

Understanding Moral Responsibility within the Context of the Free Will Debate

Stephen G. Morris, The College of Staten Island/CUNY

Despite the extensive amount of academic papers that have been written on the subject of free will, philosophers have yet to settle on a single definitive notion of what free will is. To get a glimpse of the various senses in which philosophers use the term “free will,” one need only look at Timothy O’Connor’s informative entry for “Free Will” in the *Stanford Encyclopedia of Philosophy*. Here one can find a wide spectrum of free will concepts ranging from choosing on the basis of desires and values (a compatibilist conception) to being the ultimate source of one’s actions (an incompatibilist conception).¹ The difficulty of reaching a consensus on any of these disparate notions of free will is acknowledged by Manuel Vargas who recognizes that “It is not clear that there is any single thing that people have had in mind by the term ‘free will.’”² But if different philosophers are using “free will” in different senses, one might justifiably wonder whether many of the disputes about free will among philosophers are merely verbal. David Chalmers has recently addressed this worry by pointing out how the philosophical literature on the subject of free will is “beset by verbal disputes, in a fashion that is occasionally but too rarely recognized.”³ Concern about verbal disputes in the free will debate is especially legitimate when it comes to the debate between compatibilists and incompatibilists. This point is made salient by Daniel Dennett in his book *Freedom Evolves* in which he—a professed compatibilist who views his account of free will as offering a legitimate alternative to skepticism about free will—acknowledges that should a free will skeptic accept the kind of moral outlook that Dennett defends (namely, where moral responsibility is cashed out in terms of the propriety of consequentialist punishment/reward) the skeptic’s position would be “only terminologically different from *compatibilism*.”⁴ Dennett’s observation is important for two reasons. The first is that it is an acknowledgement how at least some of the disagreements between compatibilists and incompatibilists may amount to no more than verbal differences. Secondly, it alludes to the important relationship between the concepts of *free will* and *moral responsibility*.

While philosophers disagree about how the term “free will” is best understood, they generally agree that it is closely related to moral responsibility. More specifically, free will is considered to be a necessary condition—or, more specifically, the control condition—for moral responsibility.⁵ It is unsurprising, therefore, that questions concerning moral responsibility are at the

heart of the debates between compatibilists and incompatibilists as well as those between libertarians and their opponents. In fact, one might reasonably hold that when philosophers discuss free will, it is moral responsibility with which they are primarily concerned. As Galen Strawson has said, “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.”⁶ Given that free will is understood partly by the relation it holds to moral responsibility, it stands to reason that achieving a better understanding of free will requires that we have a clear idea of the sort of moral responsibility to which free will is thought to be connected. In this essay I argue that examining the substantive differences that exist between compatibilists and incompatibilists reveals a specific notion of moral responsibility that is best suited for philosophical debates regarding free will. It is the kind of moral responsibility that could justify doling out rewards and punishments to agents on strictly retributivist grounds.

In looking for a definitive notion of moral responsibility that should be adopted by each of the major factions involved in the free will debate (compatibilists, libertarians, and skeptics), a good first step is to recognize the close connection that exists between moral responsibility and reward or punishment. Many philosophers believe that to say that one is morally responsible for either a good or bad act is to say that one should be either rewarded or punished, respectively. John Stuart Mill, for example, advocated this type of view when he claimed that “We do not call anything wrong, unless we mean to imply that a person ought to be punished in some way or other for doing it...This seems to be the real turning point of the distinction between morality and simple expediency.”⁷ The view that moral responsibility and reward and punishment are tightly linked is also shared by many contemporary philosophers. This view is especially prevalent among incompatibilists, many of whom are motivated by the intuition that it is inappropriate to inflict certain types of rewards or punishments upon determined agents for their actions. Among the incompatibilists sharing this view is Richard Joyce, who claims that “embedded in the very idea of moral judgment is a reference to a special relation holding between transgressive acts and the negative responses that they provoke (and compliant acts and the rewards they provoke). I believe we do find such a relation centrally incorporated in moral systems: We think that certain behaviors *deserve* certain responses.”⁸

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 or punishment. Following the lead of Peter Strawson, one might maintain 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, it should be pointed out that Strawson (a compatibilist) believed that reward and punishment are intimately connected to

the reactive attitudes that are central to the notion of moral responsibility. This is apparent when he says, “the preparedness to acquiesce in that infliction of suffering on the offender which is an essential part of punishment is all of a piece with this whole range of [reactive] attitudes.”⁹ Notwithstanding this, one may argue for an altered version of Strawson’s view whereby being an appropriate target of the reactive attitudes associated with praise and blame is viewed as distinct from being an appropriate target for reward or punishment. For one subscribing to this view, it seems incumbent upon her to explain what it means to say that “one is an appropriate target for the reactive attitudes” if not partly that “it would be appropriate for one to be either rewarded or punished.” And 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 any genuine debate that exists between compatibilists and incompatibilists. As I discuss in more detail later, a key intuition motivating incompatibilists—and hence, one that drives the primary controversies surrounding free will—is that it is in some sense inappropriate 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. Likewise, the incompatibilist may allow that determined agents can exercise a kind of free will that grounds this type of moral responsibility. 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 seem to bypass a primary issue upon which the incompatibilist differentiates her position from that of the compatibilist.

At this point, the question that arises concerns the specific type of reward/punishment that is justified by virtue of being morally responsible. It seems clear from the incompatibilist’s perspective that holding one morally responsible for, say, a bad act in the context of the free will debate cannot simply mean that it is appropriate to punish the agent on consequentialist grounds. If this were all there was to the concept of moral responsibility—and the concept of free will that is based on it—then it would be difficult to understand what all the bickering is about. After all, the incompatibilist can allow that there are legitimate consequentialist reasons for punishing determined agents. Hence, if an ascription of moral responsibility did no more than either affirm or allow for the propriety of rewarding/punishing a good/bad agent on consequentialist grounds, 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 incompatibilist notions of free will and moral responsibility.

A revealing perspective on what incompatibilists take moral responsibility to mean is offered by the free will skeptic Derk Pereboom, who characterizes it as follows:

...for an agent to be morally responsible for an action is for it to belong to her in such a way that she would deserve blame if she understood that it was morally wrong, and she would deserve credit or perhaps praise if she understood that it was morally exemplary. The desert sense at issue here is basic in the sense that the agent, to be morally responsible, would deserve blame or credit just because she has performed the action...and not by virtue of consequentialist considerations.¹⁰

Pereboom goes on to claim that, “The notion [of moral responsibility] that incompatibilists do claim to be at odds with determinism is rather the one defined in terms of basic desert.”¹¹ Pereboom’s statements here can be taken as an effort to clarify what the issue really is between incompatibilists and compatibilists. The issue boils down to whether or not determined beings can be morally responsible in the basic desert sense. Where one is willing to grant that this kind of moral responsibility is impossible for determined beings, Pereboom would not view this individual as being genuinely opposed to incompatibilism.

Pereboom’s discussion of the compatibilist/incompatibilist debate goes a long way towards clarifying the kind of moral responsibility that is under dispute and, in doing so, helps us identify when apparent disagreements about determinism and free agency are best construed as verbal. Disagreements between compatibilists and incompatibilists should be considered verbal whenever compatibilists use the term “moral responsibility” in a way other than that which would enable an agent to be responsible in Pereboom’s basic desert sense. It should be mentioned here that Pereboom is not alone in thinking that the endless debates between compatibilists and incompatibilists boil down to whether determined agents can be responsible in the basic desert sense. Nor is this something that only incompatibilists recognize. Michael McKenna, a prominent compatibilist, has said that, “it is easy to see where all parties are prepared to draw their lines in the sand, since it seems that what most everyone is hunting for...is the sort of moral responsibility that is desert-entailing.”¹² Recognizing the desert element at play in discussions about free will and moral responsibility is certainly useful for helping us understand what the differences are between incompatibilists and their opponents. But in order to make substantial progress towards resolving this dispute, we need to pin down exactly what it means to say that one can be morally responsible in the “basic desert sense.” For McKenna’s part, he admits that he has “no clear sense of what anyone in the free will debate means by desert.”¹³ And yet he recognizes the importance of gaining clarity on this issue when he states that “it is here [that is, the ‘unsettled questions about the nature

of desert'] where the next important controversies in the free will and moral responsibility debate will occur."¹⁴

Despite McKenna's suggestion that all parties acknowledge the centrality of desert-based moral responsibility to the primary philosophical disputes regarding free will, I would point out that many, if not most, compatibilists believe that it is possible to establish that free will exists without having to address the issue of desert. Robert Kane points out that a strain of compatibilism—often called “classical compatibilism”—holds that to be free means no more than “(1) to have the power or ability to do what you will (desire or choose or try) to do, and this entails (2) an absence of constraints or impediments [such as a mental impairment like kleptomania or being held back by an external force] preventing you from doing what you will (desire or choose or try) to do.”¹⁵ Since the classical compatibilist believes that satisfying conditions (1) and (2) are sufficient for free will, the issue of desert need never arise. While some of the most famous proponents of classical compatibilism are historical figures like David Hume and Thomas Hobbes, some more contemporary philosophers have adopted compatibilist positions that adhere quite closely to classical compatibilism. Kai Nielsen, for example, has argued that what really matters in terms of freedom and moral responsibility is “Freedom of Conduct” which he defines as, “the ability and opportunity to do what one wants to do and act in accordance with one's own rational deliberations, without constraint and compulsion.”¹⁶ Nielsen holds that this kind of important freedom is compatible with determinism. While more complex than the classical compatibilism of Hume or Hobbes (or even Nielsen's compatibilist position), Harry Frankfurt's compatibilist view—perhaps the most cited and influential version of compatibilism to date—sticks closely to the classical model's emphasis on being able to act upon one's deliberations without constraint. The important thing for Frankfurt is that one be able, roughly, to have one's actions conform to one's more reflective desires. When this occurs, Frankfurt believes that one can be said to have freedom of the will, even if the conformity of one's actions to one's higher-order desires is the result of deterministic processes. It should be mentioned that the issue of desert does not appear in the most well-known of the essays in which Frankfurt lays out his conditions for freedom of the will.¹⁷ Lest one think that compatibilist accounts in which discussions concerning desert, reward or punishment do not figure prominently are more or less absent in the contemporary literature, it is worth noting that Eddy Nahmias is a prominent compatibilist who typically foregoes discussing desert and punishment/reward in making his case. In an essay that recently appeared in the *The New York Times*, Nahmias discusses the basics of his view which bears many of the hallmarks of classical compatibilism:

Many philosophers, including me, understand free will as a set of capacities for imagining future courses of action, deliberating about one's reasons for choosing them, planning one's actions in light of this deliberation and controlling actions in the face of competing desires. We act of our own free will to the extent that we have the opportunity to exercise these capacities, without unreasonable external or internal pressure.¹⁸

So while many philosophers agree that the concept of desert should figure prominently in discussions about free will, many others apparently disagree. And even where philosophers do acknowledge the centrality of desert to the free will debate, they rarely specify what they believe "desert" should be taken to mean. Recall McKenna's admission that he is at a loss as to what anyone in the free will debate means by "desert." The difficulty of settling on a single meaning of "desert" is illustrated by the ongoing debate between Derk Pereboom and Manuel Vargas. In response to Pereboom's assertion that the basic desert sense of moral responsibility is at the heart of the disagreements between incompatibilists and compatibilists, Vargas (who takes a revisionist position on free will that basically amounts to a kind of compatibilism) has responded by stating that, "I agree that this is exactly the sense of moral responsibility with which we should be concerned."¹⁹ Yet as Pereboom points out in a response paper to Vargas, it is not clear that they have the same thing in mind when they both claim that they are concerned with a kind of moral responsibility that is desert-entailing.²⁰ In what follows, I hope to provide a specific understanding of desert that I believe adequately captures the main substantive disputes between not only compatibilists and incompatibilists, but between libertarians and skeptics as well.

One of the few overt attempts at explaining something approximating the basic desert sense of moral responsibility has been made by Galen Strawson, who claims that "true moral responsibility is 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."²¹ 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.²² While this may be true, the idea of divine retribution in the afterlife seems a plausible way of understanding the commonsense notion of desert.²³ Following Strawson's lead, I believe that the idea of deserving retribution in the afterlife accurately captures the basic desert sense of moral responsibility that incompatibilists hold to be unavailable for determined agents.²⁴ Instead of deserving either *eternal* reward or punishment, however, the key kind of question is whether two ordinary (in terms of rationality, etc.) yet determined human agents—one of which intentionally performed a bad act while the other intentionally performed a good act—should receive *any different*

treatment from one another whatsoever from a divine all-knowing judge (who didn't necessarily create the agents) on the basis of these individual actions alone.²⁵ The response that one gives to this kind of thought experiment should reveal whether or not she shares incompatibilist intuitions. The purpose of invoking the notion of a divine judge in the afterlife is 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 the two agents in this case, therefore, would seem warranted only from a retributivist, and not a consequentialist, perspective. Furthermore, since it is plausible to assume that only incompatibilists would find it improper for the divine judge to dole out different treatments to the two agents on retributivist grounds, anyone who judges that differential treatment would be improper in the case of the two determined agents would appear to be expressing incompatibilist intuitions. Likewise, asserting that it would be proper for the divine judge to treat the agents differently would indicate that one shares the compatibilist view that determined agents can be proper targets of retributivist treatment.

One might worry here that the kind of thought experiment I have presented is open to a kind of interpretation whereby a reader may still assume that the divine judge's actions may have implications that are relevant from a consequentialist perspective. For instance, one might assume that the behavior of the divine judge can somehow influence subsequent behavior on earth. If this were the case then one's judgment that it would be proper for the divine being to treat the agents differently would not necessarily indicate that the individual in question agrees that it is sometimes appropriate to reward or punish determined agents on strictly retributivist grounds. While it may be impossible to construct the kind of thought experiment that I have proposed in such a way as to ensure that readers will not mistakenly see the divine judge's actions as having consequentialist implications beyond any immediate reward or punishment that is administered, it would probably not be difficult to structure the thought experiment so that such errors are likely to be rare. The best strategy may simply be to explicitly state that the divine judge's treatment of the two agents will have no impact on anyone else's behavior or welfare beyond whatever pleasure/pain is meted out directly in the form of the divine judge's punishment(s) and/or reward(s).²⁶ The kind of thought experiment I have constructed appears to call into play the key intuitions that have led incompatibilists to reject compatibilism. Hence, where there is a genuine disagreement between these two sides, it would seem to boil down to a simple question. Namely, would it be appropriate for the aforementioned divine judge to treat these determined agents differently, whereby one is either rewarded or punished more than the other? To make her case, the incompatibilist must show us why a negative answer to this question is appropriate, while the compatibilist has the opposite burden.²⁷

In what has preceded I have suggested that the incompatibilist's position is best understood as being predicated on the notion that it can never be proper to reward or punish a determined

agent on strictly retributivist grounds. This being the case, any assertion that determined agents can be either free or morally responsible in senses that are not founded on retributivism of this sort misses the point that the incompatibilist is trying to make and is, therefore, not really addressing whatever genuine differences may exist between compatibilists and incompatibilists. Furthermore, the sense of moral responsibility that is most relevant to the debate between compatibilists and incompatibilists—that is, the type connected to retributivist justice (call it “retributivist moral responsibility”)—seems to be driving the other primary discussion related to free will—namely, whether indeterminism can enable one to be genuinely free or morally responsible. It stands to reason that the main issue between libertarians and their skeptical opponents can be boiled down to the question of whether non-determined agents can ever be proper targets of retributivist reward or punishment. After all, if the sense of moral responsibility at work in the dispute between libertarians and skeptics was not the sort that grounds retributivist justice, it would be difficult to know what the dispute is about. Were the libertarians only claiming, for example, that non-determined agents can be responsible in the sense that they can sometimes be appropriate targets of consequentialist reward and punishment, it is hard to see who might object to them. At any rate, for anyone who denies that retributivist moral responsibility is at the center of the disputes between compatibilists and incompatibilists on the one hand, and libertarians and skeptics on the other, the burden is on them to provide an alternative account of moral responsibility that can contribute to a better explanation of what the main philosophical disputes regarding free will are about. Until such an account is given, I recommend that retributivist moral responsibility should be the notion of moral responsibility adopted by all participants in the primary discussions that have come to characterize the free will debate.

Beyond the relevance of retributivist moral responsibility to the substantive disputes between both incompatibilists and compatibilists on the one hand and libertarians and their opponents on the other, another good reason for situating retributivist moral responsibility front and center in philosophical discussions about free will is that retributive desert appears to play a fundamental role in the folk concept of moral responsibility that is tightly connected with commonsense attitudes about free will. Richard Joyce makes an astute observation when he points out that “When we examine our ordinary concepts of desert and justice, what we seem to find is an idea of the world having a kind of ‘moral equilibrium.’ When a wrong is done this equilibrium is upset, and the administration of the appropriate punishment is seen as the procedure that will affect its restitution.”²⁸ The prominent role that retributive desert plays in commonsense attitudes about moral responsibility and free will is reflected in the laws that dictate how criminals are to be treated in Western democracies like the United States. In supporting his view that commonsense notions about free will in the United States are closely connected with retributivist attitudes about justice,

neuroscientist Sam Harris has pointed out how “The Supreme Court has called free will a ‘universal and persistent’ foundation for our system of law distinct from ‘a deterministic view of human conduct that is inconsistent with the underlying precepts [many of which are retributive] of our criminal justice system.’”²⁹ Given the important role that retributivist moral responsibility plays in incompatibilist and libertarian arguments about free will, as well as how retributive desert figures prominently in both common attitudes about moral responsibility and systems of criminal justice in Western democracies—both of which tend to appeal to the notion of free will for their philosophical justification—there are strong reasons for why philosophers ought to employ the retributivist sense of moral responsibility in their discussions about free will.

In responding to my claim that retributivist moral responsibility is at the center of the main philosophical disagreements surrounding the free will, one might argue that the connection between free will and retributive desert is not as straightforward as I imply since one can believe in free will without accepting that retributivism is ever warranted. I suspect that those taking this view are likely to subscribe to the existence of free will along the following lines: Since I perform certain actions that are the result of conscious choices made by me and by no one (or nothing) else, and since I am not unduly coerced in the performance of these actions, it is accurate to say that I sometimes exercise free will in—and am sometimes morally responsible for—the performance of my actions even though such freedom/moral responsibility could never render me an appropriate target for retributivist treatment. Now this is certainly a coherent position that people—especially compatibilists—can and often do take. It is essentially the classical version of compatibilism discussed earlier and is more or less the compatibilist position taken by Eddy Nahmias. If all someone means by “free will” is the power to perform actions brought about by conscious decisions in the manner described above—let “weak compatibilism” refer to the view that determined agents can exercise free will of this variety—then I would expect that virtually all incompatibilists would agree that the typical human adult, whether determined or not, is capable of exercising free will in this sense. But herein lies the problem. The fact that the vast majority of incompatibilists do not reject the ability captured by the term “free will” in this weak sense shows that this is clearly not the sense of free will that incompatibilists are rejecting so vehemently. What I am claiming the incompatibilist rejects is that we possess what I have called “retributivist moral responsibility” as well as the more robust type of free will that is associated with it (call this “retributivist free will”).

Perhaps an advocate of weak compatibilism would say that although we cannot have retributivist moral responsibility, we can nonetheless possess moral responsibility of an important type—say, the type that grounds consequentialist justice.³⁰ While this is certainly a move that is open to the weak compatibilist, it would—for reasons explained earlier—appear to eliminate any substantive differences between the weak compatibilist’s position and standard incompatibilism.

And it is clear that acknowledging that human beings possess moral responsibility of this variety would not be enough to satisfy libertarians. Beyond the fact that the weak compatibilist notion of free will under consideration seems ill-suited for helping resolve any of the primary substantive disputes in the free will debate, it also fails to track commonsense views about free will insofar as such views often serve as a justification for retributivist reward and punishment. As discussed earlier, the fact that both the general public, in Western democracies at least, as well as the criminal justice systems in such societies understand free will as providing justification for the retributivist model of justice suggests that the concept of free will at play in both folk attitudes and legal institutions in the West is different from the type of free will espoused by weak compatibilism, which is only meant to provide a justification for a weaker type of moral responsibility which focuses on consequences as opposed to retributivism. The upshot is that since weak compatibilism does not appear to be at odds with any of the primary positions taken in the free will debate, since it is disconnected from the retributivist elements inherent in commonsense notions of free will, and since it is not equipped to play the role that the concept of free will has traditionally played in Western criminal justice systems, its relevance to contemporary philosophical discussions seems minimal at best.

In this essay I have argued that philosophers ought to adopt the retributivist usage of the term “moral responsibility” (as well as the retributivist sense of free will that gives rise to it) since, as I see it, the issue of whether or not human beings are capable of possessing this type of moral responsibility lies at the heart of the primary substantive disagreements between incompatibilists and compatibilists opponents on the one hand, and libertarians and skeptics on the other. If the arguments given in this essay are correct, it follows that some so-called compatibilist accounts of free will are not really at odds with standard incompatibilism. These include accounts like that of Dennett’s in *Freedom Evolves*, where the kind of free will argued for is associated with a consequentialist, as opposed to retributivist, notion of moral responsibility. These also include accounts that subscribe to what I have called “weak compatibilism.” In fact, there does not seem to be much in the way of opposition in the mainstream philosophical literature to the substantive claims made by weak compatibilism. Since few involved in the free will debate would challenge that we sometimes exercise the kind of free will espoused by weak compatibilists, philosophers who speak of “free will” in the weak compatibilist sense run the risk of perpetuating the verbal disagreements that have obfuscated the important substantive issues in the free will debate.

To conclude, I would like to briefly touch on what the implications might be were philosophers to agree that discussions about free will should focus on whether, and if so under what circumstances, people are capable of being morally responsible in the retributivist sense. Perhaps the most important implication would be that the burden of proof in the free will debate would appear to have shifted onto compatibilists to explain how a determined agent could warrant treatment that

is strictly retributive in nature. After all, there is good reason to think that most people would share the incompatibilist intuition that it would be inappropriate to subject an agent to retributivist treatment when the agent's actions are predetermined before he is even born. The intuition that retributivist justice is inapplicable for determined agents is so strong that it suggests that most compatibilists have something other than retributivist free will in mind when they assert that determined agents are capable of possessing free will. If it is true that most people would intuitively agree that it is improper to subject determined agents to retributivist treatment, this would provide the basis of an incompatibilist defense against Nahmias and his colleagues, who claim that the incompatibilist shoulders more of a philosophical burden than the compatibilist since the incompatibilist must explain why we should accept her more "metaphysically demanding" notion of free will that requires that indeterminism be true.³¹ An incompatibilist could respond by saying that the more demanding notion of free will that she has adopted is justified in light of how the metaphysically weaker notion of free will that the compatibilist accepts—i.e., free will that does not require determinism to be false—seems intuitively mistaken given that it is retributivist free will and moral responsibility that concerns us. It should be mentioned that while Nahmias et al.'s position is motivated by their experimental surveys that elicited evidence that folk intuitions about free will and moral responsibility are amenable to compatibilism, their surveys did not solicit responses relating in any direct way to retributivism. Assuming that I am correct in arguing that it is retributivist free will that ought to be the sense of free will adopted by participants in the free will debate, one could call the conclusions that Nahmias et al. drew by way of their surveys into question since there is no guarantee that their surveys provide insight into folk intuitions about retributivist free will. Contrary to the conclusions of Nahmias and his associates, an incompatibilist could argue that it is the compatibilist who carries the greater philosophical burden since the folk share incompatibilist intuitions about retributivist free will. Of course, a compatibilist might argue that the folk would agree that retributivist treatment is sometimes warranted for determined agents. The best way to address this issue may be to follow the experimental methodology of Nahmias et al. by testing whether the folk really do hold incompatibilist attitudes about retributivist free will. Should this be an approach worth taking, the kind of thought experiment that I provided earlier seems well-suited for this task.

Notes

¹ *Compatibilists* hold free will to be compatible with the truth of causal determinism. *Incompatibilists* believe that free will is not possible if causal determinism is true. *Libertarians* are incompatibilists who maintain that human beings are capable of exercising free will. *Skeptics* are incompatibilists who deny the possibility that human beings can exercise free will.

² Manuel Vargas, "Revisionism," in *Four Views on Free Will*, edited by John Martin Fischer, Robert Kane, Derk Pereboom, and Manuel Vargas (Malden, Mass.: Blackwell, 2007) 128-129.

³ See David Chalmers, "Verbal Disputes," *The Philosophical Review* 120 (2011): 515-566, footnote 12.

⁴ Daniel Dennett, *Freedom Evolves* (New York: Viking, 2003) 97-98.

⁵ The notion of control inherent to the prevailing notion of free will basically means that an action, decision, etc., over which an agent can properly be said to exhibit free will can trace its origins in some important way to the agent herself.

⁶ Galen Strawson, "The Impossibility of Moral Responsibility," *Philosophical Studies* 75.1-2 (1994): 8.

⁷ John Stuart Mill, *Utilitarianism* (Sioux Falls, SD: NuVision, 2007) 49.

⁸ Richard Joyce, *The Evolution of Morality* (Cambridge, MA: MIT, 2006) 66-67. Other incompatibilists taking a similar view include Galen Strawson, Randolph Clarke, and Derk Pereboom among others.

⁹ Peter F. Strawson, *Freedom and Resentment and Other Essays* (London: Methuen, 1974) 23. It should be mentioned here that Strawson would probably reject the whole project of seeking to define "moral responsibility" within the context of the free will debate since he rejected a main premise on which the debate is founded. Namely, that moral responsibility is open to metaphysical justification. Given that this essay is aimed at analyzing moral responsibility strictly within the context of traditional philosophical disputes concerning free will, however, I will simply assume that Strawson is wrong and that traditional compatibilists and incompatibilists are correct in thinking that certain metaphysical claims have to hold in order for agents to be morally responsible.

¹⁰ Derk Pereboom, "Hard Incompatibilism," in *Four Views on Free Will*, 86.

¹¹ Pereboom, "Hard Incompatibilism," 86.

¹² Michael McKenna, "Compatibilism & Desert: Critical Comments on *Four Views on Free Will*," *Philosophical Studies* 144.1 (2009): 9.

¹³ McKenna, "Compatibilism & Desert," 9.

¹⁴ McKenna, "Compatibilism & Desert," 10.

¹⁵ Robert Kane, "Introduction: The Contours of Contemporary Free Will Debates," in *The Oxford Handbook of Free Will*, edited by Robert Kane (Oxford: Oxford UP, 2011)12.

¹⁶ Kai Nielsen, "The Compatibility of Freedom and Determinism," in *Free Will*, edited by Robert Kane (Malden, MA: Blackwell, 1971) 41.

¹⁷ Harry Frankfurt, "Freedom of the Will and the Concept of a Person," *Journal of Philosophy* 68.1 (1971): 5-20.

¹⁸ Eddy Nahmias, "Is Neuroscience the Death of Free Will?" *The New York Times*. November 13, 2011. Available at: <http://opinionator.blogs.nytimes.com/2011/11/13/is-neuroscience-the-death-of-free-will/>. By "control" Nahmias means basically the ability to follow through with the performance of an action that the agent decided to be best after deliberating. For a more in-depth discussion of Nahmias's compatibilist account, see "Close Calls and the Confident Agent," *Philosophical Studies* 131.3 (2006): 627-667.

¹⁹ Vargas, "Revisionism," 210.

²⁰ See Derk Pereboom, "Hard Incompatibilism and Its Rivals," *Philosophical Studies* 144.1 (2009): 26.

²¹ Strawson, "The Impossibility of Moral Responsibility," 9.

²² By *retributivism*, I am referring roughly to the justification for treatment whereby an individual is either rewarded or punished as payback for the moral rights/wrongs he has committed. Consequentialist considerations do not figure into justifications for treatment from this perspective.

²³ Incidentally, Pereboom also suggests that the basic desert sense of moral responsibility is closely connected with retributive attitudes. However, the link between them is not spelled out by him in detail. I want to suggest that to believe that one is morally responsible in the basic desert sense *just is* to harbor a retributivist attitude towards the individual in question (see below).

²⁴ In stating this I am not claiming that the position taken by incompatibilists either is, or needs to be, motivated by the belief that an afterlife actually exists.

²⁵ Where the actions in question are not the products of hypnosis, mental illness, coercion by an external agent, or other factors that are commonly agreed upon as precluding free will.

²⁶ Ruling out further consequentialist implications in this explicit way would seem to eliminate the need for adding specific stipulations such as, "No one other than the divine judge and the two agents will ever receive knowledge of the divine judge's actions in this instance," "Upon receiving whatever punishment/reward the two agents receive, the two agents cease to exist," etc.

²⁷ It may be that the responses elicited through the kind of thought experiment I have suggested will be tainted by certain people's religious commitments. Hence, it may be best to alter the scenario in such a way that the divine judge is replaced by a mortal judge who is charged with rendering a decision that is not based on any consequentialist considerations.

²⁸ Joyce, *The Evolution of Morality*, 68.

²⁹ Sam Harris, *The Moral Landscape* (New York: Free Press, 2010) 106. The passage cited by Harris appears in the U.S. Supreme Court's decision in *U.S. v Grayson* (1978).

³⁰ It should be pointed out that at least some compatibilists seem to hold that retributivist forms of free will and moral responsibility are compatible with determinism. For example, see John Martin Fischer, "Compatibilism," in *Four Views on Free Will*, especially page 82. What's more, Pereboom considers Fischer's compatibilist view—with its preservation of retributivist elements—to be "the paradigmatic compatibilist position on the sort of responsibility we have" ("Hard Incompatibilism," 202).

³¹ Eddy Nahmias, S. Morris, T. Nadelhoffer, and J. Turner, "Surveying Freedom: Folk Intuitions about Free Will and Moral Responsibility," *Philosophical Psychology* 18.5 (2005): 572.

Bibliography

Chalmers, David. "Verbal Disputes." *The Philosophical Review* 120 (2011): 515-566.

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

Dennett, Daniel. *Freedom Evolves*. New York, NY: Viking, 2003.

Fischer, John Martin. "Compatibilism." In *Four Views on Free Will*. Eds. John Martin Fischer, Robert Kane, Derk Pereboom, and Manuel Vargas. Malden, Mass.: Blackwell, 2007, 44-84.

Frankfurt, Harry. "Freedom of the Will and the Concept of a Person." *Journal of Philosophy* 68.1 (1971): 5-20.

Harris, Sam. *The Moral Landscape*. New York: Free Press, 2010.

Joyce, Richard. *The Evolution of Morality*. Cambridge, MA: MIT, 2006.

Kane Robert. "Introduction: The Contours of Contemporary Free Will Debates." In *The Oxford Handbook of Free Will*. Ed. Robert Kane. Oxford: Oxford UP, 2011, 3-44.

McKenna, Michael. "Compatibilism & Desert: Critical Comments on *Four Views on Free Will*." *Philosophical Studies* 144.1 (2009): 3-13.

Mill, John Stuart. *Utilitarianism*. Sioux Falls, SD: NuVision, 2007.

Nielsen, Kai. "The Compatibility of Freedom and Determinism." In *Free Will*. Ed. Robert Kane. Malden, MA: Blackwell, 1971, 34-46.

Nahmias, Eddy. "Close Calls and the Confident Agent." *Philosophical Studies* 131.3 (2006): 627-667.

Nahmias, Eddy. "Is Neuroscience the Death of Free Will?" *New York Times*. November 13, 2011. Available at: <http://opinionator.blogs.nytimes.com/2011/11/13/is-neuroscience-the-death-of-free-will/>.

Nahmias, Eddy, Morris, S., Nadelhoffer, T. and J. Turner. "Surveying Freedom: Folk Intuitions about Free Will and Moral Responsibility." *Philosophical Psychology* 18.5 (2005): 561-584.

O'Connor, Timothy. "Free Will." *The Stanford Encyclopedia of Philosophy* (2010). Available at: <http://plato.stanford.edu/entries/freewill/>.

Pereboom, Derk. "Hard Incompatibilism." In *Four Views on Free Will*. Eds. John Martin Fischer, Robert Kane, Derk Pereboom, and Manuel Vargas. Malden, Mass.: Blackwell, 2007, 85-125.

Pereboom, Derk. "Hard Incompatibilism and Its Rivals." *Philosophical Studies* 144.1 (2009): 21-33.

Strawson, Galen. "The Impossibility of Moral Responsibility." *Philosophical Studies* 75.1-2 (1994): 5-24.

Strawson, Peter F. *Freedom and Resentment and Other Essays*. London: Methuen, 1974.

Vargas, Manuel. "Revisionism." In *Four Views on Free Will*. Eds. John Martin Fischer, Robert Kane, Derk Pereboom and Manuel Vargas. Malden, Mass.: Blackwell, 2007, 126-165.

This collection of essays has its roots in a conference on free will and moral responsibility held at Monash University in November 2005, though only a few of the papers presented at the conference have made it into the current volume. We would like to thank both the participants at this conference and the contributors to this volume, as well as Cambridge Scholars Publishing for inviting us to put the collection together. "Understanding Moral Responsibility within the Context of the Free Will Debate." *Florida Philosophical Review*, 12(1): 68-82, 2012. "Preserving the Concept of Race: A Medical Expedient, a Sociological Necessity." "The Evolution of Cooperative Behavior and its Implications for Ethics." *Philosophy of Science*, 76(5): 915-926, 2009. "The Impact of Neuroscience on the Free Will Debate." *Florida Philosophical Review*, 9(2): 56-78, 2009. "Towards a More Empirically Informed Ethics." *Review Journal of Political Philosophy*, 7: 142-180, 2009. "Neuroscience and the Free Will Conundrum." *The American Journal of Bioethics*, 7(5): 20-22, 2007. "Canada's Assisted Human Reproduction Act: A Chimera of Religion and Politics." *The American Journal of Bioethics*, 7(2): 69-70, 2007. 2.1 Free Will and Moral Responsibility. 2.2 The Freedom to Do Otherwise. 2.3 Freedom to Do Otherwise vs. Sourcehood Accounts. One finds scholarly debate on the "origin" of the notion of free will in Western philosophy. (See, e.g., Dihle (1982) and, in response Frede (2011), with Dihle finding it in St. Augustine (354-430 CE) and Frede in the Stoic Epictetus (c. 55-c. 135 CE)). Epicurus has often been understood as seeking to ground the freedom of human willings in such indeterministic swerves, but this is a matter of controversy. If this understanding of his aim is correct, how he thought that this scheme might work in detail is not known.