Erkenn (2017) 82:837–855
DOI 10.1007/s10670-016-9846-2

Compatibilism and Retributivist Desert Moral Responsibility: On What is of Central Philosophical and Practical Importance

Gregg D. Caruso[1] · Stephen G. Morris[2]

Received: 26 January 2016 / Accepted: 22 September 2016 / Published online: 12 October 2016
© Springer Science+Business Media Dordrecht 2016

Abstract Much of the recent philosophical discussion about free will has been focused on whether compatibilists can adequately defend how a determined agent could exercise the type of free will that would enable the agent to be morally responsible in what has been called the *basic desert* sense (see, e.g., Pereboom in Living without free will, Cambridge University Press, New York, 2001; Pereboom in Free will, agency, and meaning in life, Oxford University Press, Oxford, 2014; Strawson in Freedom and belief, Oxford University Press, Oxford, 1986; Strawson in Philos Stud 75(1):5–24, 1994; Fischer in Four views on free will, Wiley, Hoboken, 2007; Vargas in Four views on free will, Wiley, Hoboken, 2007; Vargas in Philos Stud, 144(1):45–62, 2009). While we agree with Derk Pereboom and others that the compatibilist's burden should be properly understood as providing a compelling account of how a determined agent could be morally responsible in the basic desert sense, the exact nature of this burden has been rendered somewhat unclear by the fact that there has been no definitive account given as to what the basic desert sense of moral responsibility amounts to. In Sect. 1 we set out to clarify the compatibilist's burden by presenting our account of basic desert moral responsibility—which we call *retributivist desert moral responsibility* for purposes of clarity—and explain why it is of central philosophical and practical importance to the free will debate. In Sect. 2 we employ a thought experiment to illustrate the kind of difficulty that compatibilists of any stripe are likely to encounter in attempting to

✉ Stephen G. Morris
Stephen.morris@csi.cuny.edu

Gregg D. Caruso
gcaruso@corning-cc.edu

[1] Communications and Humanities Division, Corning Community College, 1 Academic Drive, Corning, NY 14830, USA

[2] Department of Philosophy, College of Staten Island, Building 2N, Room 224, 2800 Victory Boulevard, Staten Island, NY 10314, USA

explain *how* determined agents can exercise the kind of free will needed for retributivist desert moral responsibility.

In order to properly understand the substantive issues upon which compatibilists and incompatibilists disagree, it is imperative that we first get clear on what is at stake in the free will debate and what drives the dispute. In this paper we argue that what is of central philosophical and practical important in the traditional debate is the control in action required for a core sense of moral responsibility. This sense of moral responsibility is set apart by the notion of *basic desert* (see, e.g., Pereboom 2001, 2014; Strawson 1986, 1994; Fischer 2007; Clarke 2005) and is purely backward-looking and non-consequentialist. Understood this way, free will is a kind of power or ability an agent must possess in order to justify certain kinds of desert-based judgments, attitudes, or treatments in response to decisions or actions that the agent performed or failed to perform. These reactions would be justified on purely backward-looking grounds and would not appeal to consequentialist or forward-looking considerations, such as future protection, future reconciliation, or future moral formation.

While we agree, then, with Pereboom (2001, 2014) and others that the compatibilist's burden should be properly understood as providing a compelling account of how a determined agent could be morally responsible in the basic desert sense, the exact nature of this burden has been rendered somewhat unclear by the fact that there has been no definitive account given as to what the basic desert sense of moral responsibility amounts to. In this paper we present our account of basic desert moral responsibility—which we call *retributivist desert moral responsibility*—and explain why it is of central philosophical and practical importance to the free will debate. We begin in Sect. 1 by arguing that a proper understanding of the primary issues driving the free will debate requires us to adopt a conception of free will which views it as the control in action required for retributivist desert moral responsibility. To bolster our argument we consider two influential defenses of free will—those of Waller (2015) and Dennett (2003)—that explicitly sever the relevant notion of free will from retributivist moral responsibility and argue that they sidestep the primary debate between compatibilists and incompatibilists. In Sect. 2 we conclude by placing the burden on compatibilists to explain *how* determined agents can exercise the kind of free will needed for retributivist desert moral responsibility.

1 Basic Desert and Retributivism

In an attempt to understand what is truly at stake in the free will debate, a good place to begin is with the dispute between compatibilists and incompatibilists. One might reasonably think, as many philosophers do, that the heart of this dispute is about moral responsibility—i.e., whether agents who are causally determined by impersonal forces over which they have no ultimate control can ever be morally responsible for their actions. Galen Strawson, for example, writes, "It is a matter of

historical fact that concern about moral responsibility has been the main motor—indeed the *ratio essendi*——of discussion of the issue of free will" (1994, 8). In fact, the prevailing definition of "free will" in the literature today is one that defines the concept in terms of the control needed for moral responsibility. Speaking to this point, John Martin Fischer, Robert Kane, Derk Pereboom, and Manuel Vargas write: "Much of the tradition has taken 'free will' to be a kind of power or ability to make decisions of the sort for which one can be morally responsible" (2007, 1). Manuel Vargas's own definition reflects this: "'free will' is a term of art that picks out some distinctive power or capacity characteristic of morally responsible agency" (2013, 10). And Eddy Nahmias insists that "we should be concerned primarily with free will understood as the set of powers or abilities required to be morally responsible—that is, potentially to deserve blame or praise, punishment or reward" (2014, 43). Others philosophers who define free will in terms of the control condition for moral responsibility include Pereboom (2001, 2014), Strawson (1986, 1994), O'Connor (2000), McKenna (2008), Campbell (1957), Clarke (2005), Levy (2011), Richards (2000), Van Inwagen (1983), Caruso (2012), Morris (2015).

While there are some notable philosophers who reject this understanding of free will (e.g., Waller 2011, 2015), we contend that there are several distinct advantages to defining free will in this way: (a) it provides a neutral definition that virtually all parties can agree to—i.e., it doesn't exclude from the outset various conceptions of free will that are available for compatibilists, libertarians, and free will skeptics to adopt; (b) it captures the practical importance of the debate; (c) it fits with the commonsense (i.e., folk) understanding of these concepts; and, perhaps most importantly, (d) rejecting this understanding of free will makes it difficult to understand the nature of the substantive disputes that are driving the free will debate.

Before making our case for these advantages, we first need to make clear what sort of moral responsibility is at issue in the free will debate. To begin, we need to recognize that not all moral responsibility is threatened by determinism or the rejection of free will. For instance, forward-looking moral responsibility (e.g., Pereboom 2014) is not threatened. Nor is the *answerability* sense of moral responsibility defended by Scanlon (1998) and Hilary Bok (1998).[1] In trying to understand the particular sort of moral responsibility that *is* disputed in the free will debate, a good first step is to recognize the close connection that exists between moral responsibility and reward and punishment. To say that one is morally responsible for a good or bad act (in the sense relevant to free will) is to say that one *justly deserves* to be rewarded or punished, praised or blamed, for that act. This may include the expression of certain attitudes and judgments, like praise or blame, or extend upwards to more severe forms of retributive punishment.

[1] According to Thomas Scanlon, for instance, to blame an agent for an action is to judge that it reveals something about the agent's attitudes toward oneself and/or others that impairs the relations that she can have with them, and to take one's relationship with her to be modified in a way that this judgment of impaired relations justifies as appropriate (Scanlon 2009, 128–31). Derk Pereboom has argued that there is an epistemic or evidential interpretation of Scanlon that is perfectly consistent with free will skepticism (see 2012).

Many contemporary philosophers share the view that the sort of moral responsibility at issue in the free will debate is tightly linked to reward and punishment, praise and blame, in this way. It is especially prevalent among free will skeptics, many of whom are motivated by the intuition that it is inappropriate to inflict certain types of rewards or punishments upon agents who are causally determined by forces over which they have no ultimate control. Free will skeptics, however, are not the only ones who think that the central philosophical question is (and should be) about just deserts, praise and blame, and reward and punishment. The compatibilist Susan Wolf, for example, writes:

> [I]t seems to me that the philosophical problem of free will is and has been fundamentally connected to the question of whether and how the distinctive set of attitudes and the practices that constitute my kind of blame [what Wolf calls "angry blame" involving attitudes such as "resentment, indignation, guilt, and righteous anger"] can be justifiable and appropriate. We cannot understand the history of the free will debate without making reference to this set of attitudes and practices, nor can we do justice to the continued discussion of the problem if we fail to recognize that the intelligibility and legitimacy of this set of attitudes and practices in particular is at least part of what is at stake. (2011, 335)

And the libertarian Randolph Clarke similarly writes:

> If any agent is truly responsible [in the sense relevant to free will]…that fact provides us with a specific type of justification for responding in various ways to that agent, with reactive attitudes of certain sorts, with praise or blame, with finite rewards or punishments. To be a morally responsible human agent is to be truly deserving of these sorts of responses, and deserving in a way that no agent is who is not responsible. This type of desert has a specific scope and force—one that distinguishes the justification for holding someone responsible from, say, the fairness of a grade given for a performance or any justification provided by consequences. (2005, 21)

We contend that Wolf and Clarke are correct in framing the problem as one about the justification of various sorts of desert-based judgments, attitudes, or treatments—e.g., those related to just deserts, praise and blame, and the reactive attitudes (Strawson 1962) of resentment, indignation, guilt, and righteous anger.

Derk Pereboom provides a very helpful definition of the kind of moral responsibility at issue in the free will debate, which he calls *basic desert* moral responsibility and defines as follows:

> For an agent to be morally responsible for an action in this sense is for it to be hers in such a way that she would deserve to be blamed if she understood that it was morally wrong, and she would deserve to be praised if she understood that it was morally exemplary. The desert at issue here is basic in the sense that the agent would deserve to be blamed or praised just because she has performed the action, given an understanding of its moral status, and not, for example, merely by virtue of consequentialist or contractualist considerations. (2014, 2)

Pereboom's definition can be seen as an effort to clarify what the real issue is between compatibilists and incompatibilists: it boils down to whether or not determined beings can be morally responsible in the basic desert sense.

While we agree with Pereboom that it is basic desert moral responsibility that is of central philosophical and practical importance in the free will debate, it is worth mentioning that some philosophers have expressed confusion as to what *basic desert* is supposed to represent. Michael McKenna, for example, claims that he has "no clear sense of what anyone in the free will debate means by desert" (2009, 9). In order to further clarify the concept, we propose a few elucidations. First, holding one morally responsible in basic desert sense for, say, a bad act cannot simply mean that it is appropriate to punish the agent on consequentialist grounds. If this were all there was to the sort of moral responsibility at issue in the free will debate, it would be difficult to understand what all the bickering is about between compatibilists and incompatibilists. After all, even an incompatibilist like the free will skeptic can allow that there are legitimate consequentialist reasons for punishing determined agents (and other agents who they believe lack free will). For example, free will skeptics typically point out that the imposition of sanctions serves purposes other than the punishment of the guilty: it can also be justified by its role in incapacitating, rehabilitating, and deterring offenders (see, e.g., Caruso 2016a; Pereboom 2001, 2014; Pereboom and Caruso (Forthcoming); Levy 2012). If ascriptions of moral responsibility were just about consequentialist justifications for reward/punishment, we should expect both compatibilists and incompatibilists alike to attribute moral responsibility to determined agents. The fact that incompatibilists do not indicates that there is something more at stake than consequentialist considerations when it comes to the debate over free will and moral responsibility.

To further clarify the concept of *basic desert*, we believe it helpful to consider what another free will skeptic, Galen Strawson, has said about the kind of moral responsibility at issue in the free will debate. Referring to this type of moral responsibility as "true moral responsibility," Strawson has described it as "responsibility of such a kind that, if we have it, then it *makes sense*, at least, to suppose that it could be just to punish some of us with (eternal) torment in hell and reward others with (eternal) bliss in heaven" (1994, 9). In response to this suggestion, many philosophers have argued that the excessively retributivist notions of eternal suffering or eternal bliss at work here cannot accurately capture the more modest desert element seemingly at work in the commonsense understanding of moral responsibility. While this may be true, the idea of divine retribution in the afterlife seems a plausible way of understanding the notion of desert that figures centrally in the primary philosophical disputes regarding free will, since it seems safe to assume that most people would view such divine treatment as being justified—if at all—on retributivist as opposed to consequentialist grounds.

Retributivism refers roughly to the justification for treatment whereby an individual is either rewarded or punished as payback for the moral rights and wrongs he has committed. Consequentialist considerations do not figure into justifications for treatment from this perspective. Mitchell Berman aptly captures the idea behind retributivist punishments when he writes, "A person who unjustifiably and inexcusably causes or risks harm to others or to significant social

interests deserves to suffer for that choice, and he deserves to suffer in proportion to the extent to which his regard or concern for others falls short of what is properly demanded of him" (2008, 269). Following Strawson's lead, we suggest that the idea of deserving retribution in the afterlife accurately captures the kind of moral responsibility at issue in the free will debate. Instead of deserving either *eternal* reward or punishment, however, we propose that the key question with regard to basic desert moral responsibility is whether it would ever be appropriate for a divine all-knowing judge (who didn't necessarily create the agents in question) to administer differing kinds of treatments (i.e., greater or lesser rewards or punishments) to human agents in a hypothetical afterlife on the basis of actions and/or decisions that these agents performed during their lifetimes.

Regardless of one's religious views, we think it helpful to invoke the notion of an all-knowing divine judge who is responsible for doling out rewards and punishments in the afterlife, commensurate with one's deeds during his prior life, since it helps to instill the idea that any reward or punishment issued after death will have no further utility—be it positive or negative. Any differences in treatment to agents (however slight) would therefore seem warranted only from a *basic desert*, and not a consequentialist, perspective. The kind of moral responsibility that we claim is at the heart of the free will debate is the kind that could warrant such retributive treatment in the afterlife. For the sake of clarity, we will call this *retributivist desert moral responsibility*.[2]

It is important to note that retributivist desert moral responsibility does *not* entail that we, or an all-knowing judge, are *morally obligated* to blame or punish an agent who has knowingly done wrong in each and every instance. There is no contradiction involved in believing that an agent *deserves* to be blamed or punished, but since blaming or punishing her would be undesirable for other more important reasons, we ought not to do so all things considered. The question of retributivist desert moral responsibility is instead about whether it would ever be *justified* or *warranted* to blame or punish a wrongdoer on purely backward-looking, non-consequentialist grounds. As Sofia Jeppsson writes, "The claim that an agent is accountable [in the basic desert sense] for what she did does not entail that future consequences are completely irrelevant to the question of how we ought to treat her. It does entail, however, that *in the absence of sufficiently strong counter-veiling reasons*, blaming her is justified if her action was wrong and praising her justified if it was exemplary" (2016).

We should also note that our conception of *retributivist moral responsibility* includes backward-looking blame/punishment *as well as* praise/reward. This is

[2] An additional advantage of our conception of *retributive desert moral responsibility* is that it is less likely to be clouded by earthly considerations of punishment and reward. There are many reasons for people to be opposed to retributivism in our Earthly criminal justice system that has nothing to do with the existence of moral responsibility per se (see Pereboom 2014 for examples). People might, for instance, be opposed to retributivism because they believe that in real life we will often end up punishing too harshly, or that innocent people will be punished, etc. Just like people might seem more compatibilist than they actually are when considering Earthly punishment, because they allow considerations of utility to sneak in, others might seem more incompatibilist than they actually are in the same context, because they allow Earthly problems with retributivism to influence their intuitions. Invoking an all-knowing judge in the afterlife solves both problems.

perhaps somewhat atypical since retribution is often associated with blame and punishment, not with praise and rewards. We acknowledge that this broader conception of retributivism may sound odd to some ears, but we believe it is in keeping with the core justification of retributivism—which is backward-looking and non-consequentialist. But even if one thinks our use of the concept differs significantly from common usage, we provide the following stipulated definition so as to avoid confusion about what we mean: *retributivism*, as we are using the term, refers to the justification for treatment whereby an individual is either rewarded or punished as payback for the moral rights and wrongs he has committed. In our thought experiment this serves as the basis of the all-knowing judge administering different kinds of treatments (i.e., greater or lesser rewards or punishments) in a hypothetical afterlife.

Lastly, it should be pointed out that while virtually all philosophers are willing to accept that being morally responsible for a good or bad act implies that one deserves *praise* or *blame* respectively, one might balk at the suggestion that such praise/blame equates to being an appropriate target for reward/punishment. Following the lead of Strawson (1962), one might counter that deserving praise or blame only implies that it would be proper to harbor certain positive or negative reactive attitudes like gratitude or resentment towards the morally responsible individual. In response to this view, we would ask what it means to say that "one is an appropriate target for reactive attitudes" if not that "it would be appropriate for one to be either rewarded or punished"? Such a question would likely be posed by Feinberg (1970), who proposes that blame is a form of punishment. And this is because overt blame directly targets a living person and often causes emotional pain and suffering. When the justification for blame is purely backward-looking and non-instrumental, it resembles punishment in that the harm caused is believed to be justified on the grounds that the agent *deserves* it. This justification makes no appeal to forward-looking or instrumental benefits such as future safety, future reconciliation, or future moral formation (see Pereboom 2014). Our hypothetical thought experiment makes this point salient since it removes any possible forward-looking benefit that could come from blaming the agent.

At any rate, even if one can offer a coherent account of what it means to be "an appropriate target of reactive attitudes" that does not appeal to reward and punishment in some manner, an account of moral responsibility based upon such reactive attitudes alone does not seem relevant to the compatibilism/incompatibilism debate. A key intuition motivating incompatibilists—and hence, one that drives the primary controversies surrounding free will—is that it is inappropriate in some sense to reward or punish determined agents for their actions. Where one adopts a notion of moral responsibility that is not related to reward or punishment, the incompatibilist may allow that determined agents are capable of possessing responsibility of this sort (see, e.g., Pereboom 2014). But such allowances can hardly be celebrated as a victory for compatibilism since appealing to a type of moral responsibility that is detached from reward and punishment would bypass a primary issue upon which the incompatibilist differentiates his position from that of the compatibilist. For this reason, it seems appropriate in the context of the free will

debate to reject the understanding of moral responsibility favored by P.F. Strawson in favor of one in which the propriety of reward and punishment plays a central role.

Defining free will as whatever satisfies the control condition for retributivist desert moral responsibility has a number of advantages. As mentioned earlier, it not only provides a neutral definition that gives clarity to the dispute, it also captures the *practical importance* of the debate. Since much of the interest surrounding the free will debate owes its existence to how free will is relevant to questions relating to moral and socio-political issues, it would seem misguided to redefine free will in such a way that detracts from its relevance to such practical matters. To illustrate how traditional philosophical discussions about free will have real-life practical implications, consider the issue of criminal punishment. Addressing this issue, Neil Levy writes:

> Traditionally, incarceration is seen as justified, in part, by the *desert* of offenders: because they are *guilty*—morally, and not merely legally, guilty—we can impose significant sanctions on them; the more weighty the sanctions, the more such a justification is required.…But if [free will and] moral responsibility skeptics are right, agents are never deserving of the imposition of such sanctions. Thus moral responsibility skepticism has practical implications: it apparently entails that major elements of the criminal justice system are unjustified. (2012, 481)

This is just one example of the practical importance of the debate, and retributivist desert moral responsibility captures it since it maintains that free will is inextricably connected to the justification of certain kinds of judgments, attitudes, and treatments that are important to criminal law, social policy, and our interpersonal relationships. Definitions of free will that sever the relationship between free will and retributivist desert moral responsibility lose sight of the practical implications of the debate and, hence, remove much of its interest.

Furthermore, defining free will in relation to retributive desert moral responsibility also has the advantage of reflecting how ordinary folk conceive of these issues. Perhaps the strongest evidence yet linking folk conceptions of free will and retributivism comes from a set of recent studies by Shariff et al. (2014). Shariff and his colleagues hypothesized that if free will beliefs support attributions of moral responsibility, then reducing these beliefs should make people less retributive in their attitudes about punishment. In a series of four studies they tested this prediction. In Study 1 they found that people with weaker free will beliefs endorsed less retributive attitudes regarding punishment of criminals, yet their consequentialist attitudes were unaffected. In the study, two hundred and forty-four American participants (147 female; mean age = 36.81 years) completed the seven-item Free Will subscale of the Free Will and Determinism Plus scale (FAD+) (Paulhuse and Carey 2011), which measures belief in free will. In order to further measure attitudes toward retributivist and consequentialist motivations for punishment, Shariff and his colleagues had participants read descriptions of retributivism and consequentialism as motivations for punishment and then indicate on two separate Likert scales (1 = *strongly disagree,* 7 = *strongly agree*) how important retributivism and consequentialism should be in determining motivation for criminal

punishments. As predicted, Shariff et al. found that stronger belief in free will predicted greater support for retributive punishment, but was not predictive of support for consequentialist punishment. The effects remained significant when statistically controlled for age, gender, education, religiosity, and economic and social political ideology. Study 1 therefore supports the hypothesis that free will beliefs positively predict retributive attitudes, yet it also suggests that "the motivation to punish in order to benefit society (consequentialist punishment) may remain intact, even while the need for blame and desire for retribution are forgone" (Shariff et al. 2014, 7).

Study 2 further found that experimentally diminishing free will belief through anti-free-will arguments diminished retributive punishment, suggesting a causal relationship. In the study, participants were randomly assigned to one of two groups. In the anti-free will condition, participants were given a passage from Crick's (1994) *The Astonishing Hypothesis* which rejected free will and advocated for a mechanistic view of human behavior. In the neutral condition, the passage was unrelated to free will. Next, participants read a fictional vignette involving an offender who beat a man to death. Acting as hypothetic jurors, participants recommended the length of the prison sentence (if any) that this offender should sever following a 2-year, nearly 100 %-effective, rehabilitation treatment. As Shariff et al. describe:

> The notion that the offender had been rehabilitated was used in order to isolate participants' desire for punishment as retribution. The passage further focused participants on retributive, rather than consequentialist, punishment by noting that the prosecution and defense had agreed that the rehabilitation would prevent recidivism and that any further detention after rehabilitation would offer no additional deterrence of other potential criminals. (Shariff et al. 2014, 4)

As predicted, participants who read the anti-free passage recommended significantly lighter prison sentences than participants who read the neutral passage. More specifically, participants whose free-will beliefs had been experimentally diminished recommended roughly half the length of imprisonment ($\sim$5 years) compared with participants who read the neutral passage ($\sim$10 years).

Both of these findings are significant for our argument since it would be hard to explain these results if ordinary folk considered free will and moral responsibility independent issues.[3] These results indicate that there is a commonsense causal

[3] A similar link between the belief in free will and retributivist attitudes was suggested by the studies of Carey and Paulhus (2013). In the third of their studies, Carey and Paulhus presented two scenarios portraying serious crimes (child molestation and the rape of an adult woman) and tested the degree to which subjects' attitudes towards punishment of the criminals would be impacted by factors including the criminal having been abused as a child and assurance that a medical procedure would prevent the criminal from ever perpetrating similar crimes again. The fact that subjects who expressed the strongest belief in free will were essentially the only group of subjects whose attitudes towards punishment were not mitigated by environmental or consequentialist considerations led the researchers to conclude that "free will belief is related to retributivist punishment" (2013, 138).

relationship between belief in free will and retributivism, and the relationship is specific to retributive punishment and not just consequentialist punishment.[4]

Insofar as one believes that an acceptable notion of free will should not veer too far from the commonsense understanding of the term, such empirical findings lend credence to the idea that retributivist reward and punishment is one of the primary substantive issues, if not *the* primary substantive issue, at dispute in the free will debate. While some may downplay the importance of folk intuitions with regard to philosophical discussions about free will, we agree with Al Mele who holds that any adequate philosophical analysis of free will should be "anchored by common-sense judgments" since any such analysis that is completely disconnected from what the folk have to say "runs the risk of having nothing more than a philosophical fiction as its subject matter" (2001, 27).

Some philosophers, however, are more willing to discard our folk psychological intuitions and opt instead for a revisionist approach to free will. So-called *revisionism* sets out to alter the commonsense meanings of terms like "free will" and "moral responsibility" in a way that allows us to assert the existence of the general kinds of attributes that these terms attempt to capture without committing ourselves to the more philosophically problematic aspects of the commonsense notions. Revisionists in the free will debate basically hold that the way we *ought to* think about the terms "free will" and/or "moral responsibility" differs from the way the folk *actually do* think about them. Manuel Vargas (2007, 2009, 2013), for example, defends a version of revisionism which is predicated on two main theses: (1) our commonsense intuitions about free will and moral responsibility are primarily libertarian in nature (his *descriptive thesis*); and (2) we ought to revise our concepts of "free will" and "moral responsibility" in such a way as to strip them of their libertarian commitments (his *prescriptive thesis*).

From Vargas's perspective, what matters most in the free will debate is not whether human agents possess the libertarian type of free will and moral responsibility that he believes the folk typically consider themselves to have. He believes, rather, that the real importance of the concepts of free will and moral responsibility is the role they play in allowing for the effectiveness of what he calls "the responsibility system"—that is, "the responsibility norms, and their attendant social practices, characteristic attitudes and paradigmatic judgments," where such judgments include praise and blame (2007, 154). He believes that the key purpose of the responsibility system is "to get creatures like us to better attend to what moral considerations there are and to appropriately govern our conduct in light of what moral reasons those generate" (2007a, 155). Note, however, that Vargas' revisionism, while jettisoning what he takes to be our predominantly libertarian intuitions about free will, maintains the all-important link between free will and moral responsibility. As we saw earlier, he endorses a definition of free will that takes it to be "a kind of power or ability to make decisions of the sort for which one can be morally responsible." His revisionism is, therefore, not at odds with our

[4] Studies 3 and 4 further found that exposure to neuroscientific findings implying a mechanistic basis for human action—in the form of having participants read popular-science articles (study 3) or take an introductory neuroscience class in college (study 4)—similarly produced a reduction in retributivism.

proposed definition since it preserves those folk psychological intuitions *most important* to our understanding of free will. It is important to recognize that not all revisionisms are created equal. In Vargas' revisionism, some folk psychological intuitions are open for revision (e.g., our libertarian intuitions), while others remain central to our understanding of free will (e.g., those related to desert and our moral responsibility practices). We maintain that any revisionist account that severs all connection to moral responsibility and desert would be pragmatically vacuous, uninteresting, and unacceptable from the standpoint of the historical debate.

This brings us to our last, and perhaps most important, point: rejecting the link between free will and retributive desert moral responsibility would make it difficult, if not impossible, to understand the nature of the substantive disagreement between compatibilists and incompatibilists. To illustrate this point we will consider two defenses of free will—i.e., those of Waller (2015) and Dennett (2003)—that explicitly sever, in different ways, the relevant notion of free will from the kind of retributivist moral responsibility that we have been discussing. Let us begin by considering Bruce Waller's account of *restorative free will* (2015)—a revisionist account that, unlike Vargas', defends an understanding of free will that is "liberated from the burden of justifying moral responsibility" (Vargas 2015, 96).[5] According to Waller: "Free will is not unique to humans, it does not require a high level of reflective rationality, it does not support moral responsibility, and it does not involve any special *causa sui* powers" (2015, 1). On Waller's account, *free will is common to many* species, and various species manifest free will in a variety of ways. As he puts it:

> On the restorative view, the free will powers enjoyed by human animals are basically similar to those of other animals, and free will—including the human variety—is better understood by examining the common elements of free will shared by many species, rather than focusing on the distinctive *enhancements* of free will that are unique to humans. (2015, 1)

On this account, free will amounts to the ability to discriminate among and evaluate alternatives and the ability to adjust the level of behavioral variability to environmental conditions.

There are two main elements of restorative free will: *options* and *control*. Regarding the former, Waller writes: "Free will requires the capacity to effectively explore open alternatives; that includes both the capacity to *explore* open alternatives, and to *effectively* explore those alternatives" (2015, 140). The kind of options Waller has in mind, however, are not "absolute options" or "open alternatives that are chosen independently of all conditioned preferences and situational influences" (2015, 140). Such options, we are told, may be appropriate for gods but not foraging mice or humans:

> Human animals do not need alternatives that transcend our conditioned characters and our environmental influences; to the contrary, like other

[5] Waller denies that restorative free will is a revisionist account, but see Caruso (2016b) for why it is accurate to call it a form of revisionism.

> animals that evolved in and must survive in this changing world, we need options that are closely tied to changing environmental contingencies (140).

According to Waller, our characters, situations, and histories may determine our choices, but so long as the choice process does not "bypass" us (Mele 1995; Nahmias 2011) then "we have all the free will we want or need" (2015, 141). In addition to options, Waller argues we must also be able to vary the levels of exploration and control in relation to varying environmental conditions. Here again "absolute control" is not needed. Instead, Waller argues, we want control that enables us to consider alternatives, place them in some workable order, and make our own choices.

On Waller's account, open alternatives and control balance one another out as conditions change. To help explain how this balancing occurs, Waller turns to the example of foraging animals. When a food source is reliable, foraging animals continue to engage in some behavioral variability, but that variability is reduced while the control over a reliable food source is greater. As the reliability of that food source wanes, behavioral variability waxes—i.e., the animal engages in a greater variety of overall behavior, it explores more alternative paths, tries new techniques, or new combinations of behavior. Vital to the effective exercise of control, then, is the opportunity to pursue paths that are genuinely productive, and that requires adjusting to changing conditions and having access to workable and controllable paths when a path that has been favored and that seems (and perhaps was) productive is no longer beneficial or promising.

We contend that this conception of free will makes the dispute between compatibilists and incompatibilists a moot point since virtually no one in the debate denies that we have the kinds of abilities discussed by Waller. The question most philosophers are interested in—the question that lies at the heart of the debate—is whether these abilities are *enough* to justify certain kinds of desert-based judgments, attitudes, or treatments in response to decisions or actions that an agent performed or failed to perform. It's hard to see how Waller's conception of free will helps understand that debate, or frankly any other significant debate related to the historical problem of free will. Even if we grant Waller his restorative free will, it is difficult to think of anything of importance that follows from it regarding moral responsibility or our everyday practices, judgments, and attitudes. By liberating free will from moral responsibility, Waller has seemingly liberated it from all of its philosophical and practical importance.[6]

[6] Extending free will downward on the phylogenetic scale may have the effect of making us feel more connected to the rest of the living world and less special or unique, something Waller stresses, but these issues are really tangential to the historical debate and don't cut one way or the other with regard to the debate between compatibilists and incompatibilists. In the spirit of fairness, however, Waller's conception of free will—while not sufficient for desert-entailing moral responsibility and not related to what compatibilists and incompatibilists have disagreed about all along—may still be important when it comes to other issues. For example, whether animals have Waller-style free will may have moral implications for animal rights, etc. Nonetheless, we maintain that it would be problematic to understand the term "free will" along the lines suggested by Waller for the reasons provided in this essay (the term is not really relevant to the debates between compatibilists and incompatibilists, etc.). Given, however, that the capacities captured by Waller-style free will are relevant in terms of moral considerations, etc., we recommend that such capacities be noted by a term other than free will—perhaps something along the lines of "behavioral variability".

Unlike Waller, Daniel Dennett does not argue for a kind of free will that is detached from moral responsibility. Yet insofar as the kind of moral responsibility that interests Dennett would be acceptable to even a free will skeptic, we believe that Dennett-style compatibilism ought to be rejected since it is not discernable from free will skepticism in any substantive way. According to Dennett: "Free will is real, but it is not a preexisting feature of our existence, like the law of gravity. It is also not what tradition declares it to be: a God-like power to exempt oneself from the causal fabric of the physical world. It is an evolved creation of human activity and beliefs, and it just as real as such other creations as music and money" (2003, 13). This is what Dennett means when he says that *freedom evolves* (2003). On Dennett's account, free will is an evolved ability that human beings have to anticipate the consequences of our actions and to modify them accordingly. This ability, to construct counterfactuals of great complexity and to model the actions of thousands of interacting forces and agents, is what gives us free will in the only sense that matters. It allows us, requires us, to shape our actions to our moral obligations.

It's important to note that Dennett does not set out to defend the kind of retributivist desert moral responsibility we claim is important, nor does he think such moral responsibility is coherent or compatible with determinism and naturalism. In a revealing exchange with Tom Clark and Bruce Waller, on Waller's book *Against Moral Responsibility* (2012), Dennett makes this point clear. After explaining how Waller rejects retributivist desert moral responsibility, Dennett writes:

> [L]et me say right away that I agree with Waller's main conclusion in one important sense: *that* kind of absolutistic moral responsibility—insisting as it does on what I have called guilt-in-the-eyes-of-God—is incompatible with naturalism and has got to go. (2012)

Rejecting what he considers to be an untenable retributivist system of criminal punishment and the strong type of retributivist desert moral responsibility that it relies upon, Dennett proposes "a consequentialist defense of just deserts" and punishment (2012). While many compatibilists want to secure a justification for retributive blame and punishment, Dennett does not seem to be among them. Instead, he espouses a reformed non-retributivist, consequentialist concept of moral responsibility and punishment. What is of paramount importance for Dennett is that we find a naturalistic way to justify punishment so as to maintain a "secure and civil society" (2012). The justification Dennett provides for punishment, however, is decidedly different than what many compatibilists, like Fischer (1986, 1994, 2013) for instance, would endorse since he acknowledges that we must abandon retributivist desert moral responsibility and replace it with an "ultimately consequentialist" (2012) conception of punishment and reward. Given that most free will skeptics would agree, Dennett's account fails to preserve what we have argued is of central philosophical and practical importance—the claim that people deserve to be praised and blamed, rewarded and punished, on strictly retributivist and non-consequentialist grounds. In fact, Dennett's account doesn't even attempt to preserve this sort of moral responsibility and he cheerfully wishes it "good riddance" (2012).

In many ways, Dennett's justification of punishment is akin to that of the free will skeptic who defends a forward-looking justification grounded in future protection,

future reconciliation, and future moral formation (e.g., Pereboom 2001, 2014; Caruso 2016a). Recall that free will skeptics acknowledge that the imposition of sanctions serves purposes other than the punishment of the guilty: it can also be justified by its role in incapacitation, rehabilitation, and deterring offenders. The consequentialist justification Dennett provides is therefore perfectly consistent with free will skepticism. One obvious difference, however, is that the free will skeptics reject the notion of "just deserts" whereas Dennett wishes to keep it in the lexicon. We contend that the concept of just deserts is inconsistent with Dennett's reformed consequentialist account of moral responsibility. As Tom Clark writes in response to Dennett:

> Whether as consequentialists we should still talk of just deserts is debatable, given the strong deontological, retributive connotations…What you're advocating is the *practical necessity* of punishment, not its intrinsic goodness, but "just deserts" strongly implies that the offender's suffering is intrinsically good, which you don't think is the case. So I think we should drop talk of just deserts so we don't mislead people about what we believe are defensible justifications for punishment. (2012)

We believe Clark is correct. Given the canonical understanding of "just deserts" and how it is used to justify various retributive attitudes, judgments, and treatments, Dennett's use of the term lends itself to easy confusion and gives the mistaken impression that he is setting out to preserve something that he is not.

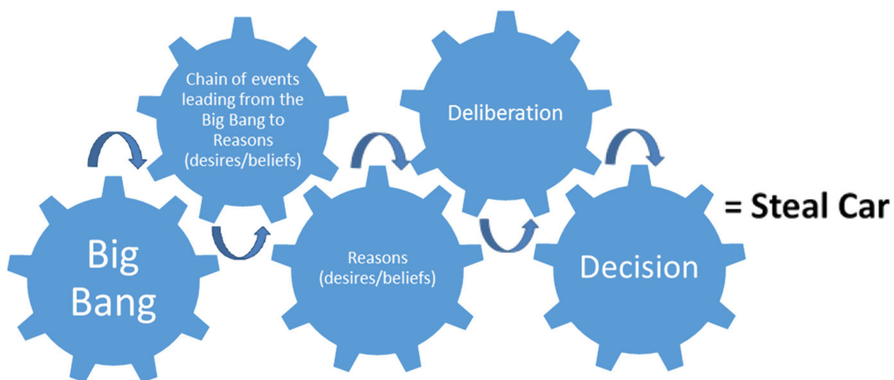
In conclusion, Dennett's brand of compatibilism fails to preserve retributivist desert moral responsibility—in fact, it rejects it outright. Furthermore, his justification for punishment, being consequentialist in nature, is completely consistent with the skeptic's rejection of free will and just deserts. While Dennett himself prefers to retain the notion of just deserts, we contend that this is misleading and potentially inconsistent with his reformist conception of moral responsibility. Drawing from this analysis, we conclude more generally that Dennett-style compatibilists who deny that determined agents could exercise retributivist free will should forego asserting the existence of "free will" in humans in order to avoid terminological confusion.

To sum up this section, then, we have argued that there are strong reasons why philosophers ought to adopt an understanding of free will which could enable agents to possess retributivist desert moral responsibility (hereafter we will refer to it as *retributivist free will* so as to differentiate it from Waller's *restorative free will*). Not only does such an understanding make sense of the kinds of pragmatic issues that figure heavily in contemporary debates about free will, it also helps to account for why many philosophers are so fervent in their rejection of compatibilism. The fact that presumably no self-identified incompatibilist would accept that a determined agent could possess retributivist desert moral responsibility, while many self-identified compatibilists would, is perhaps the best evidence that it is this kind of moral responsibility that serves as the basis of the main substantive difference between compatibilists and incompatibilists. For these reasons, we recommend that philosophers of all stripes adopt the understanding of free will that we endorse in order to avoid verbal disagreements and to preserve many of the most philosophically interesting points of dispute in the free will debate.

2 The Compatibilist's Burden

The preceding considerations suggest that a successful compatibilist account of free will would have to provide a compelling explanation of how a determined agent could possess retributivist desert moral responsibility. To face this burden squarely, the compatibilist would need to provide more than a consequentialist justification of forward-looking punishment, as Dennett does. While it is not the aim of this paper to argue that no such account is possible, we would like to make the challenge facing compatibilism vivid with the following thought experiment. We take the scenario described in it to accurately depict the kind of universe that a compatibilist would find to be consistent with free agents. We also contend that it is hard to see how it would be proper to treat the determined agents in it differently in the kind of afterlife scenario that we described earlier and, hence, how the determined agents could exercise retributivist free will. To the extent that others share our intuitions, the thought experiment helps to illuminate the difficulties facing compatibilists and, by doing so, can serve as a kind of springboard from which compatibilists can attempt to overcome the challenge that determinism poses to retributivist free will.

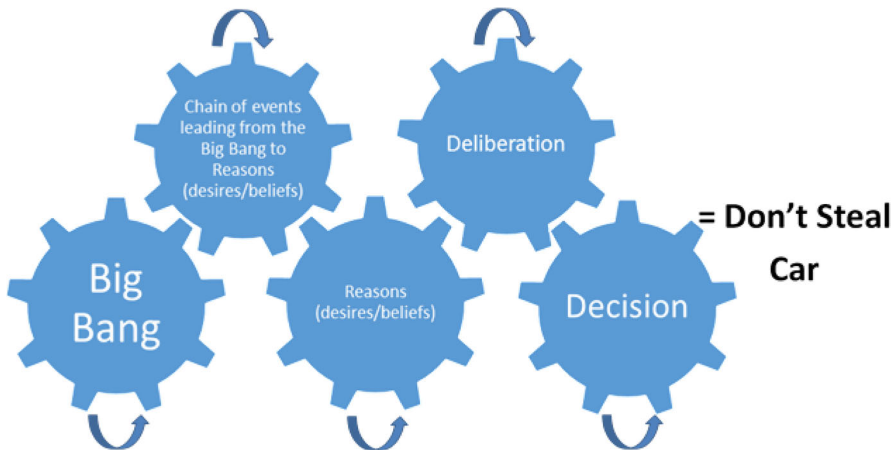
Imagine a deterministic universe (Universe A) in which every event is completely caused by what came before it, leading all the way back to the Big Bang (stipulate this as the first cause). In this universe there are intelligent human beings who, like those in this universe, can deliberate about what to do. Suppose that the actions of people in Universe A had to happen exactly the way they did given the Big Bang that created Universe A and the laws of nature. In Universe A, suppose an individual named Maya decides to steal a car after deliberating about whether or not to do so. Now it had to be the case that given the way Universe A was at the time of the Big Bang, Maya had to deliberate, decide, and act *exactly* the way she did. Think of her action as the result of a series of gear-like motions depicted as follows (where the first gear would turn the second gear a specific way, the second gear would turn the third gear a specific way, etc.):



Notice that had any one of the gears been missing (e.g., the events prior to her birth, her deliberating about what to do), she wouldn't have performed the action. Even so, each of the gears (deliberation, etc.) had to move the way they did given

the laws of nature that exist in Universe A, the physical properties of that universe, and the way the Big Bang occurred.[7]

Now imagine another universe almost identical to Universe A (call this universe "Universe B"), in which an individual named Marina decides *not* to steal a car after deliberating about whether or not to do so. As was the case in Universe A, Marina had to deliberate and decide the way she did—and not steal the car—given the way the universe was. Think of her action as the result of a series of different gear-like motions depicted as follows:



Now assume that had Universe A's Big Bang been *exactly* like the Big Bang in Universe B (and all the subsequent gears moved exactly as they did in Universe B), Maya would have ended up not stealing the car. Likewise, had Universe B's Big Bang been *exactly* like the Big Bang in Universe A, Marina would have ended up stealing the car. Suppose now that an all-knowing supernatural deity exists across both Universes A and B. Though this deity did not create either the universes or the agents within them, it is up to him to decide whether to reward or punish both Maya and Marina for their respective actions. Let us further assume that whatever punishment or reward he gives to Maya and Marina will have *no effect whatsoever* on the way that anyone acts or thinks in the future (including Maya and Marina).

At this point the question is whether it would be appropriate for the deity to punish or reward, praise or blame, Maya and Marina differently. Being incompatibilists, we believe the answer is clearly "No" despite the fact that each agent appears to fulfill the conditions for free and responsible action put forth by compatibilists. We can imagine that both Maya and Marina satisfy all the relevant compatibilist-friendly conditions for acting freely. For example, we can imagine

[7] While one may worry that the above diagram is too simplistic to serve the present purposes, we would counter that it effectively captures the relevant factors at play in a standardly conceived deterministic universe with regard to human decisions and actions. With this in mind, it is reasonable to expect that one could construct a more detailed diagram consisting of however many deterministic "gears" one wanted without altering either the general deterministic view of behavior depicted by our diagram or the relevant intuitions with regard to free will and moral responsibility that would be elicited by a diagram of this sort.

that they satisfy the compatibilist conditions proposed by Hume—i.e., neither Maya nor Marina act out of character, nor are they externally constrained, nor are the desires that motivate them irresistible (i.e., they are not internally compelled by an overriding desire). We can further imagine that both Maya and Marina meet the compatibilist condition proposed by Frankfurt (1971)—i.e., they both will, and want to will, what they do. In addition, we can assume that the actions of both satisfy the reasons-responsiveness condition advocated by John Fischer and Mark Ravizza (1998)—i.e., their desires can be modified by, and some of them arise from, the rational consideration of their reasons. We can also assume that Maya and Marina satisfy the condition advanced by Wallace (1994)—i.e., they both have the general ability to grasp, apply, and regulate their actions by moral reasons. Finally, we would like to point out that since the deity charged with administering reward/punishment in the afterlife created neither the world in which Maya and Marina live nor Maya nor Marina themselves, our thought experiment seems insulated from the objection that any incompatibilist intuitions it elicits are most likely the result of transferring responsibility for Maya's and Marina's actions onto the deity.

Despite the fact that all these compatibilist-friendly conditions are satisfied, we contend that most people would share the intuition that it would not be fair to treat Maya and Marina differently, and that this straightforwardly follows from how the factors that led to their diverging acts were (in a clear and relevant sense) outside of their control.[8] We can imagine that upon hearing that the deity has decided to punish Maya and reward Marina, a Clarence-Darrow-like figure in the afterlife earnestly protests that it is unfair to give Maya worse treatment than Marina given that their actions were equally determined by circumstances beyond their control. If one were to share this sense of injustice, it seems correct to assume that one believes that neither agent acted freely in the sense that is necessary for retributivist desert moral responsibility. In order to provide a compelling account of how a determined agent can exercise retributivist free will, it seems incumbent upon the compatibilist to explain how agents in the kind of deterministic universe portrayed above *can* be proper targets of retributivist treatment in the afterlife by a deity similar to that presented in the foregoing thought experiment.

3 Conclusion

Our primary goal in this paper has not been to argue against compatibilism but rather to clarify the kind of free will that is of central philosophical and practical importance in the debate. We have argued that there are strong reasons why philosophers ought to adopt an understanding of free will as the control in action

[8] It is worth stressing here another advantage of using punishment in the afterlife but not eternal torment in hell (as Strawson does). It is possible that one might have generally compatibilist intuitions and yet recoil at the thought of punishing someone with eternal torment in hell. We cannot be certain, then, that if one does not believe that an agent who has committed a moral wrong deserves eternal torment, this is because they have incompatibilist intuitions. It might just as well be the case that they do not believe *anyone* should deserve that much torment. Letting the punishment in the afterlife include any difference in treatment between Maya and Marina, as our view does, solves this problem.

required for *retributivist desert moral responsibility*. Adopting such a definition, we have argued, would not only avoid verbal disagreements such as those that arise with Dennett's use of *desert*, it would also preserve the most important philosophical and practical aspects of the debate. In the final analysis, we maintain that the burden is on compatibilists to explain how determined agents can be legitimate targets of the desert-based judgments, attitudes, and treatments associated with retributivist free will.

Acknowledgments The authors would like to send their sincere gratitude to Alfred Mele, Manuel Vargas, Bruce Waller, and two anonymous referees for the helpful feedback they provided during the writing of this essay.

References

Berman, M. N. (2008). Punishment and justification. *Ethics, 18*, 258–290.
Bok, H. (1998). *Descartes: His moral philosophy and psychology*. New York: New York University Press.
Campbell, C. A. (1957). *On Selfhood and Godhood*. London: Allen & Unwin.
Carey, J. M., & Paulhus, D. L. (2013). Worldview implication of believing in free will and/or determinism: Politics, morality, and punitiveness. *Journal of Personality, 81*(2), 130–141.
Caruso, G. D. (2012). *Free will and consciousness: A determinist account of the illusion of free will*. Lanham, MD: Lexington Books.
Caruso, G. D. (2016a). Free will skepticism and criminal behavior: A public health-quarantine model. *Southwest Philosophy Review, 32*(1), 25–48.
Caruso, G. D. (2016b). Review of Bruce Waller's restorative free will. *Notre Dame Philosophical Review*.
Clark, T. (2012). Clark comments on Dennett's review of Against Moral Responsibility. Naturalism.Org: http://www.naturalism.org/Wallerexchange.htm#Clark.
Clarke, R. (2005). An argument for the impossibility of moral responsibility. *Midwest Studies in Philosophy, 29*, 13–24.
Crick, F. (1994). *Astonishing hypthesis: The scientific search for the soul*. New York: Scribner.
Dennett, D. C. (2003). *Freedom evolves*. New York: Viking.
Dennett, D. C. (2012). Review of against moral responsibility by Bruce Waller and Dennett's Rejoinder to Clark. Naturalism.org: http://www.naturalism.org/DCDWallerreview.htm.
Feinberg, J. (1970). *Doing and deserving*. Princeton, NJ: Princeton University Press.
Fischer, J. M. (1986). *Moral responsibility*. Ithaca, NY: Cornell University Press.
Fischer, J. M. (1994). *The metaphysics of free will: A study of control*. Hoboken, NJ: Wiley.
Fischer, J. M. (2007). Compatibilism. In J. Fischer, R. Kane, D. Pereboom, & M. Vargas (Eds.), *Four views on free will* (pp. 44–84). Hoboken, NJ: Wiley.
Fischer, J. M. (2013). My compatibilism. In P. Russell & O. Deery (Eds.), *The philosophy of free will* (pp. 96–317). Oxford: Oxford University Press.
Fischer, J. M., & Mark, R. (1998). *Responsibility and control: A theory of moral responsibility*. Cambridge: Cambridge University Press.
Fischer, J., Kane, R., Pereboom, D., & Vargas, M. (2007). *Four views on free will*. Hoboken, NJ: Wiley.
Frankfurt, H. (1971). Freedom of the will and the concept of a person. *Journal of Philosophy, 68*(1), 5–20.
Jeppsson, S. (2016). Accountability, answerability, and freedom. *Social Theory and Practice, 42*(4), 681–705.
Levy, N. (2011). *Hard luck: How luck undermines free will and moral responsibility*. New York: Oxford University Press.
Levy, N. (2012). Skepticism and sanctions: The benefits of rejecting moral responsibility. *Law and Philosophy, 31*, 477–493.
McKenna, M. (2008). Ultimacy and sweet Jane. In N. Trakakis & D. Cohen (Eds.), *Essays on free will and moral responsibility*. Newcastle upon Tyne: Cambridge Scholars Publishing.

McKenna, M. (2009). Compatibilism & desert: Critical comments on four views on free will. *Philosophical Studies, 144*(1), 3–13.
Mele, A. (1995). *Autonomous agents*. New York: Oxford University Press.
Mele, A. (2001). Acting intentionally: Probing folk notions. In B. Malle, L. Moses, & D. Baldwin (Eds.), *Intentions and intentionality: Foundations of social cognition* (pp. 27–43). Cambridge, MA: MIT Press.
Morris, S. G. (2015). *Science and the end of ethics*. New York: Palgrave Macmillan.
Nahmias, E. (2011). Intuitions about free will, determinism, and bypassing. In R. Kane (Ed.), *The Oxford handbook of free will* (2nd ed., pp. 555–575). Oxford: Oxford University Press.
Nahmias, E. (2014). Response to Misirlisoy and Haggard and to Bjornsson and Pereboom. In W. Sinnot-Armstrong (Ed.), *Moral psychology: Free will and moral responsibility* (pp. 43–58). Cambridge, MA: MIT Press.
O'Connor, T. (2000). *Persons and causes: The metaphysics of free will*. New York: Oxford University Press.
Paulhuse, D. L., & Carey, J. M. (2011). The FAD–Plus: Measuring lay beliefs regarding free will and related constructs. *Journal of Personality Assessment, 93*, 96–104.
Pereboom, D. (2001). *Living without free will*. New York: Cambridge University Press.
Pereboom, D. (2012). Free will skepticism, blame, and obligation. In N. Tognazzini & D. J. Coates (Eds.), *Blame: Its nature and norms* (pp. 189–206). New York: Oxford University Press.
Pereboom, D., & Caruso G. D. (Forthcoming). Hard-incompatibilist existentialism: Neuroscience, punishment, and meaning in life. In G. D. Caruso & O. Flanagan (Eds.), *Neuroexistentialismism: Meaning, morals, and purpose in the age of neuroscience*. New York: Oxford University Press.
Richards, J. R. (2000). *Human nature after Darwin: A philosophical introduction*. New York: Routledge.
Scanlon, T. (1998). *What we owe to each other*. Cambridge: Harvard University Press.
Scanlon, T. (2009). *Moral dimensions*. Cambridge, MA: Harvard University Press.
Shariff, A. F., Greene, J. D., Karremans, J. C., Luguri, J., Clark, C. J., Schooler, J. W., et al. (2014). Free will and punishment: A mechanistic view of human nature reduces retribution. *Psychological Science*.
Strawson, P. (1962). Freedom and resentment. In G. Watson (Ed.), *Free will* (pp. 72–93). New York: Oxford University Press. **(reprinted)**.
Strawson, G. (1986). *Freedom and belief*. Oxford: Oxford University Press.
Strawson, G. (1994). The impossibility of moral responsibility. *Philosophical Studies, 75*(1), 5–24.
Van Inwagen, P. (1983). *An essay on free will*. Oxford: Clarendon Press.
Vargas, M. (2007). Revisionism. In J. Fischer, R. Kane, D. Pereboom, & M. Vargas (Eds.), *Four views on free will* (pp. 126–165). Hoboken, NJ: Wiley.
Vargas, M. (2009). Revisionism about free will: A statement and defense. *Philosophical Studies, 144*(1), 45–62.
Vargas, M. (2015). Desert, responsibility, and justification: A Reply to Doris, McGreer, and Robinson. *Philosophical Studies*.
Wallace, R. Jay. (1994). *Responsibility and the moral sentiments*. Cambridge, MA: Harvard University Press.
Waller, B. (2011). *Against moral responsibility*. Cambridge, MA: MIT Press.
Waller, B. (2015). *The stubborn system of moral responsibility*. Cambridge, MA: MIT Press.
Wolf, S. (2011). Blame, Italian style. In R. J. Wallace, R. Kumar, & S. Freeman (Eds.), *Reasons and recognition: Essays on the philosophy of T.M. Scanlon* (pp. 332–348). New York: Oxford University Press.