it anchors philosophical concerns in something comparatively concrete and undeniably important to our lives. This is not a sense of free will whose only implication is whether it fits with a given philosopher's particular speculative metaphysics. It is not a sense of free will that is arbitrarily attached to a particular religious framework. Instead, it is a notion of free will that understands its significance in light of the role or function it plays in widespread and recognizable forms of life.

There are two things to notice here. First, this way of characterizing free will doesn't rule out the possibility that, as Montague and Bargh suggest, free will requires something like substance dualism. Nor does it rule out Wegner's characterization of free will as nothing more than a kind of experience. What it does clarify, though, is what general sort of thing free will is supposed to be, or what is supposed to be important about the property of free will—i.e., it is supposed to be a power ordinarily necessary for

culpability.

A second point concerns the scope of the account. Although a responsibility-centric notion of free will is plausibly the notion that animates at least a plurality of philosophical discussion under the label "free will," it is not the only notion with some claim to the label. I will have little to say about other notions, but by the end of this essay it should be clear that the basic lessons about alternatives to skeptical views about free will generalize across alternative construals of the role of free will.

What is at stake here is whether responsibility-centric versions of "no free will" views are warranted or true. There is no standard terminology to refer to such views. I will interchangeably use the phrases *free will skepticism*, *free will nihilism*, and *eliminativism about free will* to refer to views on which free will does not exist, and I will use the phrase *scientific free will skeptics* to refer to those who maintain that, on scientific grounds, we should conclude that free will does not exist.[4]

None of this unduly stacks the deck against the nihilist. In treating free will in roughly functionalist terms—in offering a specific approach that makes central its role for responsibility practices, judgements, and attitudes—traditional skeptical worries remain live. For all I have said, the role in light of which we understand free will may be incoherent. Or, it might be that nothing in the world can fill the role. Or, it might be filled, but on balance there are other reasons why we are better off not deploying a notion of free will in our linguistic and social practices. If so, then free will and moral responsibility nihilism or a related view might well follow. Accepting a responsibility-centric conception of free will does not guarantee the tenability of the ensuing first-order accounts of free will. What it does ensure, though, is that we are clear on the stakes, clear on the kind of power the nihilist is denying, and clear that our target concept is plausibly connected to the long-standing philosophical and scientific traditions concerned with free will.

II.

In this section, I present three challenges to standard forms of skepticism that purport to be grounded in scientific concerns. The first concerns difficulties in the invocation of determinism. The second concerns the widespread failure of scientists to see the view known as *compatibilism* as something other than verbal or semantic subterfuge. The third involves a basic challenge to simple forms of scientific reductionism. The overriding theme in this section is that it is difficult to get an interesting scientific conclusion about free will without helping yourself to a substantive—and usually contested—account of free will. Or, to put the point differently: if science drives your free will skepticism, you better be prepared to do a lot of philosophy too.

Determinism

There are many paths from science to free will skepticism. Perhaps the most venerable path goes through determinism. Although it is a less common refrain than it once was, one sometimes still hears assertions that science in general is deterministic.[5] For example, psychologists Lee and Harris have claimed that, "most scientists are intrinsically deterministic" (forthcoming, ms 3).[6] More reservedly, some merely claim that determinism describes some or another particular field of science. For example, neuroscientist Patrick Haggard has claimed that, "Neuroscience is fundamentally deterministic in its methods, its assumptions and its outlooks" (2011, 8).

Whatever else is true, on pain of denying standard interpretations of quantum mechanics, no scientists with passing familiarity with 20th century developments in physics should think determinism is an obvious feature of *every* part of the world subjected to scientific investigation. Of course, there is a minority of scientists and philosophers who have rejected the standard interpretations of quantum mechanics.[7] Doing so, however, requires a good deal of work and at the very least it involves swimming against the dominant scientific current of the time.

Moreover, many particular subfields (including neuroscience) provide no clear home for determinism either, despite Haggard's assertions. As his fellow neuroscientists Atmanscpacher and Rotter note, "the descriptions of brain behaviour currently provided by neuroscience depend on the level and context of the descriptions. There is no clear-cut evidence for ultimately determinate or ultimately stochastic brain behaviour. As a consequence, we see no solid neurobiological basis to argue either in favour of or against any fundamental determination or openness of human decisions and actions" (2011, 85).

Although appreciation of this point is not as widespread as it perhaps should be, there is growing recognition among neuroscientists and philosophers that whatever the temptation to presuppose lurking deterministic explanations in neuroscience models, nothing in the neuroscientific data or current models of it actually supports philosophically significant conceptions of determinism (Roskies 2006; Roskies 2010; Koch 2012, esp. Ch.7).

Matters are not different in psychology more generally. There is virtually nothing in the experimental data to support deterministic generalizations of notable human behavior. It would be nothing short of a miracle if we were to regularly produce behavioral studies that isolated the variables that produce notable target behavior at rates in the 90-95% range—and this would still fall short of demonstrating determinism as it is conventionally understood.

So, whatever its venerable path to worries about free will, there are no easy generalizations to be had about science's deterministic significance. Scientists might yet give us reason for thinking determinism is true, but given the complexity of the issue in physics—where our most nuanced models of determinism are to be found—it seems unlikely that Science will issue any definitive edicts about determinism anytime soon.

Compatibilism as a definitional gambit

The path to scientific skepticism is relatively straightforward if (A) one thinks that we face a choice between determinism on the one hand, and a form of non-deterministic agency on the other and (B) one thinks there is no good evidence that we are non-deterministic agents. In the philosophical literature, *libertarianism* is the name for the sort of free will that is incompatible with determinism. So, one might conclude that the dearth of evidence for libertarianism is sufficient warrant to conclude that free will does not exist.[8]

Even allowing that one can't show that determinism is true, what this familiar "free will *or* determinism" picture leaves out is, arguably, the most influential view in the 1000+ years in which people have been grappling with these issues. The alternative view is known as *compatibilism*, and it is the view on which free will is compatible with determinism. Compatibilism's string of proponents date back to at least the ancient Stoics. Unfortunately, recognition of compatibilism as a live possibility seems largely absent in typical scientific discussions of free will. Moreover, if compatibilism gets any airing at all among scientists, it is usually to dismiss it (Nahmias forthcoming; Mele forthcoming). Although there are plausibly diverse

motivations involved in the dismissal of compatibilism, the motivation for dismissing it seems most often connected to the thought that compatibilism amounts to a kind of semantic subterfuge that sidesteps the real issue. So, for example, in this volume Susan Pockett claims that "Compatibilism is simply a definitional choice, and as such can not be either proved or disproved by any variety of science" (forthcoming, ms 7).

Pockett's claim provides an especially useful illustration in the present context because it brings together two threads of the scientific dismissal of compatibilism (1) the idea that compatibilism is something of a definitional gambit, and (2) that the commitments of compatibilist theories are somehow insulated from or orthogonal to scientific matters, and thus, the core of the free will debate. Neither are tenable.

Consider the second idea, that compatibilist accounts are cabined off from scientific matters and the core of the free will debate. As Pockett portrays it, this follows from the idea that compatibilism is a "definitional choice." If "definitional choice" just means "proposed definition" then this fact hardly suffices to show that compatibilist construals of free will are immune to scientific discomfirmation. To see why, consider a different sort of case where one might plausibly make a "definitional choice" but where it is manifestly clear that scientific findings impinge on the tenability of the choice.

Suppose I choose to define 'race' as a heritable biological kind. It does not follow that because the biological proposal for understanding race is a "definitional choice" that it can be neither proved nor disproved by any variety of science. Famously, scientists and philosophers have argued both for and against biological conceptions of race, and most of the involved participants have agreed that if we are clear about the definitional matter, we can show (or not) the existence of race, thus understood.[9]

Even if we can make non-trivial sense of the idea that compatibilism is somehow "definitional" it is hard to see how compatibilism entails isolation from scientific issues in the way Pockett claims.[10] Whether compatibilists models of agency can be squared with what we have learned from social, cognitive, and neuropsychology is something a good number of philosophers and scientists have been worried about.[11]

These concerns do not disappear if we shift from the general position of compatibilism to the particulars of individual compatibilist theories. Whether we have identifiable mechanisms for recognizing and responding to reasons (Fischer and Ravizza 1998), whether we have a general capacity to recognize moral reasons (Wallace 1994), whether we have the kind of self-knowledge required to be able to identify with our motives or values (Frankfurt 1971; Watson 1982), all these accounts invoke commitments vulnerable to experimental disconfirmation. If we could empirically demonstrate that our responsiveness to reasons was an illusion, or that we lack particular mechanisms to recognize reasons, or that we consistently lack robust self-knowledge of our motives and values, then we would, by those compatibilist accounts, lack free will. Contrary to Pockett's characterization of compatibilism as insulated from science, we should instead think that science can inform, threaten, and refine philosophical accounts of compatibilism.

Let's put aside the implications we might draw from the allegedly "definitional" nature of compatibilism. A more charitable way to characterize recurring doubts among scientists about compatibilism is to portray disputes between compatibilists and incompatibilists as engaged in something like David Chalmers' (2011) notion of a "verbal dispute." On Chalmers' account, "a dispute between two parties is verbal when the two parties agree on the relevant facts about a domain of concern and just disagree about the language used to describe that domain" (515). Perhaps the picture is this: there is

consensus that we have the powers described by compatibilists. Whether such powers are labelled 'free will' or not is a verbal dispute in the sense just given: we can call it free will, but the scientific status of compatibilist powers is hardly in dispute. What is in dispute, on the picture under consideration, is whether the additional powers postulated by libertarians of various stripes are scientifically plausible.[12]

As initially appealing as such a view of the free will problem might be, it importantly misconstrues the operative stakes in the bulk of serious work on free will over the past 40 years or so. What is (and has been) at stake in the compatibility debate is whether the powers identified by compatibilists are sufficient to support, justify, or explain attributions of responsibility (or the ability to do otherwise, or the ability to form true beliefs under deliberation—depending on the particular theorist's interests). If such powers are insufficient, then what we need to know is whether different powers—namely, the diverse powers invoked by various strands of incompatibilism—would be sufficient.

If the dispute were merely verbal, then the compatibility debate should have withered away once the involved parties recognized that the powers appealed to by compatibilists were different than those appealed to by libertarians, and that we could mark that difference by referring to "libertarian free will" and "compatibilist free will."[13] That hasn't happened, and for good reason. The disagreement between parties really is over *which* powers are sufficient to support the relevant practices, judgments, and attitudes. It is a disagreement about what sorts of agents we need to be for our suite of responsibility-characteristic phenomena to be in good standing. It is only with some grip on the answer to *that* question—what's required for moralized praising and blaming?—that we are in a position to evaluate whether we have such forms of agency. Incompatibilist and compatibilist accounts represent different camps in a substantive disagreement about that question.

Scientific verdicts about our agency typically help to resolve questions about free will only when coupled to a substantive (and usually contested) view about the requirements of free will. Given that one can't resolve this philosophical issue without doing some philosophy, it is better that such work be done straight on, rather than inadvertently. [14]

So, compatibilism is not a definitional gambit, and disputes with compatibilism are not merely verbal disputes. If one wants to show that science demonstrates the nonexistence of free will, one will have to contend with the richly developed compatibilist accounts of free will out there. To my knowledge, no scientist proclaiming the non-existence of free will has seriously undertaken this task.

Reduction and sourcehood

Here is one way to characterize much of the foregoing: most interesting claims about what science does or does not show about the free will debate turn at least in part on substantive and frequently contested philosophical positions on free will. In some cases, though, the philosophical presuppositions are less apparent. Consider the recent remarks by distinguished biologist and member of the National Academy of the Sciences, Anthony Cashmore (2010). Cashmore claims that

> Many discussions about human behavior center around the relative importance of genes and environment, a topic often discussed in terms of nature versus nurture. In concentrating on this question of the relative importance of genes and environment, a crucial component of the debate is often missed: an individual cannot be held responsible for either his genes or his environment. From this simple analysis, surely it follows that individuals cannot logically be held responsible for their behavior (4499).

At first blush, Cashmore's discussion seems to trade on a regress argument akin to the one that motivates Galen Strawson's views on the impossibility of moral responsibility (Strawson 1994). However, there are a variety of reasons to resist the conclusion that from non-responsibility for genes and environment we get non-responsibility for behavior.[15] Perhaps the most obvious is this: unless the particular inference is defended (which Cashmore doesn't try to do) it appears that he is simply helping himself to a dubious principle on which if the origin (or constituents) of something (responsible behavior) lacks a particular property, then the product (or constituted entity) lacks that property. At least as stated, that principle cannot be right. Such a principle would show that because hydrogen is not a fluid and refreshing to drink, water cannot thereby be a fluid and refreshing to drink.

So, it simply does not follow from Cashmore's "simple analysis" that individuals cannot "logically be held responsible for their behavior." But suppose we decide that this disagreement is a draw. Cashmore goes on to object that proponents of the view that we have free will simply fail to provide the kind of explanation that any serious account of free will should possess. He writes: "Whereas much is written claiming to provide an explanation for free will, such writings are invariably lacking any hint of molecular details concerning the mechanisms" (4499-4500).

Now Cashmore's engagement with philosophical accounts is fairly limited, if his references are any indication.[16] So, perhaps he just hasn't read the various accounts of, for example, the mechanisms of reasons-responsiveness or identification or value-coherence or quantum amplification or so on, that are thought to be involved in free will. Or, perhaps he is looking for a different kind of explanation. (Perhaps he means "molecular details" to mean "an account of the constituent parts of the ordinarily invoked notions" or perhaps he simply means *molecules*. The text is not clear on how this passage is meant.)

If Cashmore is simply unaware of the literature, then we can put his objection to the side. However, I suspect that he is giving expression to a deeper worry connected to broader issues in ontology. What I have in mind are worries about reductionism and the level of description at which the problem of free will arises.

I have speculated elsewhere that scientists are sometimes motivated by something philosophers call "source" intuitions, i.e., the thought that moral responsibility requires that we must be the ultimate origins of at least some strands of the causal nexus. On such accounts, free acts must be partly free of causal antecedents prior to the relevant decisions of the agent or the agent's free formulation of the relevant characterological inputs to that decision (Vargas 2009; Vargas forthcoming). For those who take such views to capture the essential element of free will, the requirement is that we must sometimes be, in some suitable fashion, the ultimate sources of what we do.

As a purely conceptual matter, there are a variety of ways we might satisfy that demand.[17] The important idea for scientific free will skeptics of the sort suggested by Cashmore is that scientific results impugn the possibility of satisfying the sourcehood requirement. The real worry, I'm suggesting, seems to be this: Cashmore and others are afraid that science shows that we are ultimately constituted by the same stuff as everything else. As a consequence, the only powers we have are those afforded to us by those (physical) things that constitute us. Since humans are composed of lower-level physical entities—molecules, as a start, but more fundamentally atoms, and even more fundamentally energy—the powers we have are limited to the powers of our constituents. Since those elements and their powers are just part of the universal causal nexus, there is no room left over for humans possessing some special, originating or ultimate source of what they do. So, free will and moral responsibility must be rejected.

On this picture, then, it is a broadly reductionist impulse that motivates a great deal of free will

and moral responsibility skepticism among scientists.[18] As they see it, the arc of the scientific worldview shows that there are no good hooks on which to hang the hope for sourcehood. Substance dualism is untenable, but so is nearly any other approach that tries to find a place for special human powers for originating causes in the universe.

Questions of reduction and emergence are thorny issues, and they involve a good deal more than debates about free will. However, there is reason to be skeptical about versions of reductionism that proceed on the basis that Cashmore and others have offered. The operative idea in Cashmore seems to be that if we don't get to molecular details in an account of the mechanisms of responsibility, we aren't properly explaining things. It is not clear why a molecular-level account (as opposed to, say, an account at the level of cognitive mechanisms on one hand, or atomic properties on another) is privileged.

The impulse to treat "lower" level explanations as privileged may be well founded in some cases. However, matters are much less simple with social-normative notions like moral responsibility (and on the present account, free will). The more plausible view of these things is that our explanatory interests and social practices play some role in settling what is a relevant and irrelevant explanation for social-normative notions.

For example, suppose a Martian came to visit and proceeded to invite us to write something up about the nature of touchdowns. We comply, writing all the customary things, such as "a touchdown is worth 6 points" and "touchdowns are earned when a player in possession of the ball during play breaks the plane of the opposing goal with the ball" and so on. We would be surprised if the Martian replied to this treatise by saying (or telepathically articulating): "Whereas much is written claiming to provide an explanation for touchdowns, such writings are invariably lacking any hint of molecular details concerning the mechanisms."

At this point, we would have good reason to insist that the Martians are looking for the wrong sort of explanation. For some kinds of explanations, we might say, what is at stake are relatively high-level phenomena. They might be *constituted* by lower level things, but the *relevant* properties and their significance to us are simply not to be found at those lower levels. If you go looking for the core of an explanation of touchdowns in the molecules of the players, the field, and the pigskin, you are just looking in the wrong place.

Of course, it is unremarkable that biologists are interested in biological things, that physicists are interested in physical things, and that sociologists are interested in social things. But there's an old carpenter's lesson here: we shouldn't think everything is a nail just because we are carrying a hammer. Perhaps Cashmore really is interested in the biology of moral responsibility for its own sake. That's fine. However, at least for any conception of agency that is interested in the powers that license praise and blame, biology is relevant only as a realizer or constituent, and not as the principle source of puzzles and solutions. At first blush, the interesting agentive and social-normative dimensions of free will and moral responsibility should be higher order properties. As in football, the molecular details matter less, and our high-level concerns may tolerate mutiple, distinct realizers when it comes to the molecular constituents of fields, football players, end zones, free will, and moral responsibility.

One *could* ask questions about the biology of free will, just as one could ask about the molecular features of touchdowns. Sometimes, the answers to such questions will be interesting—for example, the molecular constituents of footballs might have some role to play in whether touchdowns are more or less likely in some conditions and not others. Again, our theories of free will and moral responsibility are not

obviously altogether insulated from scientific developments.

What will make scientific developments interesting and informative for free will and responsibility, though, is *how* they connect to the social-normative roles that drive the basic philosophical problem. As a matter of framing the general free will problem—what powers suffice to license moralized praising and blaming?—starting from biology, physics, or some other sub-agential science is an altogether circuitous path to discoveries about what's at stake. Moreover, as we have seen, such an approach hardly avoids taking uncontested stands on a wide range of substantive philosophical issues—including the possibility of emergent causal powers, whether free will has a sourcehood requirement, how that requirement should be construed, and so on.

In sum, it really is difficult to go from science to free will skepticism without taking a stand on substantive, usually disputed philosophical matters.

III.

There is something powerfully appealing about free will nihilism, despite the shortcomings in the particular claims by scientists. Free will appears to be an artifact of a prior conception of the world, the sort of thing that would not be taken seriously if we had not inherited it from important cultural and moral authorities.

Such thoughts may be driven, in part, by the sense that close inspection of the apparently distinctive powers of humans reveals that they are not radically disconnected from the physical nature of the world. This fact alone can be unsettling to those whose image of humans is that they are radically set apart from nature. Indeed, reflection on the more general trajectory of the sciences may fuel the worry that free will is just another instance of metaphysical nonsense that the sciences gradually expunge from "the manifest image" of ourselves—that, is our widely received, pre-scientific self-conception.

These are potent considerations against any case for free will. Thus, even if we have reason to find wanting the standard paths to free will skepticism, it is not hard to believe that science threatens *something* that seems quite important to us. The question, then, is what we are to say about these matters, given dissatisfaction about the familiar paths to skepticism on the one hand, and on the other hand, uneasiness about supposing that science changes nothing. One path is to cast about for a new challenge from some or another branch of science. Perhaps we will find that there is some further distinct scientific threat to free will, unreliant on the sorts of claims canvassed thus far. Although I cannot hope to vindicate the claim here, attention to the literature on these issues strongly suggest that the recurring lesson herein tends to generalize: there is no clear path from science to free will nihilism that does not involve significant philosophical engagement.[19]

A second path is to shift gears, accepting the need to hitch free will skepticism to philosophical commitments. That is, we could pursue philosophical free will nihilism. As I noted at the outset, the scope of my arguments in this paper leave untouched the possibility that there might be good philosophical grounds for adopting free will nihilism. So, one could elect to take up the familiar philosophical issues, among them, debates about the metaphysics of agency, the analysis of capacities, the conceptual demands of freedom and responsibility, the whether and how of a raft of thought experiments, and so on. Doing so is no guarantee of a skeptical conclusion, though. Indeed, skeptical views remain something of a minority view in the literature. Antecedent to doing the work, there is no special reason to think that philosophical

free will skepticism will be sustained.

There is a third path. It is a path that accepts the skeptic's concerns about our manifest image, while insisting that we can retain a commitment to free will if we refashion our understanding of it. This is the *revisionist* path. Revisionism about free will is the view that either (A) we've misunderstood the nature of free will, or (B) we have misunderstood what it requires, or (C) at least all the conceptual and practical work for which free will has been invoked can be done with forms of agency that are compatible with the going scientific worldview.

To appreciate the basic idea of revisionism, it helps to first distinguish between an account of what we think about some domain or idea (e.g., free will) and what we ought to think about it. Call an account of the former *diagnostic*, and an account of the latter *prescriptive*. An account is revisionist to the extent to which the prescriptive account recommends a view about the subject matter that conflicts with aspects of the diagnostic account. We can say that an account is revisionist when it tells us that what we ought to think is in conflict with what we do think about that thing. Ergo: a revisionist account of free will is at odds with how we think about free will.

What making this distinction permits us to see is that the ordinary grist for the free will skeptic's mill tends to license conclusions only about how we think about free will. If I give you an example of how a universe, deterministically characterized, undermines our convictions about free will, I've only shown you something about how you and I think about free will. It remains an open question whether you and I have accurate beliefs about free will and/or whether we should change our minds about what free will requires.

In the absence of some independent account of why our ordinary convictions about free will should be thought to tell us about the Real, True, Essential nature of free will, we should only think that our reactions to experiments—whether imagined or real—tell us only about the best diagnostic theory of free will. They do not, by themselves, settle what the best prescriptive account is of how we ought to understand free will, all things considered.

One might protest that any revisionist about free will needs to explain why we should think that free will is, or could be, different than we ordinarily suppose. Why should we tolerate the idea that we somehow have importantly mistaken beliefs about free will? Isn't this just a dodge?

In reply, there are a variety of reasons why we should take seriously the possibility that the best prescriptive account of free will could conflict with how you and I intuitively understand free will. First, there is just the fact that there might not even be much agreement in our linguistic community about what free will means. You and I constitute a pretty limited data set, and it might turn out that there are important divergences in our linguistic community, or that there are notable differences across communities.[20] If so, then we would face questions about whether some or another of these usages are better or worse, and on what grounds we could plump for one or another notion of free will.

Second, it might turn out that we are confusing the metaphysics of free will (what it requires) with the "epistemological pragmatics" of our practices, that is, with the heuristics we use given our limited knowledge about people's intentions and beliefs and the like (Vargas 2006). Disentangling our conception of free will's ontology from our cognitive shortcuts used to track it in the real world is no easy task. And, to the best of my knowledge, no one has a particularly good account of when ordinary judgments about some philosophically contested matter reflect an assessment of only the metaphysics of the matter and when it amounts to the fallible deliverances of our (correctly or incorrectly) deployed epistemological heuristics.[21]

Third, there are reasonably familiar pictures in the philosophy of language for thinking that at least some terms function in ways that don't entirely depend on how the speaker intends to use the word. At least sometimes, I don't need to know what water is, in order to refer to water. It might well turn out that I have some deeply confused or mistaken beliefs about water (that it is a basic, indivisible substance, for example). Nevertheless, water exists and I can talk about it, despite my confused beliefs.

This sort of picture—a broadly referentialist picture—is not the only one that tolerates divergence between how we think about something and what it is. Suitably sophisticated "internalist" accounts allow for some flexion between beliefs and the world, as well. For example, one can hold that a term refers to whatever property it is that renders most platitudinous sentences about some subject matter true. Here, reference derives in part from what we think about things, but also from the fit our usages have with the properties of the world. In either case, there is room for 'free will' to pick out something distinct from our naive uses of the term.

Fourth, it might turn out that our traditional or naive interest in free will (assuming it is reasonably unified), doesn't uniquely underpin anything practically significant (say, deservingness for praise and blame). So, for example, suppose that *if* we had a libertarian form of agency it would suffice to justify praising and blaming. Even if we discovered we lacked libertarian free will, we might yet discover that some other form of agency or some other feature about us and our practices justifies deservingness for praise and blame. If so, whether we keep free will talk (in the present responsibility-centric sense) or not appears more a matter of convention than anything substantive about philosophy or science.

To be sure, I haven't here shown that any of these things are indeed the case in a way that definitively demonstrates the viability of revisionism about free will. That's a case I have tried to make elsewhere (Vargas 2013a). Nevertheless, the considerations presented here weigh heavily against our simply supposing that revisionism is a non-starter. If that's right, then skeptics about free will have a good deal more work to do before they have shown that we must accept nihilism about free will.

What do revisionist accounts look like? I will say more about this in a moment, but for now let us just assume they will look a lot like any number of compatibilist accounts, e.g., emphasizing elements like reasons-responsiveness, identification, or what have you. The main difference is that the revisionist does not claim that such accounts capture ordinary notions of freedom, but are instead replacement accounts, or accounts of what we ought to mean, all things considered. Indeed, the revisionist thinks that the skeptic may well be right about something important: science (or philosophy, or what have you) shows that our ordinary understanding of free will is flawed, and compatibilists are wrong to claim otherwise.

With these pieces in place, we can now begin to see how the dialectic between the revisionist and skeptic unfolds. If the free will skeptic wants to show that free will doesn't exist, as opposed to the revisionist's claim that free will exists but is (say) more like what compatibilists talk about, we need to know *why* the skeptic thinks we should conclude that free will doesn't exist.

One instructive reply is provided by a widely-cited paper by Greene and Cohen (2004) that makes just that case. They argue that "when it comes to the issue of free will itself, hard determinism is mostly correct. Free will, as we ordinarily understand it, is an illusion" (1783).[22] The core idea in their paper is that the emerging picture in neuroscience threatens free will and responsibility as we intuitively understand those terms, and that despite the law's neutrality about the metaphysics of free will, a growing appreciation of the deterministic threat will lead us to jettison our commitments to free will, responsibility, and retribution. They conclude that this will ultimately reshape the law on matters of punishment.

Perhaps neuroscience can and does threaten these things in just the way they claim. The operative question here is *why* neuroscience threatens these things. On this point, Greene and Cohen offer a familiar villain: determinism. They claim that "contrary to legal and philosophical orthodoxy determinism really does threaten free will and responsibility as we intuitively understand them" (1780). To illustrate, they introduce the idea of "Mr. Puppet."

Mr. Puppet is characterized as an engineered person, created with a particular personality and behavior profile through powerful control of his genetics and environment. Mr. Puppet's architects ensure that he has the relevant experiences and opportunities so that, by design, he commits a murder during a failed drug deal. Greene and Cohen contend that Mr. Puppet ought not be held legally responsible because once we realize the engineered nature of his actions, "it is hard to think of him as anything more than a pawn" (1780).

This is an elegant, intuitively powerful argument. However, as it stands, it does not show that determinism undermines responsibility. For starters, as Greene and Cohen themselves note, the example doesn't actually guarantee an outcome in the way we would expect from ideal deterministic control. Mr. Puppet's architects, they note, have a 95% success rate in controlling his action. Perhaps there is measurement noise, or some other error in prediction and control of a deterministic system. Still, even if we assume the control is deterministic, a similar example would show that it isn't determinism that is doing the undermining of responsibility. If Mr. Puppet's architects simply had a hard metaphysical limit of 99.9999% effective control over Mr. Puppet, given otherwise maximal control over Mr. Puppet's constitution and environment, that degree of control would presumably make us wary of assigning responsibility even if, by stipulation, Mr. Puppet's behavior was not deterministic.

So, if an example like Mr. Puppet can show a threat under both determinism and indeterminism, it looks like a mistake to present *determinism* as the threat. Instead, something like a regress worry arising from *causal embeddedness* or "being embedded in a system of causes" seems to be the root of the problem. To their credit, Greene and Cohen are not entirely insensitive to this point, sometimes appealing to the "mechanistic" nature of decision-making as the real threat.[23] (Although, notice that even if we somehow concluded that Mr. Puppet's architects could control his behavior through ectoplasmic manipulation or emotional vibrations in the ether, the same exculpatory intuitions get going. Mechanism, understood in a physicalist sense seems—like determinism—a special case of worries about causal embeddedness.)

Here, though, some of the revisionist tools mentioned above can do some work for us. Granting that we have the reactions Greene and Cohen describe for the Mr. Puppet case, this does not yet tell us anything about free will or moral responsibility, or what prescriptive account we should have of it. What it does tell us about are our received, i.e., naive *judgments* in this particular case. It is, at best, understood as a diagnostic account that is (one suspects) being taken to license a prescriptive account.

The move from diagnosis to prescription is not so easily warranted. Suppose that we could get convergence about both the substance of our judgments about the Mr. Puppet case, and that we could explain the psychological mechanisms that yield that judgment. For all that, we would not have obviously shown anything about whether people have free will . After all, showing that naive judgments about physics fall into recognizable patterns because of identifiable psychological mechanisms of judgments does not show anything about physics. So, why would showing the same about our judgments of free will show us anything about free will?

There are a number of ways to appreciate the force of this point. Consider why highlighting the

fact of causal embeddedness might undermine our existing judgments of responsibility. It could be that we have residual worries about the possibility of agency in the face of reductionism about causes. However, as we've seen in the discussion of sourcehood concerns and Cashmore's concern for molecular details, there is a lot of heavy philosophical lifting to be done if someone wants to make the case that such judgments are *correct*.

A different potential source of our exculpatory intuitions may lurk in the idea that Mr. Puppet is controlled or manipulated. The scope of control and manipulative power by Mr. Puppet's architects may drive our intuitions simply because of how remote and unusual such powers are from the circumstances in which we ordinarily deploy our responsibility judgments. It may also be the case that we are particularly disposed (whether by evolition or socialization) to detect "cheaters," or social norm violators.[24] For that matter, the name "Mr. Puppet" presumably primes our "not responsible" judgments. I imagine if we described the agent as Will Power, this would color our reactions in various ways. Finally, some of our reluctance to ascribe responsibility may simply be a byproduct of the habit of looking to find a best candidate for responsibility, generally ignoring lesser candidates. However, responsibility is not a zero-sum thing. You and I might both be fully responsible for some outcome, without either of us being solely responsible. So, to some extent, our thinking about this case may reflect more about sloppy habits of mind than careful thinking about the conditions of responsibility.

The proponent of prescription-by-diagnosis theorizing might provide replies to each of these objections in turn, but the revisionist has more to say. Let us even suppose that we can show that convergence in judgments that Mr. Puppet is not responsible does not hinge on the errors I have noted. Before we accept the truth of our initial judgment, we might wonder whether the case has been under-described. Has Mr. Puppet (or better, Will Power) been subjected to manipulations that rob him of the requisite powers to recognize and respond to reasons? If they have, then why isn't *that* the reason he is not responsible?

If Mr. Puppet/Will Power does retain his powers to recognize and respond to reasons, then why shouldn't we think he is, after all, responsible?[25] For the sake of argument, let's suppose that we persist in thinking Mr. Puppet is not responsible. We still might ask whether something about Mr. Puppet's circumstances unduly burdened the operations of those powers. Ordinary dollops of self control tend to fare badly in light of extraordinary temptations. If a problem with Mr. Puppet is that his ordinary degree of self-control has been subjected to extraordinary stress or temptation, then this might be the source of exculpatory impulses, depending on what it is that we think praise, blame, and responsibility are *for*, what they *express*, and what they *do*.

It should be evident that there are lots of places to get off the train going from diagnosis to prescription. For all that, though, one might insist that the troubles with skepticism are just as strong a reason to adopt agnosticism as they are to adopt revisionism.

It is here that the onus is on any account with revisionist aspirations. If we wish to concede that there are problematic threads to our ordinary conception of free will, while still insisting that the notion can be fruitfully recast, the revisionist needs to offer an account of the basis of the revision.

IV.

There is no single way to construct a revisionist account. In the previous section, I noted that a revisionist account can look a lot like compatibilism, absent the pretension to describe our convictions as we find them. Nevertheless, given that standard compatibilist accounts were constructed with non-revisionist aspirations, it might turn out that the best revisionist account will have some distinct contours. In what follows, I present a revisionist framework that is indebted to many compatibilist accounts, but that has its own distinctive features as well.

Start with the idea of why we care about free will. What I claimed at the outset was that, at least in many corners, the stakes are moralized praising and blaming and related assessments of moral responsibility. If free will is going to be worth the name, it will need to be the sort of thing that helps make sense of those practices and judgments. However, the skeptical scientists are right: if free will requires substance dualism, the power to initiate causal chains *ex nihilo*, choices completely causally independent of features of the world, or even decisions that always involve alternative possibilities, it seems unlikely that we have such powers. It might be nice if we did have such powers, but there is little in the book of science that would lead a disinterested viewer into thinking that we obviously have such powers. So, the question is whether anything else could be the sort of thing that helps makes sense of those practices and judgments. If there is, then we have a good candidate for a revised notion of free will, i.e., one that is not exactly what we had hoped for, but one that can do what we want from a notion of free will.

One power that we plausibly have is that we recognize and respond to reasons. What this power comes to, its limits, the circumstances of its best use, and so on, are all difficult matters. Nevertheless, the existence of this power is a familiar enough feature of our lives. If standard scientific views about us turn out to be correct, this power is physically instantiated. It need not be a property of an ectoplasmic substance or an immaterial soul. Rather, it is most plausibly a power of physical systems that, at least in principle, can be studied in considerable detail.

We need a few more ideas. One is that there is a class of reasons that are especially important to creatures like us. Call them *moral considerations*, or reasons that are tied to a characteristic suite of concerns we have, presumably derived from our imperfectly rational but social natures and the patterns of emotional reactions and judgments that follow in their wake.[26] How exactly we should characterize moral considerations, whether ordinary understandings of morality are accurate, and what exactly defines the boundaries of the moral are important questions. The present account, however, only requires that at least sometimes, we recognize moral considerations and can be moved by them. And, in the ordinary course of things, normal mature adults are such that we possess the ability to recognize and respond to moral considerations, and this is not a fact to which we are indifferent. We value this capacity in ourselves and in others, we seek to cultivate it in our children, and we generally regard it as desirable that this capacity should be maintained and even flourish. The capacity to recognize and suitably respond to moral considerations is, in many respects, a prerequisite for us to be taken seriously as fully participating members of the social sphere. We care a great deal about being seen as creatures who, to a suitable degree, recognize and respond to those considerations that our groups regard as morally significant.

Moral considerations may well be explicable in terms of more basic and familiar mechanisms. When I see you are hurt, and stop to help you, we can say that I saw a reason to stop and aid you. That explanation might decompose into more detailed explanations about the mechanisms by which I perceive and evaluate the pain of others, the interaction of those systems with my attention and sense of time pressure, and so on. At least in form, though, these do not explain away our acting for moral reasons so

much as they explain what it is to act for moral reasons.

Here is the final element we need: our ordinary practices of moralized praise and blame, the norms emphasizing how we treat each other, and our reactive dispositions to violations and supererogation in light of those norms, over time support, reinforce, extend, and fine-tune our capacity to recognize and respond to moral considerations. This does not mean that they do so in each and every instance. Nor does it mean that our responsibility norms are or should be baldly consequentialist. Indeed, part of the efficacy of these norms plausibly rely on their having robust backwards-looking elements and licensing retributive attitudes. The point is that we collectively have an interest in these norms and practices having currency in our societies. They do something for us, something that we should be loathe to lose—at least so long as we value the form of agency they support and, in time, refine.

I can now state the substantive proposal baldly: free will should be understood as the capacity to recognize and suitably respond to moral considerations. The account is revisionist, but it is not unprincipled. The capacity to recognize and suitably respond to moral considerations matters in part because it is valuable to us to be those kinds of agents who are good at such things. However, the capacity also satisfies an important desideratum. It is the sort of thing, the presence and absence of which, suffices to support the relevant roles in the characteristic patterns of judgments and practices we associate with responsibility. Where that capacity is absent we typically lack grounds for holding someone responsible, and when it is present we typically have good reason to praise and blame. Crudely put, praise and blame, as mechanisms that enhance an already present capacity, cannot fulfill that function when that capacity to recognize and respond to moral considerations is absent.

These brief remarks are only a suggestive sketch at what a fully developed revisionist account can say about the basis on which we might recast our understanding of free will and moral responsibility. The point here, though, is that such accounts exist, and they offer a good deal of what the free will skeptic insists on (Vargas 2011a; Vargas 2013a). For example, such accounts allow that contemporary science may threaten our naive view of ourselves. They also allow that compatibilism is inadequate, at least as a theory of how many of us, at least sometimes, think about free will. Moreover, the positive, revisionist aspect of the account does not ask us to pretend that free will is some anti-natural, mysterious counter-causal force. Instead, it is the kind of thing that may well be distinctive of us, but that is nevertheless entirely natural in its composition.

Science also retains a role and remains a potential threat on the account just sketched. It may yet turn out that the kinds of powers required to support our practices are insufficiently had, even on the best revisionist account. Or, we might learn that our practices only undermine the forms of agency we find morally salient. These are matters that cannot be settled by moral philosophy alone, and on these and related matters the discoveries of science will be vital for our understanding of our moral world.

This is not to say that this account does everything we might have hoped. It will not give us back our immaterial souls or our contra-causal freedom, if that's what we thought free will required. But it does get us most or even all the things for which we wanted those powers, including an explanation of why we can and should praise and blame. Revisionist free will might not be what we thought we were looking for, but it is exactly what we need, and science has not given us a reason to doubt that we have it.

V.

The principal task of this essay has not been to muster a sustained defense of the idea that there is, in fact, a set of powers and practices sufficient to justify moralized praising and blaming and ascriptions of free will. What I hope to have done is to have offered some reasons why those inclined to scientific skepticism about free will should, at worst, only embrace agnosticism, and more optimistically, may find cause to adopt a revisionist picture of free will and responsibility.

In any event, it should be clear that several familiar paths to scientific skepticism about free will are deeply problematic in ways their proponents tend not to appreciate. Moreover, there are alternatives to skepticism. For all I have said, revisionist alternatives may yet prove unsatisfying. However, there are practical reasons to aspire to revisionism over nihilism. What the free will skeptic typically claims is that we should abandon our belief in free will. As a conceptual matter, we should be careful about dispensing with ideas that are at the center of so many aspects of our social worlds. However, some experimental data suggests that disbelieving in free will promotes anti-social behavior.[27] If that is right, then we have special reason to be careful about adopting nihilism without giving revisionism a lengthy hearing.

I began by noting that it seems apparent to many that our manifest image is defective in deep, perhaps fundamental ways. Although I have only gestured at it here, it is plausible to think that there are good reasons for insisting that there is a morally central, socially vital, and conceptually defensible set of powers we have that can legitimate familiar practices of moralized praising and blaming (Vargas 2013a, esp. Ch. 5 & 8). If this latter idea is right, then a scientifically plausible account of free will is not going to be precisely what ordinarily people—and most free will skeptics—tend to think it is. Rather, a scientifically credible notion of free will will have to be revisionist, departing in specific and principled ways from strands of our troubled, fragmented ways of thinking about free will.

This still leaves us with a lot to learn from scientific work on agency. But science about the nature and limits of human agency is not the end of the free will and responsibility story any more than it was the end of the water story, the story about race, the story about morality, or the story about legality. Depending on how we think about various social and normative issues, scientific developments will impinge on those stories, and maybe even threaten familiar equilibrium points for our thinking. Nevertheless, the details of our normative interests and practices seem here to play a special role in shaping the target of our metaphysical concerns. For example, rather than concluding, as Greene and Cohen do, that knowing the causes of someone's behavior necessarily undermines our conviction that someone deserves punishment (1783), it might turn out that knowing the grounds for punishment, doing the partly *normative* work of reflecting on the justification of punishment, can firm up our commitment to punish in some cases and not others. Science might someday tell us which of those cases fall into what categories, but this will be a collaborative result, not an edict about punishment handed down solely from a lab bench.

On this picture, rather than free will being a notion disposed of by good science, free will turns out to be an issue whose shape and nature is better understood by science. None of this is to deny the possibility that there are genuine threats (from science and philosophy both) to free will. However, on the view I have sketched, the structure of scientific threats is less sweeping. Instead, it will be more about whether in these circumstances we plausibly have those powers, and whether this condition or syndrome impairs those capacities. These are typically accounts of pockets of our world, rather than accounts of the global scenery. Threats to high-level, normatively structured social phenomena are partly a matter of science, but they also remain the stuff of philosophy.

At the outset, I noted that it is something of a puzzle why revisionist-style views of free will remain the path less taken by scientists. Many scientists (including Cashmore, Montague, and Pockett) work in fields where the progress of science became possible in large part because of the willingness of researchers to entertain the thought that various familiar (if somewhat mysterious) notions could be re-written by what we learn from experimental studies of the world. What it is to be alive, the nature of the mind, and the relationship of bodies to action-initiation, were all notions that were retained and re-written in light of scientific developments. We did not ultimately jettison commitment to the mind at the first hint of difficulties with substance dualism. And, of course, we did not despair that water never existed when its indivisibility was in doubt. Why then should we do differently with free will?

There were, of course, those who thought we should do without talk about minds, mental states, causes, and so on. Moreover, there have been concepts that we have ultimately discarded as not worth saving or re-writing. Nevertheless, it seems altogether contrary to the scientific spirit to neither consider what evidence there is about actual convictions (whether folk or specialist), nor to allow for the possibility that perhaps we need to revise our understanding of free will.

There is reason to end on an optimistic note. Despite the tide of work that presumes that conflict with science always entails elimination, there are sometimes hints that at least some scientists can be sensitive to the ways in which our interests in free will may be broader than the explanatory interests of particular scientific subfields.

In that vein, John Bargh sometimes provides a model of how one might reasonably endorse limited forms of skepticism. Bargh, for example, has claimed that we do not need free will "at the psychological level" to explain a range of psychological phenomena that free will has, he thinks, been historically invoked to explain (Bargh 2008, pp. 143-45). Framing things in this manner allows for the possibility that there might be other roles or purposes for which it might make sense to admit ongoing utility to a notion of free will. If so, then a conception of free will tied to licensing moralized praise and blame can be compatible with an account of the psychology of action initiation which, by itself, need not posit an account of free will. This is, I think, one way in which serious psychological and neuroscientific work can coexist with serious philosophical work, even aspiring to be mutually informed by one another.[28]

References

Atmanspacher, Harald, and Stafan Rotter. "On Determinacy Or Its Absence in the Brain." In *Free Will and Modern Science*, edited by Richard Swinburne, 84-101. Oxford: Oxford University Press, 2011.

Bargh, John A. "Free Will is Un-Natural." In *Are We Free? Psychology and Free Will*, edited by John Baer, James C. Kaufman, and Roy F. Baumeister, 128-54. New York: Oxford University Press, 2008.

Bargh, John A., and Earp, Brian. 2009. The Will is Caused, Not 'Free'. http://www.psychologytoday.com/blog/the-natural-unconscious/200906/the-will-is-caused-not-free (accessed August 27, 2010).

Baumeister, Roy F., Masicampo, E.J., and DeWall, C. Nathan. "Prosocial Benefits of Feeling Free: Disbelief in Free Will Increases Aggression and Reduces Helpfulness." *Personality and Social Psychology Bulletin* 35, no. 2 (2009): 260-68.

Cashmore, Anthony R. "The Lucretian Swerve: The Biological Basis of Human Behavior and the Criminal Justice System." *Procedings of the National Academis of Sciences* 107, no. 10 (2010): 4459-504.

Chalmers, David J. "Verbal Disputes." *Philosophical Review* 120, no. 4 (2011): 515-66.

Clarke, Randolph. "On an Argument for the Impossibility of Moral Responsibility." *Midwest Studies in Philosophy* 29 (2005): 13-24.

Doris, John. *Lack of Character*. New York: Cambridge University Press, 2002.

Doris, John. *Talking to Ourselves: Reflection, Skepticism, and Agency*. New York: Oxford University Press, forthcoming.

Doris, John, and Murphy, Dominic. "From My Lai to Abu Ghraib: The Moral Psychology of Atrocity." *Midwest Studies in Philosophy* 31 (2007): 25-55.

Feltz, Adam. "Pereboom and Premises: Asking the Right Questions in the Experimental Philosophy of Free Will." *Consciousness and Cognition* 22, no. 1 (2013): 53-63.

Fischer, John Martin. "The Cards That Are Dealt You." *The Journal of Ethics* 10, no. 3 (2006): 107-29.

Fischer, John Martin, and Mark Ravizza. *Responsibility and Control: A Theory of Moral Responsibility*. New York: Cambridge University Press, 1998.

Frankfurt, Harry. "Freedom of the Will and the Concept of a Person." *Journal of Philosophy* 68, no. 1 (1971): 5-20.

Greene, Joshua, and Cohen, Jonathan. "For the Law, Neuroscience Changes Everything and Nothing." *Philosophical Transactions of the Royal Society of London B* 359 (2004): 1775-85.

Haggard, Patrick. "Does Brain Science Change Our View of Free Will?" In *Free Will and Modern Science*, edited by Richard Swinburne, 7-24. New York: Oxford University Press, 2011.

Honderich, Ted. *A Theory of Determinism*. New York: Oxford, 1988.

Hurley, Susan. "Is Responsibility Essentially Impossible?" *Philosophical Studies* 99 (2000): 229-68.

Koch, Christof. *Consciousness: Confessions of a Romantic Reductionist*. Cambridge, MA: The MIT Press, 2012.

Lee, Victoria K., and Lasana T. Harris. "A Social Perspective on Debates About Free Will." In *Moral Psychology, Volume 4: Free Will and Moral Responsibility*, edited by Walter Sinnott-Armstrong, Cambridge, MA: MIT Press, forthcoming.

Levy, Neil. *Hard Luck: How Luck Undermines Freedom*. New York: Oxford University Press, 2011.

Mele, Alfred. "Another Scientific Threat to Free Will?" *The Monist* 95, no. 3 (2012): 423-41.

Mele, Alfred. "Recent Work on Free Will and Science." *American Philosophical Quarterly* 45, no. 2 (2008): 107-30.

Mele, Alfred. *Effective Intentions: The Power of the Conscious Will*. New York: Oxford University Press, 2009.

Mele, Alfred. "Free Will and Substance Dualism: The Real Scientific Threat to Free Will?" In *Moral Psychology, Volume 4: Free Will and Moral Responsibility*, edited by Walter Sinnott-Armstrong, Cambridge, MA: MIT Press, forthcoming.

Montague, P. Read. "Free Will." *Current Biology* 18, no. 14 (2008): R584-R585.

Nahmias, Eddy. "Autonomous Agency and Social Psychology." In *Cartographies of the Mind: Philosophy and Psychology in Intersection*, edited by Massimo Marraffa, Mario De Caro, and Francesco Ferretti, 169-85. Berlin: Springer, 2007.

Nahmias, Eddy. "Scientific Challenges to Free Will." In *A Companion to the Philosophy of Action*, edited by Timothy O'Connor, and Constantine Sandis, 345-56. Malden, MA: Wiley-Blackwell, 2010.

Nahmias, Eddy. "Intuitions About Free Will, Determinism, and Bypassing." In *Oxford Handbook of Free Will, 2ed.*, edited by Robert Kane, 55-576. New York: Oxford University Press, 2011a.

Nahmias, Eddy. "Why 'Willusionism' Leads to 'Bad Results': Comments on Baumeister, Crescioni, and Alquist." *Neuroethics* 4, no. 1 (2011b): 17-24.

Nahmias, Eddy. "Is Free Will and Illusion? Confronting Challenges From the Modern Mind Sciences." In *Reconsidering Scientific Threats to Free Will*, edited by Walter Sinnott-Armstrong, Cambridge, MA: MIT Press, forthcoming.

Nahmias, Eddy, Coates, D. Justin, and Kvaran, Trevor. "Free Will, Moral Responsibility, and Mechanism: Experiments on Folk Intuitions." *Midwest Studies in Philosophy* XXXI (2007): 214-41.

Nelkin, Dana. "Freedom, Responsibility, and the Challenge of Situationism." *Midwest Studies in Philosophy* 29, no. 1 (2005): 181-206.

Pockett, Susan. "The Concept of Free Will: Philosophy Neuroscience, and the Law." *Behavioral Sciences and the Law* 25 (2007): 281-93.

Pockett, Susan. "If Free Will Did Not Exist, it Would be Necessary to Invent it." In *Exploring the Illusion of Free Will and Moral Responsibility*, edited by Gregg Caruso, Lanham, Maryland: Lexington Books, forthcoming.

Roskies, Adina. "Neuroscientific Challenges to Free Will and Responsibility." *Trends in Cognitive Science* 10, no. 9 (2006): 419-23.

Roskies, Adina. "Response to Sie and Wouters: A Neuroscientific Challenge to Free Will and Responsibility?" *Trends in Cognitive Sciences* 12, no. 1 (2008): 4.

Roskies, Adina. "How Does Neuroscience Affect Our Conception of Volition." *Annual Review of Neuroscience* 33 (2010): 109-30.

Sarkissian, Hagop, Chatterjee, Amita, De Brigard, Felipe, Knobe, Joshua, Nichols, Shaun, and Sirker, Smita. "Is Belief in Free Will a Cultural Universal?" *Mind and Language* 25, no. 3 (2010): 346-58.

Sommers, Tamler. *Relative Justice*. Princeton, NJ: Princeton University Press, 2011.

Spencer, Quayshawn. "What 'Biological Racial Realism' Should Mean." *Philosophical Studies* 159, no. 2 (2012): 181-204.

Spencer, Quayshawn. "Biological Theory and the Metaphysics of Race: A Reply to Kaplan and Winther." *Biological Theory* (forthcoming):

Strawson, Galen. "The Impossibility of Moral Responsibility." *Philosophical Studies* 75 (1994): 5-24.

van Inwagen, Peter. "How to Think About the Problem of Free Will." *Journal of Ethics* 12 (2008): 327-41.

Vargas, Manuel. "Philosophy and the Folk: On Some Implications of Experimental Work for Philosophical Debates on Free Will." *Journal of Cognition and Culture* 6, no. 1 & 2 (2006): 239-54.

Vargas, Manuel. "Revisionist Accounts of Free Will: Origins, Varieties, and Challenges." In *Oxford Handbook on Free Will, 2e*, edited by Robert Kane, 457-84. New York: Oxford University Press, 2011a.

Vargas, Manuel. "The Revisionist Turn: Reflection on the Recent History of Work on Free Will." In *New Waves in the Philosophy of Action*, edited by Jesus Aguilar, Andrei Buckareff, and Keith Frankish, 143-72. New York: Palgrave Macmillan, 2011b.

Vargas, Manuel. *Building Better Beings: A Theory of Moral Responsibility*. Oxford, U.K.: Oxford University Press, 2013a.

Vargas, Manuel. "Situationism and Moral Responsibility: Free Will in Fragments." In *Decomposing the Will*, edited by Till Vierkant, Julian Kiverstein, and Andy Clark, 325-49. New York: Oxford University Press, 2013b.

Vargas, Manuel. "Reconsidering Scientific Threats to Free Will." In *Moral Psychology: Volume 4: Free Will and Moral Responsibility*, edited by Walter Sinnott-Armstrong, Cambridge, MA: MIT Press, forthcoming.

Vargas, Manuel. "Review of *Effective Intentions*." *Notre Dame Philosophical Reviews* (2009): http://ndpr.nd.edu/review.cfm?id=17385.

Vohs, Kathleen D., and Schooler, J.W. "The Value of Believing in Free Will: Encouraging a Belief in Determinism Increases Cheating." *Psychological Science* 19 (2008): 49-54.

Wallace, R. Jay. *Responsibility and the Moral Sentiments*. Cambridge, MA: Harvard University Press, 1994.

Watson, Gary. "Free Agency." In *Free Will*, edited by Gary Watson, 81-95. Oxford: Oxford University Press, 1982.

Wegner, Daniel M. *The Illusion of Conscious Will*. Cambridge, MA: MIT Press, 2002.

[1] I am oftentimes skeptical that there is much use in talking about science, as though it were a single, unified thing with a consistent ontology across its fields. Nevertheless, it is striking how similar the arguments of scientists are about the non-existence of free will, whatever their field. So, despite my reservations about there being a uniform thing that is "scientific" skepticism about free will, I will use the phrase to refer to the collection of individual views of scientists that free will does not exist. In what follows, however, I try to focus on individual claims by particular scientists instead of appealing to generalizations about entire fields of scientific practice.

[2] This is a point that has been emphasized in Mele (forthcoming) and by Eddy Nahmias (2010), among others. If one looks at the data on folk beliefs, things are interestingly mixed about what people believe about free will—but you wouldn't know it from scientists discussions of why some or another bit of science shows the non-existence of free will

[3] The point here is not that no one has ever thought free will was significant in and of itself, or that it could not be separated from other concerns that we have. The claim here is only about the ordinary paths to a concern about what is evidently a metaphysical, or at least traditionally philosophical claim.

[4] The differences between views on which such non-existence is necessary, contingent, certain, or uncertain is immaterial for present purposes. However, there remains room to make use of greater terminological sensitivity to modal differences, e.g., holding that a skeptic might doubt the existence of free will, without actually believing in the non-existence of free will, or insisting that one kind of eliminativist might think free will is impossible, whereas another might think it is possible but not actual. I see nothing to be gained by insisting on these distinction in the present context.

[5] Here's a standard philosophical characterization of determinism. Let determinism stand for the following thesis: the state of the universe at any prior time and the actual laws of nature are sufficient for the state of the universe at any later time.

[6] I take it they mean, roughly, that most scientists are committed to scientific explanations being deterministic, or even more ambitiously, that all events in the world are deterministically caused.

[7] One prominent example is Honderich (1988).

[8] Elsewhere, I have argued against libertarianism on grounds that it does not have the evidence it needs in favor of the requirements it posits. See Vargas (2013a, 52-72)

[9] It does not follow that there is agreement about the definitional matter, what the folk notion of race is, and whether biologically useful phenomena map on to the folk notion (Spencer forthcoming; Spencer 2012).

[10] She writes that "nothing can kill compatibilist free will" (ms 10) and elsewhere she has written that "Compatibilism is not interested in how a behavior is caused—it simply states that, in the absence of external (and arguably also internal) compulsion, acts are said to be freely willed. In intellectual terms this is a relatively weak definition" (Pockett 2007, 292) and "Philosophical compatibilists define free will in such a way that science is irrelevant. They concentrate purely on whether or not there were constraints on a particular action or whether the actor was "free" to choose his own course. Constraints in this sense can be either external or internal." (284) These latter two characterizations of compatibilism would come as a great surprise to most self-identified compatibilists I know, even among the most enthusiastic defenders of the claim that "nothing can kill compatibilist free will."

[11] For a sampling, see Doris (forthcoming, 2002), Mele (2008, 2009, 2012), Levy (2011), Nahmias (2007, 2010), Nelkin (2005), Roskies (2010, 2008, 2006), and Vargas (2013b).

[12] Pocket claims that incompatibilism requires that "causal determinism [is] an illusion" but goes on to claim that libertarianism is distinguished from incompatibilist free will because it demands acts with "no physical antecedents" (ms 4). This way of characterizing things has the startling consequence that all event-causal libertarians and most physicalist agent causal libertarians are not libertarians.

[13] For reservations about the recent uptick in usage of this distinction, see van Inwagen (2008). For a reply, in which I deploy something like what Chalmers calls "the subscript gambit" see Vargas (2011b). See also Chalmers (2011, 532).

[14] For discussion of the difficulty of drawing substantive conclusions about various aspects of the free will problem from psychological research, see Doris and Murphy (2007), Nelkin (2005), Nahmias (2007) and Vargas (2013b)

[15] For prominent responses to the strands of Strawson's work that embrace free will nihilism, see Clarke (2005) and (Fischer 2006; Fischer 2006) One difficulty with a regression principle is that it commits us to responsibility being essentially impossible. However, there are considerations in the philosophy of language that weigh against any picture that holds that responsibility is "essentially impossible." Briefly: disagreements about the meaning and requirements of 'free will' and 'moral responsibility' cut against the plausibility of any armchair-derived proposal for responsibility being an essentially impossible property, and essentially impossible properties lack suitable "explanatory depth" for the involved practices, so that even if we had once referred to something impossible we would have likely shifted to some nearby property that actually obtains (Hurley 2000; Vargas 2011a, 466).

[16] He cites some work by Dennett and Searle. Dennett's work is obviously a reasonable place for a scientist to learn about philosophical work on free will, but the Searle text he cites has almost nothing to do with free will and moral responsibility.

[17] Substance dualism and agent causation are some of the more prominent approaches.

[18] To my knowledge, Eddy Nahmias and his collaborators were the first to consistently articulate this sort of concern in the contemporary free will debate (Nahmias et al. 2007).

[19] There are a number of candidates lurking out there. One could think that, for example, epiphenomenalism is free will's silver bullet, or that automaticity, or situationism present more powerful challenges. I agree that they are challenges. I also think they have already largely been met, but again, that the details require taking substantive positions on philosophical issues. See, among others, Doris (forthcoming), Mele (2008, 2009, 2012), Nahmias (2007, 2010), Nelkin (2005), and Vargas (2013b).

[20] The former has become increasingly plausible in light of experimental work (Nahmias 2011a; Sarkissian et al. 2010). The latter has been defended by Sommers (2011).

[21] One might deny that our metaphysical commitments and our epistemological pragmatics can yield distinct answers. That would be a most fortuitous discovery. In many ordinary cases, our commitments about the nature of some thing ("water is H_20") turns out to be rarely useful. In contrast, philosophically looser but practically efficient thoughts like "that wet, clear stuff that is refreshing to drink" tends to be relied upon a good deal, and a useful guide to finding water outside of bars.

[22] In context, it appears that the "mostly correct" claim about hard determinism doesn't reflect doubts about the existence of determinism but confidence that responsibility, understood in terms of

punishment, can persist.

[23] "We submit that [questions about free will and desert] which seem so important today, will lose their grip in an age when the mechanical nature of human decision-making is fully appreciated" (1781) and "We do not wish to imply that neuroscience will inevitably put us in a position to predict any given action based on a neurological examination. Rather, our suggestion is simply that neuroscience will eventually advance to the point at which the mechanistic nature of human decision-making is sufficiently apparent to undermine the force of dualist/libertarian intuitions" (1785, n. 6).

[24] Thanks to Eddy Nahmias for this point.

[25] As a matter of ordinary patterns of judgments about manipulation cases, it is not clear that the folk aren't compatibilists, at least sometimes. There is some evidence that even manipulated agents with intact reasoning mechanisms are sometimes judged to be responsible (Feltz 2013).

[26] Talk of morality is not intended to be supernatural. Understand morality as naturalistically as you like. However, if you aren't a naturalist about morality, it seems reasonable to assume that you don't share the familiar scientific worries about free will and moral responsibility.

[27] This is a point made in Nahmias (2011b) Nahmias argues for this in light of some interesting results detailed in Vohs and Schooler (2008) and Baumeister et al. (2009).

[28] I'm indebted to Eddy Nahmias for many conversations about these things over the years, and for feedback on this paper. His influence on my thinking about these matters is considerable. Thanks, too, to Gregg Caruso, Daniel Speak, and Stephanie Vargas for helpful feedback on an earlier draft of this paper.