

Optimistic Skepticism about Free Will

Derk Pereboom, Cornell University

penultimate draft

forthcoming in *The Philosophy of Free Will: Selected Contemporary Readings*, Paul Russell and Oisin Deery, eds., New York: Oxford University Press, 2013, pp. 421-49.

Introduction

One of the main concerns at stake in the historical free will debate is whether the sort of free will required for moral responsibility is compatible with the causal determination of our actions by factors beyond our control. Since Hume, the concern has prominently been extended to whether this sort of free will is compatible with indeterminacy in action. The position for which I argue is that free will, characterized in this way, is incompatible with this kind of causal determination, and with the type of indeterminacy of action that Hume had in mind. It is important to recognize that the term 'moral responsibility' is used in a variety of ways, and that the type of free will or control required for moral responsibility in several of these senses is uncontroversially compatible with the causal determination of action by factors beyond our control. But there is one particular sense of moral responsibility that has been at issue in the historical debate. It is this: 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 at issue here is basic in the sense that the agent, to be morally

responsible, would deserve the blame or credit just because she has performed the action, given sensitivity to its moral status, and not by virtue of consequentialist or contractualist considerations. Moral responsibility in this sense is presupposed by our retributive reactive attitudes, and it is thus the variety of moral responsibility that P. F. Strawson famously brings to the fore in his essay "Freedom and Resentment" (1962).

There are other senses of moral responsibility that have not been a focus of the free will debate. For example, an agent could be considered morally responsible if it is legitimate to expect her to respond to such questions as: "Why did you decide to do that? Do you think it was the right thing to do?" and to evaluate critically what her actions indicate about her moral character. Engaging in such interactions might well be reasonable in light of the way in which they contribute to our own and others' moral improvement (Arthur Kuflik, in conversation; for a related conceptions see Scanlon 1998 and Bok 1998). However, incompatibilists would not think that being morally responsible in this "answerability" sense is even *prima facie* incompatible with determinism. The type of moral responsibility that incompatibilists do claim not to be compatible with determinism is instead the sense characterized by basic desert and the reactive attitudes that presuppose it. From this point on, unless otherwise indicated, I will use the term 'moral responsibility' to refer to this particular type.

Spinoza (1677/1985: 440-44, 483-84, 496-97) maintained that due to very general facts about the nature of the universe we human beings lack the sort of free will required for moral responsibility. About this I think he is right. More specifically, he argues that it is because of the truth of causal determinism that we lack this sort of free will; he is thus a *hard determinist*. By

contrast, I am agnostic about the truth of causal determinism. I contend, like Spinoza, that we would not be morally responsible if determinism were true, but also that we would lack moral responsibility if indeterminism were true and the causes of our actions were exclusively states or events – this is the notion of indeterminacy of action that Hume arguably had in mind (1737/1978). For such indeterministic causal histories of actions would be as threatening to this sort of free will as deterministic histories are. However, it might be that if we were undetermined agent causes – if we as substances had the power to cause decisions without being causally determined to cause them – we would then have this type of free will. But although our being undetermined agent causes has not been ruled out as a coherent possibility, it is not credible given our best physical theories. Thus I do not claim that our having the sort of free will required for moral responsibility is impossible. Rather, I don't take a stand on whether it is possible or not. Nevertheless, since the only account on which we might in fact have this kind of free will is not credible given our best physical theories, we must take seriously the prospect that we are in fact not free in the sense required for moral responsibility. I call the resulting skeptical position *hard incompatibilism*. At the same time, I defend the optimistic view that conceiving of life without this type of free will would not be devastating to morality or to our sense of meaning in life, and in certain respects it may even be beneficial (for contrasting pessimistic views, see Smilansky 2000, and Russell 2000).

Furthermore, I reject an incompatibilism for which the availability of alternative possibilities is crucial to explaining moral responsibility, and accept instead an incompatibilism that ascribes the more significant role to an action's causal history. I argue that an agent's

moral responsibility for an action would be explained primarily by the action's having a causal history in which she is the source of her action in a specific way. I thus opt for *source* as opposed to *leeway* incompatibilism. Agent-causal libertarianism is commonly conceived as an incompatibilist position in which an agent can be the source of her action in the way required for moral responsibility, and as a result proponents of this view are typically source incompatibilists. However, one might also be a source incompatibilist and seriously doubt that we have the sort of free will required for moral responsibility, and this is the position I advocate (Pereboom 2001).

A manipulation argument against compatibilism

Defending hard incompatibilism requires facing up to compatibilism. I believe that the strongest argument against the compatibilist begins with the intuition that if someone is causally determined to act by other agents, for example, by scientists who manipulate her brain, then she is not morally responsible for that action (Taylor 1974, cf. Ginet 1990; Kane 1996; Mele 2006). The argument continues by showing that there are no differences between cases like this and otherwise similar ordinary deterministic examples that can justify the claim that while an agent is not morally responsible when she is manipulated, she can nevertheless be responsible in the ordinary deterministic examples. The non-responsibility intuition remains strong even if when manipulated the agent satisfies the conditions on moral responsibility advocated by the prominent compatibilist theories. My multiple-case argument first of all develops examples of actions that involve such manipulation and in which these compatibilist

conditions on moral responsibility are satisfied (1995, 2001). These cases, taken separately, indicate that it is possible for an agent not to be morally responsible even if the compatibilist conditions are satisfied, and that as a result these conditions are inadequate. But the argument has more force by virtue of setting out three such cases, each progressively more like a fourth, which the compatibilist might envision to be realistic, in which the action is causally determined in a natural way. An additional challenge for the compatibilist is to point out a relevant and principled difference between any two adjacent cases that would show why the agent might be morally responsible in the later example but not in the earlier one. I argue that this can't be done. I contend that the agent's non-responsibility generalizes from at least one of the manipulation examples to the ordinary case.

In the set-up, in each of the four cases Professor Plum decides to kill Ms. White for the sake of some personal advantage, and succeeds in doing so. The cases are designed so that his act of murder conforms to the prominent compatibilist conditions. This action meets certain conditions advocated by Hume: the action is not out of character, since for Plum it is generally true that selfish reasons typically weigh heavily -- too heavily when considered from the moral point of view; while in addition the desire that motivates him to act is nevertheless not irresistible for him, and in this sense he is not constrained to act (Hume 1739/1978). The action fits the condition proposed by Harry Frankfurt (1971): Plum's effective desire (i.e., his will) to murder White conforms appropriately to his second-order desires for which effective desires he will have. That is, he wills to murder her, and he wants to will to do so, and he wills this act of murder because he wants to will to do so. The action also satisfies the reasons-responsiveness

condition advocated by John Fischer and Mark Ravizza (1998): Plum's desires can be modified by, and some of them arise from, his rational consideration of the reasons he has, and if he knew that the bad consequences for himself that would result from killing White would be much more severe than they are actually likely to be, he would have refrained from killing her for that reason. Also, this action meets the condition advanced by Jay Wallace (1994): Plum has the general ability to grasp, apply, and regulate his actions by moral reasons. For instance, when egoistic reasons that count against acting morally are weak, he will typically regulate his behavior by moral reasons instead. This ability also provides him with the capacity to revise and develop his moral character over time, a condition that Alfred Mele emphasizes (1995). Now, supposing that causal determinism is true, is it plausible that Professor Plum is morally responsible for his action?

Each of the four cases I will now describe features different ways in which Plum's murder of White might be causally determined by factors beyond his control. In a first type of counterexample (Case 1) to these prominent compatibilist conditions, neuroscientists manipulate Plum in a way that directly affects him at the neural level, but so that his mental states and actions feature the psychological regularities and counterfactual dependencies that are compatible with ordinary agency (Pereboom, 2001: 121; McKenna 2008):

Case 1: A team of neuroscientists is able to manipulate Professor Plum's mental state at any moment through the use of radio-like technology. In this case, they do so by pressing a button just before he begins to reason about his situation. This causes Plum's reasoning process to be egoistic, which the neuroscientists know will deterministically

result in his decision to kill White. Plum does not think and act contrary to character since his reasoning processes are not infrequently egoistic. His effective first-order desire to kill White conforms to his second-order desires. The process of deliberation from which his action results is reasons-responsive; in particular, this type of process would have resulted in his refraining from killing White in some situations in which the reasons were different. Still, his reasoning is not in general exclusively egoistic, since he often regulates his behavior by moral reasons, especially when the egoistic reasons are relatively weak. He is also not constrained, in the sense that he does not act because of an irresistible desire – the neuroscientists do not induce a desire of this kind.

In Case 1, Plum's action satisfies all the compatibilist conditions we just examined. But intuitively, he is not morally responsible for the murder, because his action is causally determined by what the neuroscientists do, which is beyond his control. Consequently, it would seem that these compatibilist conditions are not sufficient for moral responsibility -- even if all are taken together.

This example might be filled out in response to those who have asked whether Plum in Case 1 (or in a previous version of this example) meets certain minimal conditions of agency because he is too disconnected from reality, or because he himself lacks ordinary agential control (Fischer 2004: 156; Mele 2005: 78; Baker 2006: 320; Demetriou 2010). This concern highlights the fact that in this example two desiderata must be secured at the same time: the manipulation must preserve satisfaction of intuitive conditions of agency, and it must render it plausible that Plum is not morally responsible. It turns out that these two desiderata can be met

simultaneously. Agency is regularly preserved in the face of certain involuntary momentary external influences. Finding out that the home team lost makes one irritable and more egoistic, and news of winning a prize more generous, but the conditions of agency remain intact. Still we commonly suppose that such influences are typically compatible with moral responsibility. However, we can imagine an egoism-enhancing momentary influence that preserves agency but plausibly does undermine moral responsibility. Suppose that by way of neural intervention the manipulators enhance Plum's disposition to reason self-interestedly at the requisite time, so that they know that it is causally ensured that he will decide to kill Ms. White (see also Shabo 2010: 376). Like finding out that the home team has lost, this intervention would not undermine Plum's agency, but intuitively it does render him non-responsible for his action.

Next consider a scenario more like the ordinary situation than Case 1:

Case 2: Plum is like an ordinary human being, except that neuroscientists have programmed him at the beginning of his life so that his reasoning is frequently but not always egoistic (as in Case 1), with the consequence that in the particular circumstances in which he now finds himself, he is causally determined to engage in the egoistic reasons-responsive process of deliberation and to have the set of first and second-order desires that result in his decision to kill White. Plum has the general ability to regulate his behavior by moral reasons, but in his circumstances, due to the egoistic character of his reasoning, he is causally determined to make his decision. The neural realization of his reasoning process and of the resulting decision is exactly the same as it is in Case 1

(although the external causes are different). At the same time, he does not act because of an irresistible desire.

Again, although Plum satisfies the compatibilist conditions, intuitively he is not morally responsible. So Case 2 also shows that the prominent compatibilist conditions, either individually or in conjunction, are not sufficient for moral responsibility. Moreover, it would seem unprincipled to claim that here, by contrast with Case 1, Plum is morally responsible because the length of time between the programming and the action is now great enough. Whether the programming occurs a few seconds before or forty years prior to the action seems irrelevant to the question of his moral responsibility. Causal determination by factors beyond his control plausibly explains Plum's not being morally responsible in the first case, and I think we are forced to say that he is not morally responsible in the second case for the same reason.

Imagine next a scenario more similar yet to the ordinary situation:

Case 3: Plum is an ordinary human being, except that he was causally determined by the rigorous training practices of his household and community in such a way that his reasoning processes are often but not exclusively rationally egoistic (as in Cases 1 and 2). This training took place when he was too young to have the ability to prevent or alter the practices that determined this aspect of his character. This training, together with his particular current circumstances, causally determines him to engage in the egoistic reasons-responsive process of deliberation and to have the first and second-order desires that result in his decision to kill White. Plum has the general ability to regulate his behavior by moral reasons, but in his circumstances, due to the egoistic nature of his

reasoning processing, he is causally determined make his decision. The neural realization of his reasoning process and of his decision is the same as it is in Cases 1 and 2. Here again his action is not due to an irresistible desire.

For the compatibilist to argue successfully that Plum is morally responsible in Case 3, he must adduce a feature of these circumstances that would explain why he is morally responsible here but not in Case 2. It seems there is no such feature. In all of these examples, Plum meets the prominent compatibilist conditions for morally responsible action, so a divergence in judgment about moral responsibility between these examples won't be supported by a difference in whether these conditions are satisfied. Causal determination by factors beyond Plum's control most plausibly explains the absence of moral responsibility in Case 2, and we should conclude that he is not morally responsible in Case 3 for the same reason.

Therefore it appears that Plum's exemption from responsibility in Cases 1 and 2 generalizes to the nearer-to-normal Case 3. Does it generalize to the ordinary deterministic case?

Case 4: Physicalist determinism is true – everything in the universe is physical, and everything that happens is causally determined by virtue of the past states of the universe in conjunction with the laws of nature. Plum is an ordinary human being, raised in normal circumstances, and again his reasoning processes are frequently but not exclusively egoistic (as in Cases 1-3). His decision to kill White results from his reasons-responsive process of deliberation, and he has the specified first and second-order desires. The neural realization of his reasoning process and decision is just as it is in

Cases 1-3. Again, he has the general ability to grasp, apply, and regulate his behavior by moral reasons, and it is not due to an irresistible desire that he kills White.

Given that we are constrained to deny moral responsibility in Case 3, could Plum be responsible in this ordinary deterministic situation? It appears that there are no differences between Case 3 and Case 4 that might justify the claim that Plum is not responsible in Case 3 but is in Case 4. In each of these cases Plum satisfies the prominent compatibilist conditions on moral responsibility. In each the neural realization of his reasoning process and decision is the same, although the causes differ. One distinguishing feature of Case 4 is that the causal determination of Plum's crime is not brought about by other agents (Lycan 1997). But the claim that this is a relevant difference is implausible. Imagine a further case that is exactly the same as Case 1 or Case 2, except that the Plum's states are induced by a spontaneously generated machine – a machine with no intelligent designer. Here also Plum would lack morally responsibility.

The best explanation for why the agent isn't responsible in these four cases is that he is causally determined by factors beyond his control in each. Because there is no difference between Cases 1 and 2, 2 and 3, and 3 and 4 that can explain in a principled way why he would not be responsible in the first of each pair but would be in the second, we are driven to the conclusion that he is not responsible in Case 4. The salient common factor in these cases that can plausibly explain why the agent is not responsible is that he is causally determined by factors beyond his control to act as he does. This is the best explanation for his non-responsibility in each of the cases. (See Todd 2011 for a important way to strengthen this

argument¹; for objections, see Fischer 2004, 2006; Mele 2005, 2006; Baker 2006; McKenna 2008, Demetriou 2010; Nelkin 2011; for further replies see Pereboom 2005, 2008a, 2008b. Thanks to all of these philosophers for motivating the revisions to the argument featured in this presentation).

Source incompatibilism

Why opt for a source as opposed to a leeway position? I argue that an example of the kind devised by Frankfurt supplies an effective challenge to the leeway position (Frankfurt 1969). In such examples an agent considers performing some action, but an intervener is concerned that she will not come through. So if the agent were to show some sign that she will not or might not perform the action, the intervener would arrange matters so that she would perform it anyway. Consider one of Fischer's examples: Jones will decide to kill Smith only if Jones blushes beforehand. Jones's failure to blush (by a certain time) can then function as the prior sign that would trigger the intervention that would cause her to kill Smith. Suppose that Jones acts without intervention. We might well have the intuition that she is morally responsible for killing Smith, although she could not have done otherwise than to kill Smith, and despite the fact that she could not even have formed an alternative intention. Jones could have

¹ Todd argues that proponents of manipulation arguments have assumed too heavy a burden: they do not need make it plausible that manipulated agents are not morally responsible, only that their responsibility is mitigated, for compatibilists will have as difficult a time explaining mitigation as they would explaining non-responsibility.

failed to blush, but Fischer contends that such a “flicker of freedom” is of no use to the libertarian, since it is not *robust* enough to play a part in grounding her moral responsibility (Fischer 1994, 131-59).

Here is a proposal for what is required for an alternative possibility to be robust:

Robustness: For agent to have a robust alternative to her immoral action A, that is, an alternative relevant per se to explaining why she is blameworthy for A, it must be that

(a) she instead could have voluntarily acted or refrained from acting as a result of which she would be blameless, and

(b) that for at least one such exempting acting or refraining, she is cognitively sensitive to its being available to her, with the result that she believes to some significant degree that had she voluntarily so acted or refrained she would be, or would likely be, blameless. (Pereboom 2012)²

The core intuition that underlies the proposal to ground moral responsibility in the accessibility of alternative possibilities is of the following sort: to be blameworthy for an action, the agent must have been able to do something that would have precluded her from being blameworthy for what she does, at least to the degree she’s blameworthy; she must have been able to do something that would have resulted in her being “off the hook” (Pereboom 2001: 1). Thus for

² Thanks to Dana Nelkin, James Hobbs, Jonathan Vance, and Kevin Timpe for comments that led to a revision of a previous version of this criterion (in Pereboom 2000, 2001).

an alternative possibility to be robust, it must first of all satisfy this condition: she could have willed something other than what she actually willed such that by willing it she would thereby have been precluded from the moral responsibility she actually has for the action (cf. Otsuka 1998). But in addition, the epistemic element of Robustness – that she must have been cognitively sensitive to the fact that by willing otherwise she would have been precluded from the responsibility she actually has – is motivated by the following sort of consideration.

Suppose that that the only way Joe could have avoided deciding to take an illegal deduction on his tax form -- a choice he does in fact make -- is by voluntarily taking a sip from his coffee cup, for unbeknownst to him, the coffee was laced with a drug that induces compliance with the tax code. In this situation, he could have behaved voluntarily in such a manner that would have precluded the choice for which he was in fact blameworthy, as a result of which he would have been morally non-responsible for it. But whether he could have voluntarily taken the sip from the coffee cup, not being cognitively sensitive to the fact that as a result of doing so he would have been blameless, is intuitively irrelevant to explaining whether he is morally responsible for his choice.

The most significant objection that has been raised against the earlier kinds of Frankfurt-style arguments, such as the one where blushing signals the act at issue, was initially suggested by Robert Kane and then systematically developed by David Widerker and Carl Ginet (Kane 1985: 51; 1996, 142-44, 191-92; Widerker 1995: 247-61; Ginet 1996). The general form of the Kane/Widerker/Ginet objection is this: in Frankfurt-style cases generally, the actual situation will feature a sign that occurs prior to the action at issue that signals the fact that

intervention is not required. If in the proposed case this prior sign causally determined the action, or if it were associated with some factor that did, the intervener's predictive ability could be explained. However, then the incompatibilist would not and could not be expected to have the intuition that the agent is morally responsible. But if the relationship between the prior sign and the action were not causally deterministic in such ways, then it will be the case that the agent could have done otherwise despite the occurrence of the prior sign. Either way, an alternative-possibilities condition on moral responsibility emerges unscathed.

I have proposed a type of Frankfurt-style case that avoids this objection (Pereboom 2000; 2001: 18-19; 2003, 2009a, 2012; see also Hunt 2000, 2005 for a similar example; see Fischer 2010 for an argument that the earlier cases are effective). Its distinguishing features are these: the cue for intervention must be a necessary condition for the agent's availing herself of any robust alternative possibility (without the intervener's device in place), while this cue is not itself a robust alternative possibility, and the absence of this cue at any specific time is not a sufficient condition for the agent's performing the action. Here is the most recent version of this example (Pereboom 2012):

Tax Cut: Jones can vote for or against a modest tax cut for those in his high-income group by pushing either the 'yes' or the 'no' button in the voting booth. Once he has entered the voting booth, he has exactly two minutes to vote, and a downward-to-zero ticking timer is prominently displayed. If he does not vote, he will have to pay a fine, substantial enough so that in his situation he is committed with certainty to voting (either for or against), and this is underlain by the fact that the prospect of the fine,

together with background conditions, causally determines him to vote. Jones has concluded that voting for the tax cut is barely on balance morally wrong, since he believes it would not stimulate the economy appreciably, while adding wealth to the already wealthy without helping the less well off, despite how it has been advertised. He is receptive and reactive to these general sorts of moral reasons: he would vote against a substantially larger tax cut for his income group on account of reasons of this sort, and has actually done so in the past. He spends some time in the voting booth rehearsing the relevant moral and self-interested reasons. But what would be required for him to decide to vote against the tax cut is for him to vividly imagine that his boss would find out, whereupon due to her political leanings she would punish him by not promoting him to a better position. In this situation it is causally necessary for his not deciding to vote for the tax cut, and to vote against it instead, that he vividly imagine her finding out and not being promoted, which can occur to him involuntarily or else voluntarily by his libertarian free will. Jones is sensitive to the fact that imagining this punishment scenario will put him in a motivational position to vote against. But so imagining is not causally sufficient for his deciding to vote against the tax cut, for even then he could still, by his libertarian free will, either decide to vote for or against (without the intervener's device in place). However, a neuroscientist has, unbeknownst to him, implanted a device in his brain, which, were it to sense his vividly imagining the punishment scenario, would stimulate his brain so as to causally determine the decision to vote for the tax cut. Jones's imagination is not exercised in this way, and he decides

to vote in favor while the device remains idle.³

³ As I argue in Pereboom (2012), this example can be embellished to answer one of Carl Ginet's (2002) objections to *Tax Evasion*, in particular to the (2000, 2001) version of this example. Ginet contends that at the precise time Joe makes the decision to take the illegal tax deduction, he might have been activating the necessary condition for refraining instead, and that this alternative possibility is robust: "for had J taken it, he would at t1 have been refraining from a willing – to do B [decide to take the illegal deduction] right then – such that by so refraining he would have avoided responsibility for doing B right then and would have been aware that he was avoiding responsibility for doing B right then (that being such an obvious implication of his not doing B right then, of which he of course would have been aware)." David Palmer (2011) and Christopher Franklin (2011a) also develop this kind of objection. Here, in essence, is my response. Imagine first an agent, Adam, who is causally determined to perform some immoral action during the time interval t0-t3, but the specific time during this interval he decides is up to him. Suppose he actually decides at t1. The incompatibilist has to agree that Adam is not blameworthy for making the decision at t1, but at best only responsible in some neutral sense for deciding then and not at some other time during the interval. And crucially, the reason the leeway incompatibilist must give for his not being blameworthy for deciding at t1 is that he has no (robust) alternative to making his decision by t3. According to the leeway incompatibilist it has to be the unavailability of some alternative possibility that explains why Adam is not blameworthy for his decision. In this case, causal determination is what excludes Adam's blameworthiness, and the leeway incompatibilist maintains that in general, causal

determination rules out blameworthiness because it precludes alternative possibilities. The only plausible candidate is the unavailability of an alternative to making the decision by t3, and thus on the leeway incompatibilist view, this unavailability would have to be sufficient for Adam's not being blameworthy at t1.

We can draw the following consequence for Jones's situation in *Tax Cut*. Suppose that Jones decides to vote for the tax cut at t1, a minute before the deadline t3. The leeway incompatibilist will not be able to defend the claim that Jones's deciding at t1 to vote in favor of the tax cut with the intervener's device in place is as blameworthy, and for the same reasons, as would be his deciding to vote in favor by t3 without the device in place. For with the device in place, the leeway incompatibilist cannot explain Jones's blameworthiness for making his decision at t1, but only his responsibility in a neutral sense for making the decision at t1 rather than at some other available instant. Although, as in Adam's situation, Jones does have an alternative to deciding at t1 -- for example, continuing to deliberate at t1 -- this will be insufficient to explain Jones's blameworthiness for making his decision at t1. For Jones has no robust alternative to making his decision by t3, and as in Adam's situation, for the leeway incompatibilist this will be sufficient for Jones's not being blameworthy for making his decision at t1. But it's nonetheless our strong intuition that Jones is blameworthy for deciding to vote in favor at t1, for which the leeway incompatibilist now has no explanation.

In this situation, Jones could be morally responsible – blameworthy, in particular -- for choosing to vote in favor of the tax cut by the deadline despite the fact that for this he has no robust alternative possibility.

This case does feature an alternative possibility that is accessible to the agent – Jones’s vividly imagining the punishment scenario. However, relative to his responsibility for deciding to vote in favor of the tax cut by the deadline, this alternative is not robust. First of all, absent the intervener’s device, it is not the case that by vividly imagining the scenario Jones would have avoided responsibility for deciding to vote in favor by the deadline. In these ordinary circumstances, this exercise of the imagination is compatible with his nevertheless being strongly inclined to vote for the tax cut, and indeed with actually deciding to vote in favor. Still, one might object, due to the intervener’s device, by vividly imagining the punishment scenario at some time during this interval Jones would have voluntarily done something whereby he would have avoided the blameworthiness he actually incurs. Had he exercised his imagination in this way, the device would have been activated, and he would not then have been blameworthy for deciding to vote in favor by the deadline. But Jones is not cognitively sensitive to the fact doing so would preclude him from responsibility for making the decision he does by the deadline. Moreover, he has no reason whatsoever to believe that the intervention would then take place and that as a result he would be precluded from this responsibility for his choice. Nevertheless, it remains intuitive that Jones is actually morally responsible for deciding to vote in favor of the tax cut by the deadline.

For proponents of the leeway position, the accessibility of alternative possibilities is

crucial to explaining why an agent would be morally responsible. The *Tax Cut* argument provides reason to reject the leeway view, and to affirm instead that moral responsibility would be explained primarily by the agent's being the action's source in the appropriate way. According to source incompatibilism, which I endorse, moral responsibility requires that the agent be the source of her action in a way incompatible with her being causally determined to act by factors beyond her control. It might well be that alternative possibilities -- not necessarily of the robust sort -- are entailed by her being the source of her action in this way (Pereboom 2001: 37, 2003: 197). But these alternative possibilities would not have the primary role in explaining an agent's moral responsibility. Rather, they would be a consequence of the factor that did: the agent's being the source of her action in the right way.

Against libertarianism

Defending hard incompatibilism also requires confronting libertarianism. Two contending versions of libertarianism are the event-causal and agent-causal types.⁴ In event-causal libertarianism, actions are caused solely by *events* -- such as *Joe's currently desiring to*

⁴ A third type is non-causal libertarianism, advocated by Ginet (1990), Hugh McCann (1998), and Stewart Goetz (2008). An often-cited objection to this type of libertarianism is that control in action is fundamentally a causal matter, and in particular such theories cannot secure the type of control in action required for moral responsibility (O'Connor 2000, Clarke 2003). But this position remains intriguing, especially given the sorts of concerns that arise for the other types of libertarianism.

receive a substantial tax refund, or Anne's currently believing that she can help someone in trouble. It is often assumed that all causation in the physical world is fundamentally by events, and not by things such as atoms, organisms, and agents, which we call *substances*. Although we might say, for example, that a missile – a substance – destroyed an airplane, when speak more accurately, the idea is that we should say instead that *the missile's hitting the airplane at noon yesterday* – an event – caused the destruction. If we are more precise about what it is in the physical world that causes effects, it turns out to be events, not substances. In solidarity with this position, event-causal libertarians contend that actions are caused solely by events, and indeterminacy in the production of actions by appropriate events is a highly significant requirement for moral responsibility (Kane 1996; Ekstrom 2000; Balaguer 2009, Franklin 2011b).

According to agent-causal libertarianism, free will of the sort required for moral responsibility is accounted for by the existence of agents who possess a causal power to make choices without being determined to do so (Taylor 1966, 1974; Chisholm 1976; O'Connor 2000; Clarke 2003). In this view, it is crucial that the kind of causation involved in an agent's making a free choice is not reducible to causation among events involving the agent, but is rather irreducibly an instance of a substance causing a choice not by way of events. The agent, fundamentally as a substance, has the causal power to make choices without being determined to do so.

Critics of libertarianism have contended that if actions are undetermined, agents will lack the control in action required for moral responsibility. The classical presentation of this

objection is found in Hume's *Treatise of Human Nature*, and it has become known as the “luck” objection (Hume 1739/1978: 411-2; cf. Mele 2006). There are several distinct versions of this objection (Franklin 2011b). I believe that event-causal libertarianism is undermined by one in particular, but that agent-causal libertarianism might well evade it (cf. O’Connor 2000, Clarke 2003). The strongest challenge to the agent-causal position is rather one based on our best physical theories. Our choices produce physical events in the brain and in the rest of the body, and these events are, according to these theories, governed by physical laws. A libertarian view must make it credible that our choices could be free in the sense it advocates given the evidence we have about these physical laws. The concern is that agent-causal libertarianism does not meet this standard.

The version of the luck objection that in my view reveals the deepest problem for event-causal libertarianism is what I call the *disappearing agent* (DA) objection (2001, 2004, 2007):

DA objection: Consider a decision made in a context in which moral reasons favor one action, prudential reasons favor a distinct and incompatible action, and the net strength of these sets of reasons are in close competition. On an event-causal libertarian picture, the relevant causal conditions antecedent to the decision – agent-involving events -- would leave it open whether the decision will occur, and the agent has no further causal role in determining whether it does. With the causal role of the antecedent events already given, whether the decision occurs is not settled by any causal factor involving the agent. In fact, given the causal role of all causally relevant antecedent events, *nothing settles* whether the decision occurs. Thus, plausibly, on the event-causal

libertarian picture, agents lack the control required for moral responsibility.

The objection is not that agents will have no causal role in producing decisions, but that causal role that available to agents will be insufficient for the control moral responsibility demands. For on the event-causal libertarian view, the agent will “disappear” at the exact point at which moral responsibility for her decision requires her to exercise control.

To illustrate, consider Kane’s example of a businesswoman – let’s call her Anne – who has the option of deciding to stop to help an assault victim, whereupon she would be late for an important meeting at work, or not deciding to stop, which would allow her to make it to the meeting on time. For simplicity, suppose the relevant antecedent conditions are, against stopping, *Anne’s desiring at t not to annoy her boss*, and *Anne’s believing at t that if she is late for the meeting her boss will give her a difficult time*; and for stopping, *Anne’s desiring at t to help people in trouble*, and *Anne’s belief that she can be effective in helping the assault victim*. Suppose the motivational force of each of these pairs of conditions is for her about the same. On an event-causal libertarian theory, with the causal role of these antecedent conditions already given, both Anne’s deciding to stop and her not deciding to stop are significantly probable outcomes. Suppose she in fact decides to stop. There is nothing else about Anne that can settle whether the decision to stop occurs, since in this view her role in producing a decision is exhausted by antecedent states or events in which she is involved. If at this point nothing about Anne can settle whether the decision occurs, then, plausibly, she lacks the control required for moral responsibility (in the basic desert sense) for it. So it seems that on an event-causal libertarian view there is no provision that allows the agent to have control over

whether the decision occurs or not (in the crucial sorts of cases), and for this reason she lacks the control required for moral responsibility for it.

Libertarians agree that an action's resulting from a deterministic sequence of causes that traces back to factors beyond the agent's control would rule out her (basic-desert) moral responsibility for it. The deeper point of the luck objection is that if this sort of causal determination rules out moral responsibility, then it is no remedy simply to provide slack in the causal net by making the causal history of actions indeterministic. Such a move would yield one requirement for moral responsibility -- the absence of causal determinism for decision and action -- but it would not supply another -- sufficiently enhanced control (Clarke 1997, 2003). In particular, it would not provide the capacity for an agent to be the source of her decisions and actions that, according to many incompatibilists, is unavailable in a deterministic framework.

The agent-causal libertarian's solution is to specify a way in which the agent could have this enhanced control, which involves the power to settle which of the antecedently possible decisions actually occurs. The suggested remedy is to reintroduce the agent as a cause, this time not merely as involved in events, but rather fundamentally as a substance. The agent-causal libertarian claims that we possess a special causal power -- a power for an agent, fundamentally as a substance, to cause a decision without being causally determined to cause it (Chisholm 1966, O'Connor 2000, Clarke 2003, Griffiths 2010).

I argue that the agent-causal position has not been shown to be incoherent (Pereboom 2004). However, can agent-causal libertarianism be reconciled with what we would expect

given our best physical theories? Consider first the supposition that all the events in the physical world are governed by deterministic laws. In this agent-causal picture, when an agent makes a free decision, she causes the decision without being causally determined to do so. On the path to action that results from this undetermined decision, changes in the physical world, for instance in her brain or some other part of her body, are produced. But it would seem that we would at this point encounter divergences from these laws. For the changes in the physical world that result from the undetermined decision would themselves not be causally determined, and they would thus not be governed by deterministic laws. One might object that it is possible that the physical changes that result from every free decision just happen to dovetail with what could in principle be predicted on the basis of the deterministic laws, so nothing actually occurs that diverges from these laws. But this proposal would seem to involve coincidences too wild to be credible. For this reason, agent-causal libertarianism is not plausibly reconcilable with the physical world's being governed by deterministic laws.

On some interpretations of quantum mechanics, however, the physical world is not in fact deterministic, but is rather governed by probabilistic statistical laws. But wild coincidences would also arise on this suggestion. Consider the class of possible actions each of which has a physical component whose antecedent probability of occurring is approximately 0.32. It would not violate the statistical laws in the sense of being logically incompatible with them if, for a large number of instances, the physical components in this class were not actually realized close to 32% of the time. Rather, the force of the statistical law is that for a large number of instances it is correct to *expect* physical components in this class to be realized close to 32% of the time.

Are free choices on the agent-causal libertarian model compatible with what the statistical law would have us to expect about them? If they were, then for a large enough number of instances the possible actions in our class would almost certainly be freely chosen near to 32% of the time. But if the occurrence of these physical components were settled by the choices of agent-causes, then their actually being chosen close to 32% of the time would amount to a wild coincidence. The proposal that agent-caused free choices do not diverge from what the statistical laws predict for the physical components of our actions would be so sharply opposed to what we would expect as to make it incredible (for objections to this argument, see O'Connor 2003, 2008; Clarke 2003: 181, n. 31, 2011; for replies see Pereboom 2005).

At this point, the libertarian agent-causalist might propose that there are indeed divergences from the probabilities that we would expect absent agent-causes, and that these divergences are located at the interface between the agent-cause and that part of the physical world that it directly affects -- an interface very likely to be found in the brain. But the issue for this proposal is that we have no evidence that such divergences occur. This difficulty yields a strong reason to reject this approach.

Thus the various kinds of libertarianism face significant problems. Because compatibilism is vulnerable to the argument from manipulation cases, the position that remains is hard incompatibilism, which denies that we have the sort of free will required for moral responsibility. The concern for this stance is not that there is considerable empirical evidence that it is false, or that there is a powerful argument that it is somehow incoherent, and false for

that reason. Rather, the crucial questions it faces are practical: What would life be like if we believed it was true? Is this a sort of life that we can cope with?⁵

Hard incompatibilism and wrongdoing

Accepting hard incompatibilism requires giving up our ordinary view of ourselves as blameworthy for immoral actions and praiseworthy for those that are morally exemplary. At this point one might object that giving up our belief in moral responsibility would have harmful consequences, perhaps so harmful that thinking and acting as if this skeptical view is true is not a feasible option. So even if the claim that we are morally responsible turns out to be false, there might yet be weighty practical reasons to believe that we are, or at least to treat people as if they were morally responsible.

For instance, one might think that if we gave up the belief that people are blameworthy and praiseworthy, we could no longer legitimately judge any actions as wrong or even bad, or as right or good. But this thought seems mistaken. Even if we came to believe that some perpetrator of genocide was not blameworthy due to a degenerative brain disease, we would still hold that his actions were morally wrong, or at least that it was very bad that he acted as he did. So, in general, denying blameworthiness would not appear to threaten judgments of

⁵ In responding to these questions, I have been inspired by others who have done excellent work in answering them, including Spinoza (1677/1985), Galen Strawson (1986), Ted Honderich (1988); Bruce Waller (1990), and Saul Smilansky (2000).

wrongness or badness, and, likewise, denying praiseworthiness would not seem to undercut assessments of rightness or goodness (for a contrary view, see Haji (1998, 2002)).

Perhaps treating wrongdoers as blameworthy is often required for effective moral education and improvement. If we resolved never to treat people as blameworthy, we might be left with insufficient leverage to reform immoral behavior (Nichols 2007; for a response see Pereboom 2009b). But this proposal would have us treat people as blameworthy -- by, for example, expressing anger toward them just because of what they have done -- when they do not deserve it, which would seem morally wrong. If people are not morally responsible for immoral behavior, treating them as if they were seems unfair. However, it is possible to achieve moral reform by methods not threatened by this sort of unfairness, and in ordinary situations such practices could arguably be as successful as those that presuppose moral responsibility. Instead of treating people as if they deserved blame, the free will skeptic can turn to moral admonition and encouragement, which presuppose only that the offender has done wrong. These methods can effectively communicate a sense of right and wrong, and they might well issue in salutary reform.

But does hard incompatibilism have resources adequate for contending with criminal behavior? Here it would appear to be at a disadvantage, and if so, practical considerations might generate good reasons to treat criminals as if they were morally responsible. First, if this skeptical view is true, a retributivist justification for criminal punishment is unavailable, for it asserts that a criminal deserves pain or deprivation just for committing the crime, while hard

incompatibilism rejects this basic-desert claim. And retributivism is among the most naturally compelling ways to justify criminal punishment.

By contrast, a theory that justifies criminal punishment on the ground that punishment educates criminals morally is not threatened by hard incompatibilism specifically. Thus one might suggest that the free will skeptic should endorse a view of this kind. However, we lack strong empirical evidence that punishing criminals results in moral education, and without such evidence, it would be wrong to punish them to achieve this aim. It is generally wrong to harm a person for the sake of realizing some good without strong evidence that the harm will produce the good. In addition, even if we had impressive evidence that punishment is effective in morally educating criminals, we should prefer non-punitive ways of achieving this aim, whether or not criminals are morally responsible.

According to deterrence theories, punishing criminals is justified for the reason that it deters future crime. The two most-discussed deterrence theories, the utilitarian view and the version that grounds the right to punish on the right to self-defense and defense of others, are not imperiled by hard incompatibilism *per se*. But they are questionable on other grounds. The utilitarian theory, which claims that punishment is justified when and because it maximizes utility, faces well-known objections. It would require punishing the innocent when doing so would maximize utility; in certain situations it would prescribe punishment that is unduly severe; and it would authorize harming people merely as means to the safety of others. The kind of deterrence theory that grounds the right to punish in the right of individuals to defend themselves and others against immediate threats (Farrell 1985: 38-60) is also objectionable. For

when a criminal is sentenced to punishment he most often does not pose an immediate threat to anyone, since he is then in the custody of the law, and this fact about his circumstances distinguishes him from those who can legitimately be harmed on the basis the right of self-defense and defense of others.

There is, however, an intuitively legitimate theory of crime prevention that is neither undercut by the skeptical view, nor threatened by other sorts of considerations. This theory draws an analogy between the treatment of criminals and the treatment of carriers of dangerous diseases. Ferdinand Schoeman (1979) argues that if we have the right to quarantine carriers of serious communicable diseases to protect people, then for the same reason we also have the right to isolate the criminally dangerous. Notice that quarantining a person can be justified when she is not morally responsible for being dangerous to others. If a child is infected with a deadly contagious virus that was transmitted to her before she was born, quarantine can still be legitimate. Now imagine that a serial killer poses a grave danger to a community. Even if he is not morally responsible for his crimes (say because no one is ever morally responsible), it would be as legitimate to isolate him as it is to quarantine a non-responsible carrier of a serious communicable disease.

It would be morally wrong to treat carriers of communicable diseases more severely than is required to protect people from the resulting threat. Similarly, it would be wrong to treat criminals more harshly than is required to protect society against the danger posed by them. Moreover, just as moderately dangerous diseases may allow for only measures less intrusive than quarantine, so moderately serious criminal tendencies might only justify

responses less intrusive than detention. Furthermore, I suspect that a theory modeled on quarantine would not justify measures of the sort whose legitimacy is most in doubt, such as the death penalty or confinement in the worst prisons we have. It would also demand a degree of concern for the rehabilitation and well-being of the criminal that would alter much of current practice. Just as society must seek to cure the diseased it quarantines, so it would be required to try to rehabilitate the criminals it detains. In addition, if a criminal cannot be rehabilitated, and if protection of society demands his indefinite confinement, there would be no justification for making his life more miserable than is needed to guard against the danger he poses.

Meaning in life

If the skeptical view about free will is true and we came to recognize this, could we legitimately retain a sense of achievement for successes that make our lives fulfilled, happy, or worthwhile (Honderich 1988)? It might be argued that on the supposition of this position there would be no genuine achievements, for an agent cannot have an achievement for which she is not also praiseworthy. However, achievement is not as closely tied to praiseworthiness as this objection would have it. If an agent hopes to achieve success in a project she undertakes, and if she accomplishes what she hoped for, intuitively this outcome would be an achievement of hers even if she is not praiseworthy for it -- although the sense in which it is her achievement might be diminished. For instance, if teacher hopes that her efforts will result in well-educated children, and they do, there remains a clear sense in which she has achieved what she hoped for -- even if it turns out she is not praiseworthy for what she has accomplished.

One might be concerned that accepting hard incompatibilism would instill an attitude of resignation to whatever the future holds in store and would undcut our motivation for achievement. But this is not clearly correct. Even if what we understand about our behavioral dispositions and our environment provides evidence our futures turning out in a particular way, it can often be reasonable to hope that they will turn out differently. For this to be so, it may sometimes be important for us to lack complete knowledge of our dispositions and environmental conditions. Suppose that someone reasonably believes that he has a disposition that would an impediment to realizing something he hopes to achieve. But because he does not know whether this disposition will in fact have this effect, it remains open for him – that is, not ruled out by anything he knows or believes – that another disposition he has will allow him to transcend the impediment. For example, imagine that someone aspires to become a successful politician, but he is concerned that his fear of public speaking will keep this from happening. He does not know for sure whether this fear will in fact frustrate his ambition, since it is open for him that he will overcome this problem, perhaps due to a disposition for resolute self-discipline in transcending obstacles of this sort. Thus he might reasonably hope that he will overcome his fear and succeed in his ambition.

At the same time, one might concur with Saul Smilansky that although determinism allows for a limited foundation of the sense of self-worth that derives from achievement and virtue, the free will skeptic's perspective can nevertheless be "extremely damaging to our view of ourselves, to our sense of achievement, worth, and self-respect," especially when it comes to achievement in the formation of one's own moral character. In response to this concern,

Smilansky argues that it would be best for us to foster the illusion that we have free will (Smilansky 1997, 2000). I agree with Smilansky that there is a kind of self-respect that presupposes that we have the sort of free will required for moral responsibility, and that this self-respect would be undermined if hard incompatibilism were true. I question, however, whether Smilansky is right about how damaging it would be for us to relinquish this sort of self-respect, and whether an appeal to illusion is required as a remedy.

Note first that our sense of self-worth -- our sense that we have value and that our lives are worth living -- is to a non-trivial extent due to features we possess not produced by our will, let alone by free will. People place great value on natural beauty, native athletic ability, and intelligence, none of which result from our voluntary efforts. We also value efforts that are voluntary in the sense that they are willed by us -- in productive work and altruistic behavior, and in the formation of moral character. But how much does it matter to us that these voluntary efforts are also *freely* willed? Perhaps Smilansky overestimates how much we care.

Consider how someone comes to have a good moral character. Not implausibly, this character was formed to some significant degree by upbringing, and the belief that this is so is widespread. Parents typically regard themselves as having failed in raising their children if they turn out with immoral dispositions, and parents often take great care to bring their children up to prevent such a result. Accordingly, people often come to believe that they have the good moral character they do largely because they were raised with love and skill. But those who come to believe this about themselves seldom experience dismay because of it. People tend not to become dispirited upon coming to believe that their good moral character is not their

own doing, and that they do not deserve significant praise or respect for it. By contrast, they often come to feel more fortunate and thankful. Suppose, however, that there are some who would be overcome with dismay. Would it be justified or even desirable for them to foster the illusion that they nonetheless deserve praise and respect for producing their moral character? I suspect that most people would eventually be able to accept the truth without incurring significant loss. All of this, I believe, would also hold for those who come to believe that they do not deserve praise and respect for producing their moral character because they are not, in general, morally responsible in the basic-desert sense.

Emotions, reactive attitudes, and personal relationships

P. F. Strawson (1962) contends that the justification for judgments of blameworthiness and praiseworthiness has its foundation in the reactive attitudes -- emotional reactions to how people voluntarily behave -- attitudes such as moral resentment, guilt, gratitude, forgiveness, and love. Because moral responsibility has this type of foundation, the truth or falsity of determinism is irrelevant to whether we are justified in regarding agents as morally responsible. This is because these reactive attitudes are required for the kinds of interpersonal relationships that make our lives meaningful, and so even if we were able to give up the reactive attitudes, we would never have sufficient practical reason to do so. Strawson believes that it is in fact psychologically impossible for us to relinquish our reactive attitudes altogether, but in a limited range of cases we can adopt what he calls the "objective attitude," which he conceives as a cold and calculating stance:

To adopt the objective attitude to another human being is to see him, perhaps, as an object of social policy; as a subject for what, in a wide range of sense, might be called treatment; as something certainly to be taken account, perhaps precautionary account, of; to be managed or handled or cured or trained; perhaps simply to be avoided... The objective attitude may be emotionally toned in many ways: it may include repulsion or fear, it may include pity or love, though not all kinds of love. But it cannot include the range of reactive feelings and attitudes which belong to involvement or participation with others in interpersonal human relationships; it cannot include resentment, gratitude, forgiveness, anger, or the sort of love which two adults can sometimes be said to feel reciprocally, for each other. (Strawson 1962)

Strawson suggests that if determinism did threaten our reactive attitudes, and we were able to give them up, we would face the prospect of adopting this objective attitude toward everyone, and as a result our interpersonal relationships would be damaged. But because we have extremely good practical reasons for maintaining these relationships, we would never have sufficient practical reason to adopt the objective attitude in most cases. As a result we would not have sufficient reason to relinquish our reactive attitudes, and thus to stop regarding people as morally responsible.

I think it's plausible that if we persistently maintained an objective attitude toward others, as Strawson describes it, our interpersonal relationships would be seriously threatened.⁶ However, I deny that we would have good reason to adopt this stance if we came

⁶ For a contrary view, see Sommers 2007; Shabo (2011) defends Strawson against Sommers.

to accept the skeptical view about free will. Some of the reactive attitudes would then in fact be challenged, because some of them, such as moral resentment and indignation, would have the false presupposition that the person who is the object of the attitude is morally responsible in the basic desert sense. But the reactive attitudes we would want to retain either would not be threatened in this way, or else have analogues or aspects that would not have false presuppositions. The attitudes that would survive do not amount to the objective attitude, and they would be sufficient to sustain good interpersonal relationships.

It is plausible that to a certain degree moral resentment and indignation are beyond our power to affect. Even supposing that a free will skeptic is thoroughly committed to morality and rationality, and that she is admirably in control of her emotions, she might well be unable to eliminate these attitudes. So we might continue to expect people to be morally resentful or indignant in certain circumstances. But at the same time we have the ability to prevent, temper, and sometimes to dispel these attitudes, and given the skeptical conviction, we might do so for the sake of morality and rationality. Modifications of this sort might well be good for interpersonal relationships.

One might reply that in relationships moral resentment and indignation are crucial to effective communication of wrongdoing, and if we dispelled or modified these attitudes, relationships would be damaged. However, when someone is wronged in a relationship, she typically has further attitudes not threatened by the skeptical view whose expression can have the communicative role at issue. These attitudes include being alarmed or distressed about

what another has done, and moral concern, sadness, or sorrow for him (Pereboom 2009b; for a contrary view, see Nichols 2007). Moral resentment is thus not clearly required for effective communication in interpersonal relationships.

Forgiveness might seem to presuppose that the person being forgiven is blameworthy, and if this is so, this attitude would also be threatened by hard incompatibilism.⁷ But certain key features of forgiveness would not be endangered, and they are sufficient to sustain the role forgiveness has in relationships. Suppose a friend repeatedly mistreats you, and in consequence you decide to end your relationship with him. However, he then apologizes to you, indicating his recognition that his actions were wrong, a wish that he had not mistreated you, and a commitment to refrain from the immoral behavior. Because of this you decide not to end the friendship. In this case, the aspect of forgiveness that is consistent with the skeptical position is a willingness to cease to regard past immoral behavior as a reason to weaken or end a relationship. A feature often associated with forgiveness that would be undercut is the disposition to disregard the friend's blameworthiness. But since the skeptic denies blameworthiness (in the sense that involves basic desert), she no longer needs a willingness to disregard blameworthiness for good interpersonal relationships.

One might object that accepting hard incompatibilism would jeopardize the self-directed attitudes of guilt and repentance, and that this would be especially bad for relationships. Without guilt and repentance, we would not only be incapable of restoring

⁷ See Nelkin (2008, 2011) for reasons to think that forgiveness does not have this presupposition.

relationships damaged due to wrongdoing, but, in addition, it would become more difficult to restore the moral integrity of those who have done wrong. For without the attitudes of guilt and repentance, we would no longer have the psychological mechanisms that can play these roles. Note first, however, that it would be because guilt essentially involved a belief that one is blameworthy that this attitude would be challenged by the skeptical view. It is for this reason that repentance would also seem to be (indirectly) threatened, for a sense of guilt would seem required to motivate repentance. Imagine, however, that in your relationship with another person you have acted immorally, but because you endorse skepticism about free will, you do not believe that you are blameworthy in the basic-desert sense. Instead, you acknowledge that you were the agent of wrongdoing, you feel genuine sorrow because of what you have done to another person, and you deeply regret having acted as you did. Moreover, because you are committed to doing what is right and to your own moral improvement, you resolve not to act in this way in the future, and you communicate this to the other person. None of these measures threatened are threatened by hard incompatibilism (see Pereboom 2009b for further discussion).

Gratitude would seem to presuppose that the agent to whom one is grateful is morally responsible for a beneficial act, whereupon the skeptical view would jeopardize gratitude. But as in the case of forgiveness, certain core aspects of this attitude would remain unaffected, and these aspects can provide what is required for good interpersonal relationships. Gratitude involves, first of all, being thankful toward someone who has acted beneficially. True, being thankful toward someone often involves the belief that she is praiseworthy for some action in

the basic desert sense. Still, one can be thankful to a young child for some kindness without believing that she is morally responsible for it. This aspect of gratitude could be retained even without the presupposition of praiseworthiness. Usually gratitude also involves joy in response to a person for what she has done. But no feature of hard incompatibilism poses a threat to the legitimacy of this sort of joy and of expressing it. Expressing joy can bring about the same sense of harmony and goodwill typically produced by a sense of gratitude unmodified by free will skepticism, and thus on this point the skeptical view is not at a disadvantage.

Is the kind of love that mature adults have for each other in good relationships endangered by a hard incompatibilist conviction, as Strawson's line of argument suggests? Consider first whether for loving someone it is significant that the person who is loved possesses and exercises free will in the sense required for moral responsibility. Parents love their children rarely, if ever, for the reason that they possess this sort of free will, or decide to do what is right by free will, or deserve to be loved due to freely-willed action. When adults love each other, it is also very seldom, if at all, for these sorts of reasons. In addition to moral character and action, factors such as intelligence, appearance, style, and resemblance to certain others in one's personal history all might play a part. Plausibly, morally admirable qualities are particularly important in occasioning, enriching, and maintaining love. But even if there is an aspect of love that we see as a deserved response to morally admirable qualities, it is unlikely that love would be diminished at all if we came to believe that these qualities are not brought about or sustained by freely willed decisions. Such admirable qualities are loveable whether or not we think of agents as deserving praise for having them.

One might contend that we want to be freely loved by others – to be loved by them by their own free will. However, the love that parents have for their children is typically independent of the parents' will altogether, and we don't regard love of this sort as deficient. Kane acknowledges this fact about parents' love, and he recognizes that romantic love is similar in this respect. But he argues that there is a kind of love we very much want that would be ruled out if love were always causally determined by factors beyond our control (Kane 1996: 88). The plausibility of Kane's idea might be enhanced by reflecting on how you would react if you found out that someone you love was causally determined to love you by a benevolent manipulator.

Setting aside *free* will for a moment, we might ask: when does the will play any role at all in engendering love? When a relationship is disintegrating, people will sometimes decide to try to restore the love they once had for one another. When a student finds himself in conflict with a roommate from the outset, he might choose to do what he can to improve the relationship. When a marriage is arranged, the partners may decide to take steps to promote love for each other. In these kinds of circumstances, we want others to make decisions that would produce or maintain love. However, this is not to say that we would want such choices to be freely willed in the sense required for moral responsibility. It is not clear that appreciable value would be added by such a decision's being free in this sense. Moreover, although in some circumstances we might want others to make such relationship-enhancing decisions, we would typically prefer love that did not require choices of this sort. This is so not only for intimate

romantic relationships — where it is undeniable -- but also for friendships and for relationships between parents and children.

But imagine that Kane's view was vindicated, and we did desire a kind of love that is freely willed in the sense required for basic-desert moral responsibility. We would then desire a kind of love that would be impossible if hard incompatibilism were true. Still, the sorts of love not undercut by this skeptical position would be sufficient for good interpersonal relationships. If we can aspire to the kind of love parents typically have for their children, or the type romantic lovers share, or the sort had by friends who are devoted to each other, and whose friendship was deepened through their interactions, then the prospect of fulfillment through interpersonal relationships remains intact (see Pereboom 2009b for further discussion).

Thus a skeptical conviction does not pose a threat to good interpersonal relationships. It might well undermine certain attitudes that typically have a role in such relationships. Moral resentment and indignation would likely be irrational for the free-will skeptic, since these attitudes would have presuppositions she believes to be false. But these attitudes are either not required for good relationships, or they have analogues that could play their typical beneficial role. Moreover, love -- the attitude most essential to good interpersonal relationships -- does not appear to be placed at risk by such a view at all. Love of another involves, fundamentally, wishing for the other's good, taking on her aims, and a desire to be together with her, and none of this is undercut by hard incompatibilism.

The good in hard incompatibilism

The skeptical view also holds out the prospect of substantial benefits for human life. Of all the attitudes associated with the belief that we are morally responsible, anger seems most closely connected with it. Discussions about moral responsibility typically focus not on how we regard morally exemplary agents, but rather on our attitudes toward wrongdoers. Examples designed to elicit a strong intuition that an agent is morally responsible usually feature action that is especially malevolent, and the intuition typically involves sympathetic anger. It might be, then, that our attachment to the belief that we are morally responsible derives to a significant degree from the role anger has in our emotional lives. Perhaps we sense that giving up the assumption of responsibility is threatening because the rationality of anger would be unjustified as a result.

The type of anger at issue is the sort directed toward a person who is believed to have behaved immorally, and it comprises both moral resentment and indignation. Let us call this attitude *moral anger*. Not all anger is moral anger. One kind of non-moral anger is directed toward someone because his abilities are lacking in some way or he has performed poorly in some situation. At times we are angry with machines for malfunctioning. Sometimes our anger has no object. But most human anger is moral anger.

Such anger has a significant part in our moral psychology. It motivates us to resist abuse, discrimination, and oppression. But expression of moral anger often has harmful effects for those toward whom it is directed, and also for those expressing the anger. Frequently its expression is intended to cause little else than emotional or physical pain. As a result, moral

anger has a tendency to damage relationships, hinder the functioning of organizations, and unsettle societies. In extreme cases, it can motivate people to torture and kill.

The realization that expression of moral anger can be damaging generates a strong requirement that it be morally justified when it occurs. The demand to morally justify behavior that is harmful is generally very strong, and expression of moral anger is often harmful. This demand is made more urgent by our degree of attachment to this emotion, which is fueled by the satisfaction we frequently have in expressing it. Often we justify expression of such anger by contending that wrongdoers are morally responsible in the basic-desert involving sense for what they have done. If we became convinced that we do not have the sort of free will required for moral responsibility, we would regard such justifications as illegitimate. About the view that we lack this sort of free will Spinoza says: “this doctrine contributes to the social life insofar as it teaches us to hate no one, to disesteem no one, to mock no one, to be angry at no one...” (1677/1985: 490). Given the concerns to which expression of moral anger gives rise, our coming to believe that we lack this sort of free will may on balance be a good thing.

References

- Baker, Lynne R. (2006). “Moral Responsibility without Libertarianism,” *Noûs* 40, pp. 307-30.
- Balaguer, Mark (2009). *Free Will as an Open Scientific Problem*, Cambridge, MA: MIT Press.
- Bok, Hilary (1998). *Freedom and Responsibility*, Princeton NJ: Princeton University Press.
- Chisholm, Roderick (1976). *Person and Object*, La Salle IL: Open Court.

- Clarke, Randolph (1996). "Agent Causation and Event Causation in the Production of Free Action," *Philosophical Topics* 24, pp. 19-48.
- Clarke, Randolph (2003). *Libertarian Theories of Free Will*, New York: Oxford University Press.
- Clarke, Randolph (2011). "Are We Free to Obey the Laws?" *American Philosophical Quarterly*, forthcoming.
- Demetriou, Kristin (2010). "The Soft-Line Solution to Pereboom's Four-Case Argument," *Australasian Journal of Philosophy* 88, pp. 595-617.
- Ekstrom, Laura W. (2000). *Free Will: A Philosophical Study*, Boulder, CO: Westview.
- Farrell, Daniel M. (1985). "The Justification of General Deterrence," *Philosophical Review* 104, pp. 367-94.
- Fischer, John Martin (1994). *The Metaphysics of Free Will*, Oxford: Blackwell.
- Fischer, John Martin, and Mark Ravizza (1998). *Responsibility and Control: A Theory of Moral Responsibility*, New York: Cambridge University Press.
- Fischer, John Martin (2004). "Responsibility and Manipulation," *The Journal of Ethics* 8; pp. 145-77; reprinted in Fischer (2006).
- Fischer, John Martin (2006). *My Way*, New York: Oxford University Press.
- Fischer, John Martin (2010). "The Frankfurt Cases: The Moral of the Stories," *Philosophical Review* 119, pp. 315-336.
- Fischer, John Martin, Robert Kane, Derk Pereboom, and Manuel Vargas (2007). *Four Views on Free Will*, Oxford: Blackwell.

- Frankfurt, Harry G. (1969). "Alternate Possibilities and Moral Responsibility," *Journal of Philosophy* 66, pp. 829-39.
- Frankfurt, Harry G. (1971). "Freedom of the Will and the Concept of a Person," *Journal of Philosophy* 68, pp. 5-20.
- Franklin, Christopher (2011a). "Neo-Frankfurtians and Buffer Cases: The New Challenge to the Principle of Alternative Possibilities," *Philosophical Studies* 152, pp. 189–207
- Franklin, Christopher (2011b). "Farewell to the Luck (and Mind) Argument," *Philosophical Studies*, forthcoming.
- Ginet, Carl (1990). *On Action*, Cambridge: Cambridge University Press.
- Ginet, Carl (1996). "In Defense of the Principle of Alternative Possibilities: Why I Don't Find Frankfurt's Arguments Convincing," *Philosophical Perspectives* 10, pp. 403-17.
- Ginet, Carl (2002). "Review of *Living without Free Will*," *Journal of Ethics* 6, pp. 305-9.
- Goetz, Stewart (2008). *Freedom, Teleology, and Evil*, London: Continuum
- Griffiths, Meghan (2010). "Why Agent-Caused Actions are not Lucky," *American Philosophical Quarterly* 47, pp. 43-56.
- Haji, Ishtiyaque (1998). *Moral Accountability*, New York: Oxford University Press.
- Haji, Ishtiyaque (2002). *Deontic Morality and Control*, Cambridge: Cambridge University Press.
- Honderich, Ted (1988). *A Theory of Determinism*, Oxford: Oxford University Press.
- Hume, David. (1739/1978) *A Treatise of Human Nature*, Oxford: Oxford University Press.
- Hunt, David (2000). "Moral Responsibility and Unavoidable Action," *Philosophical Studies* 97, pp. 195-227

- Hunt, David (2005). Moral Responsibility and Buffered Alternatives, *Midwest Studies in Philosophy* 29, pp. 126-45.
- Kane, Robert (1985). *Free Will and Values*, Albany: SUNY Press.
- Kane, Robert (1996). *The Significance of Free Will*, New York: Oxford University Press.
- Lycan, William G. (1997). *Consciousness*, Cambridge MA: MIT Press.
- McCann, Hugh (1998). *The Works of Agency*, Ithaca: Cornell University Press.
- McKenna, Michael (2003). "Robustness, Control, and the Demand for Morally Significant Alternatives," in *Freedom, Responsibility, and Agency: Essays on the Importance of Alternative Possibilities*, Michael McKenna and David Widerker, eds., Aldershot UK: Ashgate, pp. 201-16.
- McKenna, Michael (2008). "A Hard-line Reply to Pereboom's Four-Case Argument," *Philosophy and Phenomenological Research* 77, pp. 142-59.
- Mele, Alfred (1995). *Autonomous Agents*, New York: Oxford University Press.
- Mele, Alfred (2005). "A Critique of Pereboom's 'Four-case' Argument for Incompatibilism," *Analysis* 65, pp. 75-80.
- Mele, Alfred (2006). *Free Will and Luck*, New York: Oxford University Press.
- Moya, Carlos (forthcoming). "On the Very Idea of a Robust Alternative."
- Nelkin, Dana (2008). "Responsibility and Rational Abilities: Defending an Asymmetrical View," *Pacific Philosophical Quarterly* 89, pp. 497-515.
- Nelkin, Dana (2011). *Making Sense of Freedom and Responsibility*, Oxford: Oxford University Press.

- Nichols, Shaun (2007). "After Compatibilism: a Naturalistic Defense of the Reactive Attitudes," *Philosophical Perspectives* 21, pp. 405-28.
- O'Connor, Timothy (2000). *Persons and Causes*, New York: Oxford University Press.
- O'Connor, Timothy (2003). "Review of *Living without Free Will*," *Philosophical Quarterly* 53, pp. 308-10.
- O'Connor, Timothy (2008). "Agent-Causal Power," in Toby Handfield, ed., *Dispositions and Causes*, Oxford: Clarendon Press, pp. 189-214.
- Otsuka, Michael (1998). "Incompatibilism and the Avoidability of Blame," *Ethics* 108, pp. 685-701.
- Palmer, David (2011). "Pereboom on the Frankfurt Cases," *Philosophical Studies*, forthcoming.
- Pereboom, Derk (1995). "Determinism *Al Dente*," *Noûs* 29, pp. 21-45.
- Pereboom, Derk (2000). "Alternative Possibilities and Causal Histories," *Philosophical Perspectives* 14, 2000, pp. 119-37.
- Pereboom, Derk (2001). *Living without Free Will*, Cambridge: Cambridge University Press.
- Pereboom, Derk (2003). "Source Incompatibilism and Alternative Possibilities," in *Freedom, Responsibility, and Agency: Essays on the Importance of Alternative Possibilities*, Michael McKenna and David Widerker, eds., Aldershot UK: Ashgate, pp. 185-99.
- Pereboom, Derk (2004). "Is Our Conception of Agent Causation Incoherent?" *Philosophical Topics*. 32, pp. 275-86.
- Pereboom, Derk (2005). "Defending Hard Incompatibilism," *Midwest Studies in Philosophy* 29,

pp. 228-47.

Pereboom, Derk (2007). "Hard Incompatibilism," and "Response to Kane, Fischer, and Vargas," in Robert Kane, John Martin Fischer, Derk Pereboom, and Manuel Vargas, *Four Views on Free Will*, Oxford UK: Blackwell, pp. 85-125, 191-203.

Pereboom, Derk (2008a). "A Hard-Line Reply to the Multiple-Case Manipulation Argument," *Philosophy and Phenomenological Research* 77, 2008, pp. 160-70.

Pereboom, Derk (2008b). "Defending Hard Incompatibilism Again," *Essays on Free Will and Moral Responsibility*, ed. Nick Trakakis and Daniel Cohen, Newcastle UK: Cambridge Scholars Press, pp. 1-33.

Pereboom, Derk (2009a). "Further Thoughts about a Frankfurt-Style Argument," *Philosophical Explorations* 12 (2010), pp. 109-18

Pereboom, Derk (2009b). "Free Will, Love and Anger," *Ideas y Valores* 141, pp. 5-25.

Pereboom, Derk (2012). "Frankfurt Examples, Derivative Responsibility, and the Timing Objection," *Philosophical Issues* nn, pp.

Russell, Paul (2000). "Compatibilist-Fatalism", in *Moral Responsibility and Ontology*, ed. Ton van den Beld, Dordrecht: Kluwer, pp. 199-218.

Scanlon, Thomas (1998). *What We Owe to Each Other*, Cambridge: Harvard University Press.

Schoeman, Ferdinand D. (1979). "On Incapacitating the Dangerous," *American Philosophical Quarterly* 16, pp. 27-35.

- Shabo, Seth (2010). "Uncompromising Source Incompatibilism," *Philosophy and Phenomenological Research* 80, pp. 349-83.
- Shabo, Seth (2011). "Incompatibilism and Personal Relationships: Another Look at Strawson's Objective Attitude," *Australasian Journal of Philosophy*, forthcoming.
- Smilansky, Saul (1997). "Can a Determinist Help Herself?," C.H. Manekin and M. Kellner, ed., *Freedom and Moral Responsibility: General and Jewish Perspectives*, College Park MD: University of Maryland Press, pp. 85-98.
- Smilansky, Saul (2000). *Free Will and Illusion*, Oxford: Oxford University Press.
- Sommers, Tamler (2007). "The Objective Attitude," *Philosophical Quarterly* 57, pp. 321-41.
- Spinoza, Baruch (1677/1985). *Ethics*, in *The Collected Works of Spinoza*, Edwin Curley, ed. and tr., Volume 1, Princeton NJ: Princeton University Press.
- Strawson, Galen (1986). *Freedom and Belief*, Oxford: Oxford University Press.
- Strawson, Peter F. (1962). "Freedom and Resentment," *Proceedings of the British Academy* 48 (1962), pp. 1-25.
- Taylor, Richard (1966). *Action and Purpose*, Englewood Cliffs: Prentice-Hall.
- Taylor, Richard (1974). *Metaphysics*, 4th ed., Englewood Cliffs: Prentice-Hall.
- Todd, Patrick (2011). "A New Approach to Manipulation Arguments," *Philosophical Studies* 152, pp. 127-33.
- van Inwagen, Peter (1983). *An Essay on Free Will*, Oxford: Oxford University Press.
- Wallace, R. Jay (1994). *Responsibility and the Moral Sentiments*, Cambridge MA, Harvard University Press.

Waller, Bruce (1990). *Freedom without Responsibility*, Philadelphia: Temple University Press.

Widerker, David (1995). "Libertarianism and Frankfurt's Attack on the Principle of Alternative Possibilities," *Philosophical Review* 104, pp. 247-61.

Widerker, David (2006). "Libertarianism and the Philosophical Significance of Frankfurt Scenarios," *Journal of Philosophy* 103, pp. 163-87.