

Revisionism

Manuel Vargas

Forthcoming in *Wiley Companion to Free Will*,
ed. Joseph Campbell, Kristin Mickelson, & V. Alan White

ABSTRACT: A revisionist theory of free holds that the correct account of free will is at odds with our ordinary understanding of free will's nature, concept, or associated practices. This chapter canvasses developments in revisionist theorizing about free will, including motivations, methodology, various forms of revisionism, and notable objections and replies to revisionism about free will.

1. Introduction

A theory of x is revisionist if the truth of the theory's account of x is in conflict with commonsense views about that thing. Those commonsense views that are repudiated by the revisionist theory may include views about x 's theoretical commitments, the nature and status of practices involving x , and the characteristic propositional attitudes and inferences involving x . In its canonical forms, revisionist theories are *preservationist*, i.e., accounts that seek to retain the term, concept, or practice for which revision is proposed. Revisionist theories of some phenomenon x typically contrast with *eliminativist* accounts, which seek to expunge talk, thinking, or practices of x . However, revisionist accounts also conflict with *conventional* accounts, where neither elimination nor revision is called for. Revisionist theories have been proposed for a wide range of things, including race, personal identity, propositional attitudes and various notions within the philosophy of science (for discussion, see Nichols 2015: 56-62 and Vargas 2013a: 73-75).

A revisionist theory of *free will* is one that holds that the correct account of free will is, in some or another way, at odds with our ordinary understanding of free will's nature, concept, or associated practices. One most versions, though, revisionism constitutes a rejection of eliminativism (Strawson 1986; Pereboom 2001; Levy 2011; Caruso 2015), in maintaining that we can and should preserve characteristic talk, judgments, and practices bound up with free will. Revisionist theories in a domain depart from conventional theories in maintaining that a philosophically satisfactory account of free will is in conflict with ordinary understandings, theories, or practices concerned with free will.

Typically, there is a *methodological* component to revisionism; the revisionist theory is the ensuing distinctive *substantive* results that come about from adopting the methodology. Methodologically, revisionism depends on distinguishing between what we—the folk, or competent language users, or ordinary practitioners—think about something from what, all things considered, we *ought* to think about that thing. Substantively, a theory is revisionist to the extent that the substance of the all-things-considered proposal conflicts with some aspect of folk or “commonsense” thinking. A theory is not revisionist if, for example, the theorist has mischaracterized folk views of x and offered an all-things-considered theory of x that conflicts with that mischaracterization. Such an account would not satisfy the requirement of being in conflict with actual folk views. Similarly, a view would fail to be revisionist if the prescriptive (or all-things-considered) proposal involved resolving or refining commitments within the constraints of existing indeterminacies in folk thinking. Such “revisions” are generally too trivial to count as the basis of a distinctive theoretical or methodological option.

The range of revisions can vary. Revisionism about conceptual commitments, for example, might not entail revisionism about existing social practices. Or, it might only entail revisions about those practices in very tightly circumscribed conditions. An account might be revisionist about the meaning of a concept, without being revisionist about how the concept is to be deployed in everyday usage. This sort of revisionism might insist that the account preserves paradigmatic inferences about free will, apart from those concerned with its nature.

Within a single axis of revision, there is room for more and less substantive revisions. A weakly revisionist theory might allow that some portion of the folk are in error about some peripheral conceptual matter, without those errors being part of the core commitments of some concept or practice. A theory might be weakly revisionist if it amounts to a correction of some common comparatively downstream confusion about what our foundational or core commitments in some domain really come to. Perhaps some defenses of the conditional analysis of 'can' were like this. In contrast, on a view according to which there is something conceptually foundational or central, important for the practice, reference-fixing, widespread, or normatively authoritative about the element the proposal revises away from, then the degree of revisionism will be correspondingly stronger. So, depending on what one is a revisionist about (e.g., concepts, attitudes, practices), the degree of one's revisionism, and the extent to which others regard the putatively erroneous element as more and less central, one's revisionism may be a more or less salient.

Revisionism about free will has been motivated by a variety of worries regarding putative elements of received commitments about the concept, meaning, or nature of free will and morally responsible agency. For example, revisionists have argued that folk conceptions of *agency* are metaphysically implausible (Doris 2015a; Nichols 2015), that ordinary conceptual and metaphysical commitments about *freedom* and *responsibility* are implausible (Vargas 2004; Vargas 2013a), that considerations about phenomenology (Deery 2015), reference (Heller 1996; McCormick 2013; McCormick 2015), and natural kinds (Deery, 2020) may favor a form of revisionism, and that eliminativist views neglect a more moderate and plausible alternative than the strong forms of agency demanded by some accounts of free will (Hurley 2000; Vargas 2005a).

The basis of doubts about received or pre-revised notions of agency, freedom, and responsibility have also varied. To make the argument that ordinary or folk views of free will are untenable, philosophers have drawn from various sources, including: experimental data concerning folk notions (Vargas 2013a; Nichols 2015); considerations from the history of ideas (Nichols 2007b) and culture and religion (Vargas 2013a; Vargas 2016); reflections on

phenomenology and language (Deery 2015; McCormick 2016); the presumptions of moral concepts (Smart 1961; Arneson 2003), and reflections on the fit between folk notions and findings in cognitive science and experimental science (Doris 2002; Nichols and Knobe 2007; Knobe and Doris 2010; Vargas 2013c; Weigel 2013; Doris 2015a; Nichols 2015, Faucher 2016, Vargas 2017).

Perhaps the most common form of revisionism about free will reflects a concern in the compatibility debate, claiming that the philosophically preferred view is one that conflicts with standing folk convictions about free will. One might allow that folk thinking has incompatibilist strands, in whole or in part, but prescribe compatibilism as a solution. To be sure, one can be a revisionist about free will without being committed to a specifically compatibilist revisionism. For example, a revisionist theorist might hold that everyday thinking about free will is committed to *causa sui* agency (Strawson 1986), but that our best philosophical theory supports a form of event-causal libertarianism as a suitable revisionist replacement. That said, it is unclear the extent to which some or any libertarians currently conceive of their accounts as revisionist (Cf. Vargas 2005a: 418, 424 n.8).

Revisionism about free will need not be grounded in aspects of the compatibility debate. For example, some accounts of free will maintain that the most appealing conception of it will not be sufficient to support retributive elements in our ordinary practices, and these theorists are often prepared to acknowledge that they are potentially revisionists on this point (Wallace 1994: 228; Scanlon 1998: 275; see also the discussion in Nelkin 2011: 175-177). Alternately, one might isolate grounds for revisionism in what we think about agency more generally. For example, one might hold that free and responsible agency exists but that it does not involve, for example, the psychological features we ordinarily take to ground it (Cf. Doris 2015a).

Despite its distinguished pedigree in the philosophy of mind, philosophy of race, and in the philosophy of science, it is only in recent years that revisionism has been regarded as a distinctive family of views in debates about free will and moral responsibility (Cf. McCormick 2016). Important precursors to self-described revisionist theories of free will include accounts by Smart (1961), Bennett (1980), Heller (1996), Singer (2002), Hurley (2000), and Arneson (2003). Among contemporary theorists, philosophers who have identified their proposals as revisionist in some or another way, including Walter (2004), Vargas (2004; 2005a; 2013), Nichols (2015), Weigel (2013), McCormick (2013; 2015), Doris (2015a). Others have explicitly (as we saw above with Wallace, Scanlon, and Nelkin) or tacitly (McGeer 2015) signaled an

openness to their accounts being construed (in part) as a form of revisionism about free will and/or moral responsibility.¹

In the context of debates about free will, revisionism's appeal seems to be grounded in its putative resources for (1) explaining disagreements about conflicting intuitions and the persistence of debates about free will, and (2) re-anchoring our understanding of the epistemic, metaphysical, and normative basis for free will (and/or moral responsibility). Neither (1) nor (2) are uniquely claimed by revisionist theories. However, those who are drawn to the more ambitious forms of revisionist accounts pursue them partly because they believe that they enable one to acknowledge some powerful aspects of everyday thinking (and thus, to remain in conversation with those gripped by these convictions) without thereby making their positive or prescriptive accounts of free will beholden to folk convictions. To be sure, whether the rehabilitation of common sense seems required, interesting, and/or promising depends a great deal on whether one is inclined to think there is something infelicitous about ordinary convictions about free will. Conventional compatibilists and libertarians tend not to.

In what follows, I will treat *free will* as a placeholder term, without committing to any particular view of how it is to be understood. However, depending on how one understands the various involved ideas, revisionism about free will may entail revisionism about moral responsibility, responsible agency, deliberation, and various notions of moral desert. Given this fact of the state of play in contemporary philosophical discussions, I will sometimes move between talk of revisionism about free will, and talk of revisionism about responsible agency and moral responsibility. This is not ended to reflect an endorsement of a substantive view about how to characterize free will. Rather, it is only an effort to capture the range of philosophical approaches that might plausibly be construed as revisionist about free will.²

¹ Dennett (1984; 2003) has an important but ambiguous relationship to revisionism. Although his account is associated with the slogan "the varieties of free will worth wanting" (the subtitle of his first book on free will), and perhaps popularly thought of as an early proponent of revisionism, the actual arguments and the views advanced in his accounts tend to overlook the differences in revisionism and conventional compatibilism. Moreover, when pressed, Dennett has tended to disavow the revisionist reading of his work. For a more detailed discussion of the difficulty of treating Dennett as a conventional revisionist, see Vargas (2005b).

² The nature of free will and its characterization is a contentious issue. Some are inclined to provide a core characterization in terms of moral responsibility (as in Fischer et al. 2007), some in terms of desert (Pereboom 2001), some in terms of deserving retribution (Caruso and Morris 2017). Whether these are mistakes is not something on which this discussion needs to take a stand. For some relevant discussions, though, see van Inwagen (2008), Balaguer (2010), Vargas (2011), and Franklin (2015).

The structure of the rest of this chapter is as follows: First, I discuss some of the methodological foundations of revisionist accounts. Then, I discuss various grounds that have been offered for revisionism. In a subsequent section, I consider some objections directed at recent revisionist developments. Finally, I conclude with thoughts on the direction of revisionist theorizing.

2. Methodological foundations

It is undoubtedly true that within the literature on free will, methodological approaches have varied considerably (Double 1996). However, it is also certainly true that a concern for ordinary convictions about freedom, and for grounding philosophical arguments in those convictions, has been a widespread feature of philosophical work on free will (Nahmias et al. 2006; Nichols 2006; Nichols 2007a; Nahmias 2011; Björnsson and Persson 2013). A robust philosophical preoccupation with “intuitiveness” and coherence with folk commitments in the context of free will might explain why interest in convictions about free will were an especially prominent feature of the first wave of experimental philosophy and empirically informed approaches to moral psychology (Nahmias 2006; Doris et al. 2007; Nahmias et al. 2007; Roskies and Nichols 2008).

A wide range of philosophical methodologies maintain that an adequate philosophical account of many, perhaps most, phenomena begins from the everyday meaning, usage, or thinking about the particular phenomenon in question. More ambitiously (or perhaps its reverse?) some have thought that the elucidation of some “ordinary” or “common sense” concepts *just is* the task of metaphysics, and sometimes, philosophy more generally (Strawson 1959; Lewis 1973; Jackson 1998). Others have thought that the proper task of philosophy should seek to explain how widely accepted ideas could be true (Nozick 1981). Coherence with folk commitments is a crucial, even constitutive desideratum of these philosophical methodologies. Given the history of debates about free will in the analytic tradition—including debates about the meaning of ‘can’ in the 1950s and 1960s, and the metaphysics of ability in the 1960s and 1970s (Vargas 2011)—it is perhaps unsurprising that a good deal of the literature has been structured by sometimes explicit but oftentimes tacit presumptions about the importance of folk thinking about free will. Against this background, revisionism represents a methodological break in the consensus about the role of folk commitments for philosophical theory-building.

Revisionist methodologies usually start by distinguishing at least two questions. First, what is the status of commonsense thinking about the thing in question? Call this the *diagnostic question*. Second, what sort of theory should we have about that thing, all things considered? Call this the *prescriptive question*. The disjunctive nature of the diagnostic question—what is

the nature and plausibility of ordinary commitments about *x*?—can be further separated into two component questions. Shaun Nichols (2015) helpfully distinguishes between a *descriptive* question and *substantive* question within the broader diagnostic question. The descriptive question concerns our (typically folk, but potentially philosophical) beliefs in some domain. The substantive question concerns the status of those beliefs—as true, false, implausible, and so on. What is helpful about subdividing the diagnostic question into two further questions, descriptive and substantive, is that it is one thing to identify folk commitments, it is another to decide that they are implausible or in need of revision. In practice, revisionist accounts tend to proceed by first identifying some problematic aspect of common sense and/or folk notions (of freedom, race, propositional attitudes, etc.), and then moving to a positive proposal for theorists (and perhaps the folk) to go forward without the problematic aspect.³

At this point, some might be inclined to protest that the purported distinctiveness of revisionist theorizing is hardly distinctive at all. Many philosophers accept something like the method of reflective equilibrium, on which theoretical revisions of considered beliefs and principles is permissible. Even on conceptual analysis-style accounts of philosophical methodology, revisions that depart from aspects of common sense or folk beliefs can be a relatively pedestrian feature of conceptual analysis, where the stakes can be something “suitable close to our ordinary conception” (Jackson 1998: 31). Moreover, relatively radical revisions can be licensed under particular kinds of conditions. For example, some are prepared to allow for the possibility of infinite concrete possible worlds on the grounds that this sort of counterintuitiveness enables a reduction in the number of basic entities posited by the theory (Lewis 1973: 87). Thus, isn’t revisionism just fussiness about something that is already methodologically a commonplace?

When the alternative is confusion, fussiness can be a necessary virtue. Revisionist accounts are often motivated by the thought that a failure to be cautious about whether and when one is intending to revise away from problematic folk notions has led to cross-talk and methodological confusions about what accounts of free will are endeavoring to provide (Vargas 2011; Nichols 2015). The thought seems to be this: in domains where folk beliefs, convictions, or practices are taken to have special weight in constructing a theory of that domain (e.g., moral and social categories), revisionist theories face special burdens licensing departures from folk thinking. Thus, it can be particularly useful to identify that the theory is revisionist, because it will have

³ Compare McCormick’s formulation of revisionism: “Revisionism is the view that we can and should distinguish between what we think about moral responsibility and what we ought to think about it, that the former is in some important sense implausible and conflicts with the latter, and so we should revise our concept accordingly” (2016: 109).

those special burdens. Relatedly, there are certain intuitions or judgments that the theory may not be in the business of trying to preserve or capture. So, marking the theory as revisionist is a way of being cautious about what is being revised and what is not thought to be in need of revision.

In short, failure to mark such differences is less a matter of declining to be fussy than it is a failure of intellectual hygiene. To leave the revisions unmarked runs the risk of presenting revisionist proposals as descriptions of folk commitments, and folk commitments as (tacitly and adequately) grounded prescriptions when they are not. (What is “suitably close to our ordinary conception” to count as our ordinary conception, anyway?) Conceptual analysis, philosophical explanation, and the method of reflective equilibrium too often fail to mark out what is putatively a part of everyday thinking and what is the product of revisionary philosophical labor. In failing to mark such distinctions, we risk running together ordinary convictions with philosophical innovation, and we leave ourselves open to unnoticed changes in topic. Distinguishing revisionist and conventional proposals clarifies the stakes and the options.

None of this is to deny that there are phenomena where conflicts with folk commitments are rightly regarded as an entirely unremarkable matter of course. In such cases—contemporary physics is a particularly clear example—there is little reason to identify the account as revisionist. However, in domains where accounts are not always understood to be revisionist, and where there is a heavy reliance on folk commitments in the construction of theories, there is special reason to be explicit about whether and when one is being revisionist. Thus, although revisionist theorizing can be found in a wide range of theoretical domains, its salience, degree, and significance varies by the context of our theorizing about those things.

One can be a revisionist about the meaning of terms, about the metaphysics associated with a term, about the concept, about judgments and inferences connected to the term, about practices associated with it, and even about emotional attitudes characteristic of the term and its associated practices. So, it would be useful to have a neutral term covering all these cases. Unless there is some matter where more precision is useful, in what follows I will generally speak of revisionism about commitments, intending this to cover the panoply of more particular varieties of revisionism. I will also put to one side some potential complexities concerning unintentional revisions, or revisions that are products of inconsistency in theoretical design. The general form of revisionism that is of interest in what follows is intentional revisionism away from some plausible understanding of current practices, beliefs, dispositions of judgment and attitudes—i.e., our commitments.

Some early accounts of free will revisionism (e.g., Vargas 2005a) failed to distinguish what we might call the *existence question* from the prescriptive question. The existence question concerns whether or not the considered property exists, or is instantiated. Different answers to the existence question are compatible with a revisionist prescriptive proposal. To see why, consider that one might hold that the best prescriptive account of *x* is revisionist, but that even after revision, that we should conclude that it does not exist.

The idea of true love might serve as an example. Suppose Juan correctly believes that true love is widely regarded as a monogamous tie between two souls that are perfectly fitted to one another. However, his personal view might be revisionist, a view according to which true love is compatible with both the non-existence of souls and non-monogamous romantic relationships. For example, Juan might accept a threshold theory of love, according to which crossing some threshold of mutual fulfillment and affection some distance short of perfection constitutes true love. Given the foregoing, it would be available to him to discover that he is a polyamorist materialist who has more than one true love. He might accept this revisionist theory for years without difficulty. However, suppose that Juan becomes a divorce lawyer and after a decade of practice he becomes jaded about the possibility of true love. If so, he might eventually conclude that true love, even on his revisionist account of it, does not exist. Perhaps some further fact—something discovered in the accumulating emotional debris of divorce proceedings—leads him to think that no one ever attains the threshold of mutual fulfillment and affection he took to be (his revised) notion of true love. So, Juan might be a revisionist about true love even while accepting that true love does not exist.

To be sure, most prescriptive proposals are offered on the presumption that the target property exists. In many contexts, there would be little that is theoretically appealing about a proposal to modify our thinking or practices concerned with *X*, if one did not think the proposed modification succeeded in tracking a viable (that is, existing or potentially existing) property in the world.

3. Arguments for revisionism

In this section, I canvas some of the arguments for revisionism that have appeared in the recent philosophical literature. These include arguments from two-dimensional semantics, a debunking theory of the psychological mechanisms involved in incompatibilist convictions, doubts about the scientific plausibility of ordinary conceptions of agency, and various combinations of these views.

One motivation for incompatibilist pictures of free will is the idea that ordinary phenomenal experience has a libertarian character. That is, in ordinary deliberation about what to do, the phenomenal character of such deliberation is best understood as having libertarian commitments. Some have gone on to argue that this phenomenal character is reference-fixing (Caruso 2015). However, drawing from work by David Chalmers on two-dimensional semantics, Deery (2015) argues that phenomenal experience may well be libertarian, but that it also comes with “imperfect content” that provides veridicality conditions that appeal to whatever it is that normally causes those libertarian experiences. The punch line is that there is good reason to think *those* properties—i.e., those properties that ordinarily cause libertarian experiences—are in fact compatibilist-friendly properties, unthreatened by determinism. The result is a kind of error theory about folk phenomenology—it presents as libertarian in a way that cannot be sustained. However, it is also a kind of success theory about the properties that produce the error-ridden phenomenology. Deery goes on to recommend that our theories track this latter property, even though it conflicts with the evident phenomenology of choice-making.

Nichols (2015) maintains that the problem of free will arises because of some peculiarities of the psychology of explanation in the context of human action. On Nichols’ account, we have “an explanatory compulsion,” according to which we believe every action or event must have a cause. This compulsion to explanation lends itself to seeing human action as a causal (deterministic, in particular) product of prior events. However, in the case of human action the explanation yields a distinctive result. In first-personal cases of action explanation, when I consider why I acted the way I did, there is no immediately identifiable causal origin for my action. Nichols holds that when we cannot identify the causal sources of our own actions, we take this as evidence that our actions lack deterministic causes (42-50). The result is the genesis of the free will problem, namely, a conflict between our normal deterministic picture of explanation and a robust sense of our own action being indeterministic, i.e., the “indeterministic intuition.” However, our introspection is a flawed guide to the causal factors involved in our decision making, and thus, “our intuition of indeterminism counts for nothing” (53).

Vargas (2013a) argues that folk commitments about free will and moral responsibility are not compatibilist nor incompatibilist in a uniform way. Instead, there is oftentimes diversity across and within individuals about these commitments. Moreover, various rational, deliberative, and cognitive pressures regiment or systematize those commitments in ways that support compatibilist and incompatibilist convictions in the philosophical domain (2013a: 49-51). However, the persistence of the free will problem, as a philosophical problem readily available to ordinary reflection, arises in large part because of the pervasiveness of specifically

libertarian-friendly commitments (alternative possibilities, and perhaps a kind of sourcehood) that are supported by distinctive psychological and cultural scaffolding (2013a: 40-43; 2016). These commitments are not necessarily had by everyone. Nevertheless, various aspects of human deliberation, explanation, and cultural narratives make those commitments appealing to a wide range of people. Evidence of the widespread presence of those libertarian commitments (about the basis of responsibility, about the nature of agency, and about the phenomenology of deliberation) can be found in experimental work (Nichols and Knobe 2007; Weigel 2013), broadly sociological data (Collins 1998; Nichols 2007b; Bourget and Chalmers 2014; Vargas 2016), and in the readiness with which people understand talk of abilities in the context of philosophical arguments for incompatibilism.

Gunnar Björnsson and Karl Persson (2012, 2013) have developed an account of the explanatory framework of human action according to which the widespread presence of incompatibilist commitments in the folk, at least with respect to talk about responsibility, is a byproduct of shifts between explanatory frameworks. As they see it, the ordinary concept of moral responsibility is bound up with our interests in identifying relevant motivational structures that play a suitable systematic role in action explanation. Thus, the ordinary concept of responsibility has been “shaped by conscious and unconscious interests and concerns that govern our practices of holding people responsible” (2013: 328). However, broadly incompatibilist regress arguments (e.g., Strawson 1994, Pereboom 2001: 112-117) and appeals to determinism get their force from shifting attention away from the explanatory concerns that are salient for responsibility attributions.

Björnsson and Persson (2012) outline different possible readings of their account, including interpretations on which it is an account of the contextualist semantics of responsibility, an account on which it yields a story about our judgments of responsibility but not the (realist) nature of responsibility, and an account of some dispositions structuring expressivist views of responsibility. As they note, issues here are delicate. A variety of philosophical commitments bear on how we should understand the upshots of their approach, including: whether we have an independent account of the reference of moral terms, or moral responsibility in particular; accounts of the function of the involved concepts and practices of holding responsible; and the normative preconditions for holding people responsible (346-348).

Perhaps the most natural reading of Björnsson and Persson’s model is that it suggests an error theory for incompatibilist convictions, and correspondingly, a picture according to which the referent of talk about responsibility picks out some property compatible with a deterministic picture of the world. On their account, humans are prone to adopting faulty incompatibilist

conclusions, and that absent some special remediation for a natural cognitive tendency, there will always be people who erroneously come to accept incompatibilist commitments about moral responsibility.

If that is how it is read, Bjornsson and Persson's picture suggests a form of connotational revisionism (Vargas 2013a: 88), according to which terms like 'free will' and 'moral responsibility' successfully refer, but for which we (pre-revision) had a variety of false beliefs about the nature of the thing referred to. However, the Bjornsson and Persson model also admits of another possibility, according to which the tension in explanatory frames fails to provide a determinate referent for talk of moral responsibility. On this account, the picture might favor a kind of denotational revision, that is, where there is a change in referent.⁴ Bjornsson and Persson take no stand on these matters, but their proposal suggests a powerful and principled basis for adopting revisionism about moral responsibility.

In recent work, John Doris (2015a) has adopted a revisionist account of responsible agency, but on grounds disconnected from the traditional compatibility debate.⁵ Doris holds that ordinary thought and many strands of philosophical theorizing about agency are structured by a background commitment to view that the exercise of human agency consists in judgment and behavior ordered by self-conscious (and accurate) reflection about what to think and do. However, a substantial body of experimental considerations suggest that the core of this view—something Doris calls *reflectivism*—is false. He holds that our responsibility practices can and should be retained—they are “felicitous,” as he says—but that an accurate account of their basis involves a radically different understanding of the nature of human agency and the way in which responsibility practice interacts with responsible agency.

Although Doris is a kind of pluralist about responsible agency—he allows it may be attained in a variety of ways—he also argues that there is a relatively radical alternative conception of our agency that we should accept, something he characterizes as *collaborativism*. Collaborativism is

⁴ Normally, denotational revision is characterized by the idea that there is some nearby property that is a suitable basis for various practices, judgments, and attitudes that are impugned by one's diagnostic account, and the revisionist proposal is a recommendation to shift reference to explicitly track that nearby property. So, this would be a special case of reference shift, from an indeterminate referent to a determinate referent.

⁵ Doris disassociates free will from moral responsibility (2015a: 9-11), and says little about free will specifically. So, his account does not obviously entail revisionism about free will. Given what he says about reflectivism and agency, however, it is not difficult to imagine recasting most of his arguments against folk conceptions of responsibility as arguments against folk conceptions of free will.

the idea that optimal human reasoning is facilitated by social interaction (103). As Doris sees it, the process of reasons-giving, and various intra- and interpersonal explanations of action and statements of values function less as an accurate statement of the springs of action than as a kind of social scaffolding that serves to coordinate future action and publicly establish commitments to particular standards of action evaluation. This is not intended as a characterization of how we understand ourselves, but rather, a characterization of how we ought to understand ourselves, given an empirically informed conception of our agency. Doris characterizes his view as conservative about the practice of moral responsibility but revisionary about the theoretical underpinnings (158). However, his revisionism is tempered by the thought that it is exceedingly difficult to demarcate folk theories about agency (i.e., whether and to what degree an account is revisionist may be a fundamentally indeterminate matter) (Cf. Doris 2015b).

4. Doubts about revisionist approaches

Most explicitly revisionist accounts of free will have proposed prescriptive theories of free will that are compatible with the truth of the thesis of determinism. Moreover, one can convert nearly any conventional compatibilist into a revisionist account simply by acknowledging that folk thinking has incompatibilist commitments, and that one's compatibilist proposal is intended as a correction of defective folk thinking. So, one might wonder, is this evidence that revisionism is simply compatibilism by a different name?

In reply: no. As noted at the outset, what makes a theory revisionist is that it is in conflict with commonsense commitments. Revisionism is silent on the content of those commitments, and one can be a revisionist and offer an incompatibilist prescriptive theory. So, revisionism is not just compatibilism. It is true that conventional compatibilist theories of free will can be readily converted to revisionist proposals. However, such proposals come with distinctive costs. For example, the pressure to identify the grounds for making the revision becomes unavoidable for any newly revisionist theory, as does the explanation of why the proposed revision does not amount to a changing of the topic. If the aspirations of a theory change, what the theory has to show changes.

If revisionism is not simply a form of compatibilism, one might charge that revisionists are nevertheless incompatibilists. That is, even if revisionism does not by itself entail incompatibilism, given that most contemporary revisionists hold that ordinary convictions are in some or another significant way contain incompatibilist elements, then it seems that they are incompatibilists by another name. Whatever the appeal of this objection may be, it relies on a

controversial picture of the relationship between folk beliefs and reference, one that a number of revisionists have sought to contest. Whether a given revisionist proposal is, in the end, incompatibilist depends on some delicate issues to which we will return, below, when we consider connotational and denotational forms of revisionism.

A third possibility lurks. A critic might insist that *everyone* is, to some extent, a revisionist.⁶ All philosophical theories depart to some degree, in some way, from the particulars of folk thinking. In reply: we need to be careful to track the distinctive nature of revisionist theories. What makes a theory revisionist in the sense adumbrated here is not merely the fact of a departure from commonsense. Instead, it is the fact that the truth of the proposal is in *conflict* with commonsense. To be sure, conflict can come in degrees, so theories can be more and less revisionist. We have also seen that the target of revision can vary. For example, as noted above, one might be revisionist about our understanding of the nature of free will without being a revisionist about particular practices that implicate free will.

Whether any revisions are interesting or philosophically salient partly depends on the background presumptions about how common conflicts with commonsense are expected to be in philosophical theorizing. Within normative ethics and philosophy of physics, such conflicts are a matter of course. So, revisionism is less notable. Within philosophy of mind, the gradual retreat from eliminativism to revisionism about folk psychological terms took longer, despite the widespread influence of Quinean holism. However, within contemporary debates about free will—that is, a literature still heavily structured by the methods and presumptions of 1960s and 1970s analytic metaphysics—revisionism remains a minority position.

The distinctiveness of revisionism is more apparent if one considers the history of the subfield. For example, it would be grossly uncharitable to interpret most of the historical proponents of the conditional analysis as aspiring to reform meaning. Similarly, to assert that most current compatibilist accounts are revisionist in the sense under consideration would also be to say that most contemporary compatibilist accounts are entirely unresponsive to the burdens of motivating their revisions, of explaining why they aren't changes in topic, and in acknowledging that folk thinking is indeed incompatibilist in significant measure. These would be serious lacunae in contemporary compatibilist accounts. The fact of their absence in virtually all

⁶ This objection was rightly raised by Michael McKenna (2009) against previous formulations of revisionism (e.g., Vargas 2005a) where revisionism was wrongly characterized in terms of any departure from commonsense commitments.

compatibilist accounts that do not explicitly identify as revisionist would seem to justify incompatibilist complaints that compatibilism is a wretched subterfuge.

It is more plausible to hold that revisionist compatibilists really are up to something different than traditional compatibilists.

5. Recent issues in revisionist theorizing

Two recent debates about the forms revisionist theories can take—a matter of the conceptual space available to revisionists—merit some discussion. One is Nichols' argument for the possibility of a view he calls *discretionism* and a second is McCormick's argument for restricting revisionism to denotational revisionism, or a kind of *replacementism*. Nichols' approach might be thought of as an attempt to expand the scope of options available to revisionists whereas McCormick has argued for a narrower range of options available to the revisionist.

Nichols' version of revisionism about free will and moral responsibility is subtle and innovative. As we saw above, he argues that folk thinking has a robust and readily identifiable commitment to incompatibilist conceptions of agency. However, he maintains that for moral and pragmatic reasons, the attitudes and practices we associate with responsibility (including moral anger and retributive impulses) should be retained. What makes the account revisionist—as opposed to fictionalist or illusionist (as in Smilansky 2000)—is that he thinks that our incompatibilist intuitions can be severed from those practices, so that our practices retain adequate normative standing. His is a kind of conceptual revisionism with little or no revisionism about the practice.

One innovation of Nichols' account is that he endorses a kind of discretionism, a view according to which variability in the reference conventions of the target term permit one to selectively adopt eliminativist discourse in one moment, and non-eliminativist discourse in another. The underlying idea is that some terms exhibit distinct reference conventions (e.g., having both a causal-historical reference convention *and* a distinct descriptivist convention). Nichols thinks 'free will' and 'moral responsibility' are such terms (2015: 166). What makes the account discretionist is that one is licensed or permitted to select among these conventions at will—at least in some contexts.⁷ This referential flexibility is appealing, Nichols thinks, because it affords us particular psychological advantages. During one's long dark nights of the soul, discretionism allows us to remind ourselves that we are not free agents of the sort that figure in libertarian

⁷ Were there a uniformly privileged reference convention, it is less clear that discretionism is permissible in the same way.

accounts of free will, and thus, that we cannot be responsible for our failings in the way characterized by those accounts. In daily life, however, it allows us to rightly insist that people can be blameworthy and that they can deserve our condemnation.

Whether discretionism is a stable view remains a matter of some dispute (see McCormick 2017; Vargas 2017; for a reply see Nichols 2017). Here is one concern: even if there are multiple (non-privileged) reference conventions for 'free will' and 'moral responsibility', it isn't clear that referential discretionism gets us anything different than a non-discretionist form of revisionism. After all, the non-discretionist revisionist can already say (in chorus with the discretionist) that one is not free in the way specified by the incompatibilist—i.e., that one does not have the properties that the incompatibilist says are required for free will and moral responsibility. So, the discretionist is not thinking a new thought on those long, dark nights of the soul when she concludes that she is not free and responsible as specified by a given incompatibilist account. Moreover, the sense in which the discretionist is putatively joining the eliminativist in thinking that one is not morally responsible cannot be the same sense in which an eliminativism worth the name makes that claim. On pain of there being insufficient reason to eliminate the practice and discourse of freedom and responsibility (and thus, to be a distinctive view), eliminativism must deny that our current freedom and responsibility-characteristic attitudes and practices are in good normative standing. Nichols' discretionist, however, does not (and cannot) think *that*. The entire point of Nichols' positive account just is a defense of integrity and value of those attitudes and practices. So, the virtues of discretionism as a distinctive form of revisionism, or perhaps as an alternative to familiar forms of revisionism, remain unclear.

In a series of recent papers, Kelly McCormick (2013; 2015; 2016) has deepened our understanding of the resources and challenges facing revisionist accounts of free will and moral responsibility, and developed new resources for responding to a range of concerns about revisionist approaches to theorizing. If she is right, connotational revisionism is unstable. On her account, the only viable form of revisionism must involve the relatively ambitious project of re-anchoring the referent of 'free will' and 'moral responsibility' in some nearby property.⁸

⁸ To recap: the difference between denotational and connotational revisionism is the difference between revisionist accounts according to which error-ridden folk commitments can be jettisoned without affecting reference (connotational revisionism) and revisionist accounts according to which error-ridden folk commitments *cannot* be jettisoned without some kind of reference change (denotational revisionism). Put differently, denotational revisionism concedes that the target term has no extension, but goes on to recommend a reference shift. Connotational revisionism holds that the target term has an extension, but that we have had false beliefs about the extension.

McCormick argues that connotational revisionism faces a kind of dilemma: either the to-be-revised-away content is widespread (and thus, it is implausible to insist that it is neither constitutive nor essential) or it is insufficiently widespread to ground a substantive conflict between the diagnostic and prescriptive elements of the account (2016: 117). For this reason, she favors a denotational analysis. However, denotational revisionism faces challenges of its own. First, one might doubt that denotational revisionism is distinct from eliminativism. Second, there is some reason to be skeptical whether connotational revisionism will go quietly into the night.

Is denotational revisionism really eliminativism in all but name? Notice that the denotational revisionist must grant that the pre-revised term lacks an extension. That fact, one might think, suffices for showing that any denotational revisionism is eliminativist. If revisionism just is restricted to denotational revisionism, then revisionism collapses into eliminativism, and there is nothing left that could count as a revisionist theory of free will.

McCormick allows that there is a sense in which denotational revisionism is indeed a kind of eliminativism. This result is misleading, however, because there remains a substantive difference between the position of the eliminativist and the denotational revisionist: the denotational revisionist is ultimately a kind of success-theorist, holding that there is reason to re-anchor talk in a nearby property (for example, one that does not involve powers invoked by libertarian accounts). Unlike the eliminativist, when we implement the revisionist's proposal, we *will* successfully refer. Moreover, our referring to this new property will preserve the characteristic structure of practices, attitudes, and judgments at stake in our concern for whether we are free (or responsible). So, a successful denotational revisionism might just as well be labeled *replacementism*, she notes, thereby highlighting that the term and its linguistic and social functions remain intact in light of identifying a suitable alternative property to anchor them.

Here's one way to see the point: there is a big difference between patching a boat with new parts, and dis-assembling the boat because one no longer regards it as seaworthy. As long as the characteristic function of the term is preserved, patching the function with a different referent remains a reasonable strategy. Thus, so long as we retain the plausible social or psychological functions of terms like 'free will' and 'moral responsibility', there will be a substantive difference between revisionist and eliminativist proposals.

As a matter of the sociology of those who self-describe as eliminativists and revisionists, McCormick's observation about the aspirational differences between eliminativism and

revisionism is surely right. Yet, a puzzle lurks. Consider the point made at the outset of this chapter, about how one might be a revisionist about true love, yet ultimately conclude that true love does not exist.⁹ A similar situation seems possible for free will. A theorist might appeal to an adequate replacement property to ground a proposed reference shift. This would suffice to make the theorist a denotational revisionist. However, that theorist might later conclude that there is a further reason for thinking that the new target property also fails to obtain. If no further revision seems viable, eliminativism is the order of the day. Thus, it seems, the difference between revisionism and eliminativism, if there is one, is not the difference between a success theory and an error theory.

Suppose McCormick grants that eliminativism and revisionism can, in fact, overlap in the particular way exemplified in the true love case. She would still be right that the ambitions of the proposals remain distinct, even in the rare limit case where a theory is both eliminativist and revisionist. Even in the true love case, the revisionist undertakes her revision with the thought that it is promising as a success theory; this is not true of the eliminativist. Even if eliminativism ends up being the right account, this is a further discovery. Despite the fact that revisionist theories are picked out by the conflict between folk and all-things-considered theorizing—and although they are ordinarily undertaken with an aspiration of providing a success theory, and the best accounts will give some reason for thinking things are as the account prescribes—whether reality cooperates remains a further thing.

A second family of challenges for McCormick's account concerns whether the argument against connotational revisionism succeeds. Here is one reason for doubt. Suppose 'free will' is a natural kind (Heller 1996; Deery 2020)—or at least a *genuine kind* (see Spencer 2012). On such an account, free will refers to property not fixed by a description but instead by a suitable set of externalist conditions—e.g., a causal chain, or a homeostatic cluster that underpins a function. If that is the sort of term 'free will' is, then discourse about free will can tolerate considerable and widespread errors about the nature of free will and still refer. Similarly, talk and beliefs about water can refer despite beliefs that it is, for example, one of the four basic elements of the world. To be sure, this line of response depends on moral responsibility being a particular kind of thing. At the same time, it does seem a kind of cost to McCormick's argument against connotational revisionism that it requires that free will is not a natural or genuine kind.

⁹ Here, as there, we can ignore a complexity about the difference between a shared concept, term, or practice and whatever the individual correlates of these things may be. Revisionism of the sort discussed here concerns proposals about the former (collective) versions of these things, so any example about an individual case will be, to that extent, modestly misleading.

Is connotational revisionism possible without free will being a genuine kind, or being picked out by a causal-historical chain? There are a variety of descriptivist strategies for accommodating false beliefs in such a way that reference is retained even in the face of some amount of false beliefs about the nature of the thing in question (Cf. Lewis 1983). On these accounts, so long as enough of the description of free will (or moral responsibility, or . . .) is true, or there is some property that renders enough of the platitudes true, jettisoning some false beliefs need not disrupt reference. If reference can be preserved, even in the face of the falsity of some platitudes under descriptivist reference conventions, then there is little reason to think widespread commitments, cannot also be jettisoned without loss of the referent. In sum, the possibility of connotational revisionism does seem not turn on whether free will is a genuine kind, referred to by a causal-historical chain (Cf. Vargas 2013a: 91-6).

6. Further developments

Perhaps the most visible unresolved issues facing revisionist theories concerns the role of revisionism about free will in connection with some normative ideas. This concluding section briefly canvasses three issues where revisionist work has interesting implications for recent debates. They include the relationship of free will to moral responsibility, disputes about desert, and the basis on which one might decide between revisionism and eliminativism.

As noted at the outset, a fair number of self-described revisionist accounts seem to be focused on issues that are readily construed as centrally about moral responsibility, rather than free will. Given the varied conceptions of what kind of thing free will is, how it is to be characterized, and the diverse roles played by the term in popular and technical discourse (Vargas 2011: 152-155; Vargas 2013b: 179-181), one can imagine that some theorists will be tempted to be eliminativist about free will and revisionist about responsibility.

Within the literature on moral responsibility, desert has become an increasingly important locus of disagreement. This was an issue noted by Smart (1961), Wallace (1994: 228), and Scanlon (1998: 275), but only recently has it become a matter of sustained discussion (see Caruso and Morris 2017; McKenna 2009; McKenna 2012: 114-120; Roskies and Malle 2013; Nahmias 2014: 55 n.3; Doris 2015b). In particular, for theories that seek to ground the normative basis of responsibility in broadly consequentialist justifications (theories that in recent years have tended to be revisionist), it is an open question whether such accounts have the resources to capture our ostensibly commonsense concern for desert, and if not, whether

desert itself admits of a revisionist reconstruction (see McGeer 2012; McGeer 2015; Vargas 2015, as well as McCormick in this volume).

The fate of revisionist theories is partly tied to one's conception of the aspirations for philosophical theories of free will and moral responsibility. Roughly speaking, there seem to be two competing impulses: (1) the desire to rescue our default conceptual or theoretical commitments and (2) the aspiration to ground the integrity of practices, attitudes, or other relatively concrete phenomena. For many eliminativists, the impossibility of vindicating ordinary theoretical commitments is the basis for the eliminativism and the basis for rejecting the possibility of grounding our practices. In contrast, for revisionists, the significance of rescuing our naïve theoretical commitments pales in comparison to the significance of grounding our practices. By such lights, the eliminativist's concern for vindicating folk conceptions of agency, free will, desert, and moral responsibility is like a hungry person's fascination with accurate portraits of food. It is an explicable interest, but the wrong preoccupation when there is real food readily available to eat.

Even so, eliminativism may yet retain some allure. Suppose that we conclude that primary philosophical stakes are practical and not theoretical—that is, they concern some subset of identifiable judgments, practices, and attitudes, and *not* the contents of our naive theories about the foundations of these things. Even if we decide that revisionists can provide an adequate ground for those practices, we might still ask whether, all things considered, we do well to keep these things at all. Might we be better off without the responsibility-characteristic practices, attitudes, and judgments? Eliminativism might yet be the better alternative.

One reply is the one given by P.F. Strawson (1962): it involves holding that the underlying attitudes implicated in responsibility practices are simply implastic pieces of our psychological architecture. Many eliminativists, and at least some revisionists, are inclined to regard this sort of claim with suspicion. There is a live issue here about how to measure the benefits of free will and responsibility practices, attitudes, and judgments, and what the metric should be for deciding whether we do better to keep or eliminate them. Still, one can hold on to basic Strawsonian impulse, and adopt revisionism (Vargas 2004).

Eliminativists (like Pereboom and Caruso) will extol the putative benefits of a world without free will and the attendant responsibility practices. Revisionists who think there are important normative and social payoffs of retaining these things (e.g., McGeer, Nichols, Vargas) will insist that the benefits of systematically foregoing responsibility have been overstated, and indeed, that the purported benefits of eliminativism are at best local, and often byproducts of free-

riding on social practices structured by moral responsibility. Gandhi's and Martin Luther King's apparent lack of resentment, indignation, and retributivism are undoubtedly appealing and this fact may seem to favor the psychological picture advocated by eliminativists. To be sure, pacificism and diminished reactivity may seem particularly appealing when the alternative is violent, angry, and vindictive. Yet a revisionist reply remains available: the current appeal to us of not blaming, and the appeal of non-blaming exemplars, may depend on us tacitly holding fixed in our imagination the goods secured by blame-propelled practices. If so, it is unclear whether our admiration for non-blaming can indeed be sustained in a systematic way if we are clear on what goods are lost with the jettisoning of blame. If that is right, then eliminativism's moral appeal is cashing a check that draws on blame's account.¹⁰

¹⁰ My thanks to Gregg Caruso, Kelly McCormick, Stephanie Vargas, and the editors of this volume for feedback on prior drafts of this essay.

References

- Arneson, Richard J. 2003. "The Smart Theory of Moral Responsibility and Desert." Pp. 233-258 in *Desert and Justice*, edited by Serena Olsaretti. Oxford: Oxford.
- Balaguer, Mark. 2010. *Free Will as an Open Scientific Problem*. Cambridge, MA: MIT Press.
- Bennett, Jonathan. 1980. "Accountability." in *Philosophical Subjects*, edited by Zak Van Straaten. New York: Clarendon.
- Björnsson, Gunnar and Karl Persson. 2012. "The Explanatory Component of Moral Responsibility." *Nous*. 46 (2):326-354.
- . 2013. "A Unified Empirical Account of Responsibility Judgments." *Philosophy and Phenomenological Research*. 87 (3):611-639.
- Bourget, David and David J. Chalmers. 2014. "What do philosophers believe?" *Philosophical Studies*. 170 (3):465-500.
- Caruso, Gregg. 2015. "Free Will Eliminativism: Reference, Error, and Phenomenology." *Philosophical Studies*. 172 (10):2823-2833.
- Caruso, Gregg D. and Stephen G. Morris. 2017. "Compatibilism and Retributive Desert Moral Responsibility: On What is of Central Philosophical and Practical Importance." *Erkenntnis*. 82(4):837-855.
- Collins, Randall. 1998. *Sociology of Philosophies*. Cambridge, MA: Belknap.
- Deery, Oisín. 2015. "The Fall from Eden: Why Libertarianism Isn't Justified by Experience." *Australasian Journal of Philosophy*. 93 (2):319-334.
- Deery, Oisín. 2020. *Naturally Free Action*. Oxford, U.K.: Oxford University Press.
- Doris, John. 2002. *Lack of Character*. New York: Cambridge University Press.
- . 2015a. *Talking to Our Selves: Reflection, Ignorance, and Agency*. New York: Oxford University Press.
- . 2015b. "Doing Without (Arguing about) Desert." *Philosophical Studies*. 172 (10):2625-2634.
- Doris, John, Joshua Knobe, and Robert L. Woolfolk. 2007. "Variantism about Moral Responsibility." *Philosophical Perspectives*. 21 (1):183-121.
- Double, Richard. 1996. *Metaphilosophy and free will*. New York: Oxford University Press.
- Faucher, Luc. 2016. "Revisionism and Moral Responsibility for Implicit Attitudes." Pp. 115-145 in *Implicit Bias and Philosophy, Volume 2: Moral Responsibility, Structural Injustice, and Ethics*, edited by Michael Brownstein and Jennifer Saul.
- Fischer, John Martin, Robert Kane, Derk Pereboom, and Manuel Vargas. 2007. *Four views on free will*. Malden, MA: Wiley-Blackwell.
- Franklin, Chris. 2015. "Everyone Thinks That an Ability to do Otherwise is Necessary for Free Will and Moral Responsibility." *Philosophical Studies*. 172 (8):2091-2107.

- Heller, Mark. 1996. "The Mad Scientist Meets the Robot Cats: Compatibilism, Kinds, and Counterexamples." *Philosophy and Phenomenological Research*. 56:333-337.
- Hurley, Susan. 2000. "Is Responsibility Essentially Impossible?" *Philosophical Studies*. 99:229-268.
- Jackson, Frank. 1998. *From Metaphysics to Ethics: A Defense of Conceptual Analysis*. New York: Oxford University Press.
- Knobe, Joshua and John Doris. 2010. "Strawsonian Variations: Folk Morality and the Search for a Unified Theory." in *The Handbook of Moral Psychology*, edited by John Doris and The Moral Psychology Research Group. Oxford: Oxford University Press.
- Levy, Neil. 2011. *Hard Luck: How Luck Undermines Freedom*. New York: Oxford University Press.
- Lewis, David. 1973. *Counterfactuals*. Cambridge: Harvard University Press.
- . 1983. "How To Define Theoretical Terms." Pp. 78-95 in *Philosophical Papers*. New York: Oxford University Press.
- McCormick, Kelly. 2013. "Anchoring a revisionist account of moral responsibility." *Journal of Ethics and Social Philosophy*. 7 (3):1-19.
- . 2016. "Revisionism." Pp. 109-120 in *Routledge Companion to Free Will*, edited by Kevin Timpe, Meghan Griffith, and Neil Levy. New York: Routledge.
- . 2015. "Companions in Innocence: Defending a New Methodological Assumption About Moral Responsibility." *Philosophical Studies*. 172 (2):515-533.
- . 2017. "Why We Should(n't) Be Discretionists About Free Will." *Philosophical Studies* 174 (10):2489-2498.
- McGeer, Victoria. 2012. "Co-reactive attitudes and the making of moral community." Pp. 299-326 in *Emotions, Imagination, and Moral Reasoning*, edited by R. Langdon and C. Mackenzie. New York: Psychology Press.
- . 2015. "Building a Better Theory of Responsibility." *Philosophical Studies*. 172 (10):2635-2649.
- McKenna, Michael. 2009. "Compatibilism and Desert: Critical Comments on Four Views on Free Will." *Philosophical Studies*. 144:3-13.
- . 2012. *Conversation and Responsibility*. New York: Oxford University Press.
- Nahmias, Eddy. 2006. "Folk Fears about Freedom and Responsibility: Determinism and Reductionism." *Journal of Culture and Cognition*. 6 (1-2):215-238.
- . 2011. "Intuitions about Free Will, Determinism, and Bypassing." Pp. 55-576 in *Oxford Handbook of Free Will, 2ed*, edited by Robert Kane. New York: Oxford University Press.
- . 2014. "Response to Misirlisoy and Haggard and to Björnsson and Pereboom." Pp. 43-57 in *Moral Psychology Volume 4: Free Will and Moral Responsibility*, vol. Moral Psychology: Volume 4: Free Will and Moral Responsibility, edited by Walter Sinnott-Armstrong. Cambridge, MA: MIT Press.

- Nahmias, Eddy, D. Justin Coates, and Trevor Kvaran. 2007. "Free Will, Moral Responsibility, and Mechanism: Experiments on Folk Intuitions." *Midwest Studies in Philosophy*. XXXI:214-241.
- Nahmias, Eddy, Stephen Morris, Thomas Nadelhoffer, and Jason Turner. 2006. "Is Incompatibilism Intuitive?" *Philosophy and Phenomenological Research*. 73 (1):28-53.
- Nelkin, Dana Kay. 2011. *Making sense of freedom and responsibility*. Oxford: Oxford University Press.
- Nichols, Shaun. 2006. "Folk Intuitions on Free Will." *Journal of Cognition and Culture*. 6 (1 & 2):57-86.
- . 2007a. "After Incompatibilism: A Naturalistic Defense of the Reactive Attitudes." *Philosophical Perspectives*. 21:405-428.
- . 2007b. "The Rise of Compatibilism: A Case Study in Quantitative History of Philosophy." *Midwest Studies in Philosophy*. 31:260-270.
- . 2015. *Bound: Essays on Free Will and Responsibility*. New York: Oxford University Press.
- . 2017. "Replies to Kane, McCormick, and Vargas." *Philosophical Studies* 174 (10):2511-2523.
- Nichols, Shaun and Joshua Knobe. 2007. "Moral Responsibility and Determinism: The Cognitive Science of Folk Intuitions." *Nous*. 41 (4):663-685.
- Nozick, Robert. 1981. *Philosophical Explanations*. Oxford: Clarendon Press.
- Pereboom, Derk. 2001. *Living Without Free Will*. Cambridge: Cambridge University Press.
- Roskies, Adina and Bertram Malle, F. 2013. "A Strawsonian Look at Desert." *Philosophical Explorations*. 16 (2):133-152.
- Roskies, Adina and Shaun Nichols. 2008. "Bringing Responsibility Down to Earth." *Journal of Philosophy*. 105 (7):371-388.
- Scanlon, T. M. 1998. *What We Owe to Each Other*. Cambridge, MA: The Belknap Press of Harvard University Press.
- Singer, Ira. 2002. "Freedom and Revision." *Southwest Philosophy Review*. 18 (2):25-44.
- Smart, J.J.C. 1961. "Free Will, Praise, and Blame." *Mind*. 70:291-306.
- Smilansky, Saul. 2000. *Free Will and Illusion*. New York: Clarendon.
- Spencer, Quayshawn. 2012. "What 'Biological Racial Realism' Should Mean." *Philosophical Studies*. 159 (2):181-204.
- Strawson, Galen. 1986. *Freedom and belief*. Oxford: Oxford University Press.
- . 1994. "The Impossibility of Moral Responsibility." *Philosophical Studies*. 75:5-24.
- Strawson, P. F. 1959. *Individuals, an Essay in Descriptive Metaphysics*. London: Methuen.
- . 1962. "Freedom and Resentment." *Proceedings of the British Academy*. XLVIII:1-25.
- van Inwagen, Peter. 2008. "How to Think about the Problem of Free Will." *Journal of Ethics*. 12:327-341.

- Vargas, Manuel. 2004. "Responsibility and the Aims of Theory: Strawson and Revisionism." *Pacific Philosophical Quarterly*. 85 (2):218-241.
- . 2004. "Libertarianism and Skepticism about Free Will: Some Arguments Against Both." *Philosophical Topics*. 32 (1&2):403-426.
- . 2005a. "The Revisionist's Guide to Responsibility." *Philosophical Studies*. 125 (3):399-429.
- . 2005b. "Compatibilism Evolves? On Some Varieties of Dennett Worth Wanting." *Metaphilosophy* 36 (4): 460–75.
- . 2011. "The Revisionist Turn: Reflection on the Recent History of Work on Free Will." Pp. 143-172 in *New Waves in the Philosophy of Action*, edited by Jesus Aguilar, Andrei Buckareff, and Keith Frankish. New York: Palgrave Macmillan.
- . 2013a. *Building Better Beings: A Theory of Moral Responsibility*. Oxford, U.K.: Oxford University Press.
- . 2013b. "If Free Will Does Not Exist, Then Neither Does Water." Pp. 177-202 in *Exploring the Illusion of Free Will and Moral Responsibility*, vol. Exploring the Illusion of Free Will and Moral Responsibility, edited by Gregg Caruso. Lanham, Maryland: Lexington Books.
- . 2013c. "Situationism and Moral Responsibility: Free Will in Fragments." Pp. 325-349 in *Decomposing the Will*, edited by Till Vierkant, Julian Kiverstein, and Andy Clark. New York: Oxford University Press.
- . 2015. "Desert, Responsibility, and Justification: Reply to Doris, McGeer, and Robinson." *Philosophical Studies*. 172 (10):2659-2678.
- . 2016. "The Runeberg Problem: Theism, Libertarianism, and Motivated Reasoning." Pp. 27-47 in *Libertarianism and Free Will: The Interplay of Religious Belief and Free Will*, edited by Kevin Timpe and Daniel Speak. New York: Oxford University Press.
- . 2017. "Implicit Bias, Moral Responsibility, and Moral Ecology." Pp. 219-247 in *Oxford Studies in Agency and Responsibility 4*, edited by David Shoemaker. New York: Oxford University Press.
- Wallace, R. Jay. 1994. *Responsibility and the Moral Sentiments*. Cambridge, MA: Harvard University Press.
- Walter, Henrik. 2004. "Neurophilosophy of Moral Responsibility: The Case for Revisionist Compatibilism." *Philosophical Topics*. 32 (1&2):477-532.
- Weigel, Chris. 2013. "Experimental Evidence for Free Will Revisionism." *Philosophical Explorations*. 16 (1):31-43.