

Why We Have No Free Will and Can Live Without It

DERK PEREBOOM

Derk Pereboom teaches philosophy at Cornell University.

1. OUTLINE OF HARD INCOMPATIBILISM

Baruch Spinoza (1677/1985: 440–44, 483–84, 496–97) maintained that due to certain general facts about the nature of the universe, we human beings do not have the sort of free will required for moral responsibility. I agree. More exactly, he argues that it is because causal determinism is true that we lack this sort of free will; he is thus a *hard determinist*. By contrast, the position I defend is agnostic about causal determinism. I contend, like Spinoza, that we would not have the sort of free will required for moral responsibility if causal determinism were true, but also that indeterministic theories do not significantly improve the prospects for this sort of free will. Consequently, we need to take seriously the verdict that we lack the sort of free will required for moral responsibility. I call the resulting view *hard incompatibilism*. In addition, I argue that a conception of life without this sort of free will need not exclude morality or our sense of meaning in life, and in some respects it could even be beneficial.

2. AGAINST COMPATIBILISM

The case for hard incompatibilism involves arguing against two competing positions. The first of these is *compatibilism*, which claims that free will of the type required for moral responsibility is compatible with determinism. Compatibilists typically maintain, in addition, that we do in fact have this sort of free will. The second is *libertarianism*, which contends that although the sort of free will required for moral responsibility is not compatible

with determinism, it turns out that determinism is false, and we do have this kind of free will.

Compatibilists typically attempt to formulate conditions on agency intended to provide an account of what it is to be morally responsible for an action. These conditions are compatibilist in that they allow for an agent to be morally responsible for an action even when she is causally determined to act as she does. For instance, David Hume and his followers specify that morally responsible action be caused by desires that flow from the agent's "durable and constant" character, and that the agent not be constrained to act, at least in the sense that the action not result from an irresistible desire (Hume 1739/1978: 319–412). Harry Frankfurt proposes that moral responsibility requires that the agent have endorsed and produced her will to perform the action in the right way. More specifically, she must have a second-order desire—that is, a desire to have a particular desire—to will to perform it, and her will must be her will because she has this second-order desire (Frankfurt 1971). John Fischer argues that morally responsible action must result from a rational consideration of the reasons at issue; among other things, the agent must be receptive to the reasons present in a situation, and she must be responsive to them to the degree that in at least some situations in which the reasons are different, she would have done otherwise (Fischer 1994). Finally, Jay Wallace proposes that moral responsibility requires that the agent have the general ability to grasp, apply, and regulate her behavior by moral reasons (Wallace 1994). Each of these compatibilists intends for his conditions to be sufficient for an agent's

moral responsibility when they are supplemented by some fairly uncontroversial additional necessary conditions, such as the provision that the agent understands that killing is morally wrong.

In my view, the best type of challenge to 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. This intuition remains strong even if she meets the compatibilist conditions on moral responsibility just canvassed. The following "four-case argument" first of all develops examples of actions that involve such manipulation, in which these compatibilist conditions on moral responsibility are satisfied (1995, 2001). These cases, taken individually, 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 additional force, by way of setting out three such cases, each progressively more like a fourth scenario which the compatibilist might envision to be realistic, in which the action is causally determined in a natural way. The further challenge for the compatibilist is to point out a difference between this fourth scenario and one or more of the manipulation examples that shows why the agent might be morally responsible in the ordinary case but not in the manipulation examples. My contention is that non-responsibility generalizes from at least one of the manipulation cases to the ordinary one.

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. We design the cases so that his act of murder conforms to the prominent compatibilist conditions. Plum's action meets the Humean conditions, since for him purely selfish reasons typically weigh heavily—much too heavily as judged 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. It fits the condition proposed by Frankfurt: 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 Fischer: Plum's desires are 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 this reason. In addition, this action meets the condition advanced by Wallace: Plum retains the general ability to grasp, apply, and regulate his behavior by moral reasons. For instance, when egoistic reasons that count against acting morally are relatively weak, he will usually regulate his behavior by moral reasons instead. This ability even provides him with the capacity to revise and develop his moral character over time. Now, supposing that causal determinism is true, is it plausible that 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.

Case 1: Professor Plum was created by neuroscientists, and they can manipulate him directly through the use of radio-like technology, but he is as much like an ordinary human being as is possible given these unusual features. The neuroscientists manipulate him to undertake the process of reasoning by which the desires at play in his act of murder are brought about and modified. They do this by pushing a series of buttons just before he begins to reason about his situation, thereby causing his reasoning process to be rationally egoistic. Plum does not think and act contrary to character since his reasoning process is typically rationally 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 egoistic reasons were different. Still, he is not exclusively rationally egoistic, since he typically

regulates his behavior by moral reasons when the egoistic reasons are relatively weak— weaker than they are in the current situation. He is also not constrained, in the sense that he does not act because of an irresistible desire—the neuroscientists do not provide him with 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 taken together.

A compatibilist might resist this conclusion by arguing that although in Case 1 the process resulting in the action satisfies all of the prominent compatibilist conditions, yet Plum's relevant states are directly produced by the manipulators moment by moment—he is locally manipulated—and this is the aspect of the story that undermines his moral responsibility. In reply, could a time lag between the manipulators' activity and the production of the states in the agent plausibly make the crucial difference as to whether an agent is morally responsible? If the neuroscientists did all of their manipulating during one time interval and, after an appropriate length of time, the relevant states were produced in him, would he only then be morally responsible? It is my sense that such a time lag, all by itself, would make no difference to whether an agent is responsible.

Let us now 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* to weigh reasons for action so that he is often but not exclusively rationally egoistic, with the consequence that in the circumstances in which he now finds himself, he is causally determined to engage in the reasons-responsive process of deliberation and to have the set of first- and second-order desires that result in his killing White. Plum has the general ability to regulate his behavior by moral reasons, but in his circumstances the

egoistic reasons weigh very heavily for him, and consequently he is causally determined to murder White. But at the same time he does not act because of an irresistible desire.

Here again, although Plum meets each of the compatibilist conditions, it is intuitive that he is not morally responsible. Thus Case 2 also shows that the prominent compatibilist conditions, either separately or in conjunction, are not sufficient for moral responsibility. Furthermore, 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 or fifty years before the action seems irrelevant to the question of his moral responsibility. Causal determination by factors beyond Plum's control plausibly explains his lack of moral responsibility in the first case, and I believe that 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 an 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 so that he is often but not exclusively rationally egoistic (exactly as 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 his character. Consequently, Plum is causally determined to engage in the reasons-responsive process of deliberation and to have the first- and second-order desires that result in his killing White. Here again he has the general ability to grasp, apply, and regulate his behavior by moral reasons, but in these circumstances the egoistic reasons are very powerful, and so the training practices of his upbringing, in conjunction with background circumstances, deterministically result in his act of murder. Nonetheless, he does not act on an irresistible desire.

If a compatibilist wishes to contend that Plum is morally responsible in Case 3, he needs to point

to a feature of these circumstances that would explain why he is morally responsible here but not in Case 2. But it seems that there is no such feature. In each of these examples, Plum meets all the prominent compatibilist conditions for morally responsible action, so a divergence in judgment about moral responsibility between these examples will not be supported by a difference in whether these conditions are satisfied. Causal determination by factors beyond his control most plausibly explains the absence of moral responsibility in Case 2, and we are constrained to conclude that Plum is not morally responsible in Case 3 for the same reason.

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

Case 4: Physicalist determinism is true—everything in the universe is in some sense physical, and every event is rendered inevitable by virtue of the past states of this physical universe in conjunction with causal processes governed by laws of nature. Plum is an ordinary human being, raised in normal circumstances, and he is typically but not exclusively rationally egoistic (just as egoistic as in Cases 1–3). His act of murdering White results from his engaging in the reasons-responsive process of deliberation, and he has the specified first- and second-order desires. Although he possesses the general ability to grasp, apply, and regulate his behavior by moral reasons, in these circumstances the egoistic reasons weigh very heavily for him, and he is thus causally determined to kill White. However, it is not due to an irresistible desire that he commits this act of murder.

Given that we need to deny moral responsibility in Case 3, could Plum be responsible in this more ordinary case? There would seem to be no differences between Case 3 and Case 4 that could serve to justify the claim that Plum is not responsible in Case 3 but is in Case 4. One distinguishing feature of Case 4 is that the causal determination of Plum's crime is not brought about by other agents (Lycan 1997: 117–18). However, the claim that this is a relevant difference is implausible. Imagine

a further case that is exactly the same as, say, Case 1 or Case 2, except that Plum's states are induced by a spontaneously generated machine—a machine that has no intelligent designer. Here also Plum would not be morally responsible.

The best explanation for the intuition that Plum is not morally responsible in the first three cases is that his action is produced by a deterministic causal process that traces back to factors beyond his control. Because his action is also causally determined in this way in Case 4, we should conclude that here again he is not morally responsible. So by this argument, Plum's non-responsibility in Case 1 generalizes to non-responsibility in Case 4. We should conclude, I think, that if an action results from any deterministic causal process that traces back to factors beyond the agent's control, then she will lack the control required to be morally responsible for it.

3. EVENT-CAUSAL LIBERTARIANISM AND THE LUCK OBJECTION

Let us now consider libertarianism, the variety of incompatibilism that claims that we do have the sort of free will required for moral responsibility. There are two major types of libertarianism, the event-causal and the agent-causal versions. In event-causal libertarianism, actions are caused solely by events—events such as *Joe's desiring at noon to have lunch*, or *Mary's believing today that if she sells her stock tomorrow she will maximize her gains*. Now it is often supposed that all causation in the physical world is by events, and not by *things*, such as stars, machines, and agents, which we call *substances*. Although one might say, for example, that a bomb—a substance—caused damage to the building, when we want to speak more accurately, we say instead that *the bomb's exploding at a certain time*—an event—caused the damage. So if we think carefully 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 libertariansim maintains that all actions are caused solely by events, and further, that some type of indeterminacy in the production of actions by appropriate

events is the decisive requirement for moral responsibility (Kane 1996; Ekstrom 2000).

Critics of libertarianism have contended that if actions are not causally determined, agents cannot be morally responsible for them. Its 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: 411–12). The key idea is that if, holding fixed all of the conditions that precede an action, the action could either occur or not, then whether it does occur is a matter of chance or luck, and the agent cannot be morally responsible for it (Hume 1739/1978: 411–12).

I contend that event-causal libertarianism is undermined by the luck objection. Intuitively, for an agent to be morally responsible for a decision, she must exercise a certain type and degree of control in making that decision. In an event-causal libertarian picture, the relevant causal conditions prior to a decision—events that involve the agent—leave it open whether this decision will occur, and the agent has no further causal role in determining whether it in fact does. Accordingly, whether the decision occurs or not is in this sense a matter of luck, and the agent lacks the control required for being morally responsible for it.

To illustrate, consider Robert Kane's example of a businesswoman—let's call her Anne—who has a choice between stopping to help out an assault victim, as a result of which she would be late for an important meeting, or not stopping, which would allow her to make it to a meeting on time (Kane 1996). For simplicity, suppose the causally relevant events that immediately precede the action are, against stopping, *Anne's desiring not make her boss angry*, and *Anne's believing that if she is late for the meeting her boss will be angry with her*; and for stopping, *Anne's desiring to help people in trouble*, and *Anne's belief that she can help the assault victim*. Imagine that the motivational force of each of these pairs of prior events is for her about the same. On an event-causal libertarian view, with the causal influence of these events already in place, both Anne's deciding to stop and her not deciding to stop remain significantly probable outcomes. Suppose that she in fact decides to

stop. There is nothing else about Anne that can settle whether the decision to stop occurs, because in this theory her role in producing a decision is exhausted by these prior events. If nothing about Anne can settle whether the decision occurs, then she will not have the control required for moral responsibility for it. This might be called *the problem of the disappearing agent*. On an event-causal libertarian theory, no provision allows the agent to have the right sort of control over whether the decision occurs, 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 moral responsibility for it. One 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 prerequisite for moral responsibility—the absence of causal determinism for decision and action—but it would not supply another—sufficiently enhanced control (Clarke 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.

4. AGENT CAUSAL LIBERTARIANISM AND AN OBJECTION FROM OUR BEST PHYSICAL THEORIES

What needs to be added to the event-causal libertarian story is a further causal involvement of the agent in the making of the decision, a causal involvement that would enhance her control in making a decision over what is present in event-causal and deterministic contexts. It is this enhanced control that would solve the problem of the disappearing agent, which is highlighted by the luck objection against event-causal libertarianism. The agent-causal libertarian's solution is to specify a way in which the agent could have this enhanced control. 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 (Chisholm 1976; O'Connor 2000; Clarke 2003). 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 do so. The proposal is that if Anne had this power, by exercising it she would be able to settle which of the two competing decisions occurs—both of which remain as open possibilities given only the causal role of the events. In this way she would have the enhanced control required to be morally responsible for her decision.

But can agent-causal libertarianism be reconciled with what we would expect given our best physical theories? If the agent-causal position is true, then 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 example in the agent's brain or some other part of her body, are produced. But if the physical world were generally governed by deterministic laws, it seems that here we would 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 the 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 happens that diverges from these laws (Kant 1781/1997: 532–46). But this proposal would seem to involve coincidences too wild to be believed. For this reason, agent-causal libertarianism is not plausibly reconciled with the physical world's being governed by deterministic laws.

On the standard interpretation of quantum mechanics, however, the physical world is not in fact deterministic, but is rather governed by probabilistic statistical laws. Some philosophers have defended the claim that agent-causal libertarianism can be reconciled with physical laws of

this sort (Clarke 2003: 181). However, 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 percent 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 percent of the time. Are free choices on the agent-causal libertarian model compatible with what the statistical law leads 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 close to 32 percent 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 percent 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 run so sharply counter to what we would expect as to make it incredible.

At this point, the libertarian might propose that there actually do exist divergences from the probabilities that we would expect without the presence of agent-causes, and that these divergences are to be found at the interface between the agent-cause and that which it directly affects—an interface that is likely to be found in the brain. The problem for this proposal, however, is that we have no evidence that such divergences occur. This difficulty, all by itself, provides a strong reason to reject this approach.

It is sometimes claimed that our experience of deliberating and choosing provides us with good evidence for the broader thesis that we have libertarian free will. Perhaps, then, if we could have libertarian free will only if we were agent causes, then this evidence from our experience would count in favor of the existence of divergences from what our best physical theories predict. But Spinoza remarks, "experience itself, no less than reason,

teaches that men believe themselves free because they are conscious of their own actions, and ignorant of the causes by which they are determined . . ." (Spinoza 1677/1985: 496). Spinoza maintains that we believe our decisions are free only because we are ignorant of their causes. The lesson to draw from Spinoza here is that the evidence from experience that is apt to generate a belief that we have libertarian free will would be just the same if decisions were instead causally determined and we were ignorant of enough of their causes. For this reason, this evidence experience provides for our having libertarian free will is not especially impressive. This consideration counts strongly against the proposal that such evidence gives us reason to believe that the divergences in question exist.

On the other hand, nothing we've said conclusively rules out the claim that because we are agent causes, there exist such divergences. We do not have a complete understanding of the human neural system, and it may turn out that some human neural structures are significantly different from anything else in nature we understand, and that they serve to ground agent causation. This approach may be the best one for libertarians to pursue. But at this point we have no evidence that it will turn out to be correct.

Thus all versions of libertarianism face serious difficulties. Earlier, we saw that compatibilism is vulnerable to an 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 view is not, I think, that there is significant empirical evidence that it is false, or that there is a good argument that it is somehow incoherent, and false for that reason. Rather, the 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 tolerate?

5. HARD INCOMPATIBILISM AND WRONGDOING

Accepting hard incompatibilism demands giving up our ordinary view of ourselves as blameworthy for immoral actions and praiseworthy for actions

that are morally exemplary. At this point one might object that giving up our belief in moral responsibility would have very harmful consequences, perhaps so harmful that thinking and acting as if hard incompatibilism is true is not a feasible option. Thus even if the claim that we are morally responsible turns out to be false, there may yet be weighty practical reasons to believe that we are, or at least to treat people as if they were.

For instance, one might think that if we gave up the belief that people are blameworthy, we could no longer legitimately judge any actions as wrong or even bad, or as right or good. But this seems mistaken. Even if we came to believe that some perpetrator of genocide was not morally responsible because of some degenerative brain disease he had, we would still maintain that his actions were morally wrong, and that it was extremely bad that he acted as he did. So, in general, denying blameworthiness would not at the same time threaten judgments of wrongness or badness, and, likewise, denying praiseworthiness would not undermine assessments of rightness or goodness.

Perhaps treating wrongdoers as blameworthy is often required for effective moral education and improvement. If we resolved never to treat people as blameworthy, one might fear that we would be left with insufficient leverage to reform immoral behavior. Still, this option would have us treat people as blameworthy—by, for example, expressing anger toward them because of what they have done—when they do not deserve it, which would seem *prima facie* morally wrong. If people are not morally responsible for immoral behavior, treating them as if they were would seem to be unfair. However, it is possible to achieve moral reform by methods that would not be 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 deserve blame, the hard incompatibilist 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 can 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 yield strong reasons to treat criminals as if they were morally responsible. First of all, if hard incompatibilism is true, a retributivist justification for criminal punishment is unavailable, for it asserts that the criminal deserves pain or deprivation just for committing the crime, while hard incompatibilism denies this claim. And retributivism is one of 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. So one might suggest that the hard incompatibilist could endorse a view of this kind. However, we lack significant empirical evidence that punishing criminals brings about moral education, and without such evidence, it would be wrong to punish them in order to achieve this goal. In general, it is wrong to harm a person for the sake of realizing some good in the absence of impressive evidence that the harm will produce the good. Moreover, even if we had impressive evidence that punishment was effective in morally educating criminals, we should prefer non-punitive ways of achieving this result, if they are available—whether or not criminals are morally responsible.

Deterrence theories have it that punishing criminals is justified for the reason that it deters future crime. The two most-discussed deterrence theories, the utilitarian version and the one that grounds the right to punish on the right to self-defense, are not undermined by hard incompatibilism *per se*. Still, they are questionable on other grounds. The utilitarian theory, which claims that punishment is justified because it maximizes utility (i.e., the quantity of happiness or pleasure minus the quantity of unhappiness or pain), faces well-known challenges. It would seem at times to require punishing the innocent when doing so would maximize utility; in certain situations it would appear to prescribe punishment that is unduly severe; and it would authorize harming people merely as means to the well-being, in this case the safety, of others. The

sort of deterrence theory that grounds the right to punish in the right of individuals to defend themselves against immediate threats (Farrell 1985: 38–60) is also objectionable. For when a criminal is sentenced to punishment he is most often not 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 of the right of self-defense.

There is, however, an intuitively legitimate theory of crime prevention that is neither undercut by hard incompatibilism, nor threatened by other sort 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.

Clearly, 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. In addition, 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. Moreover, it would 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 needed to guard against the danger he poses.

6. MEANING IN LIFE

If hard incompatibilism is true, could we legitimately retain a sense of achievement for what makes our lives fulfilled, happy, satisfactory, or worthwhile, and our hold on our hopes for making these sorts of achievements in our lives? (Honderich 1988) It might be argued that if hard incompatibilism is true, there can be no genuine achievements, for an agent cannot have an achievement for which she is not also praiseworthy. However, achievement, and our hope for achievement, is not as closely connected to praiseworthiness as this objection supposes. If an agent hopes to achieve success in some project, 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 at the same time the sense in which it is her achievement may be diminished. For example, if someone hopes that her efforts as a teacher 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 anything she does.

One might think that hard incompatibilism would instill an attitude of resignation to whatever the future holds in store. But this isn't clearly right. Even if what we know about our behavioral dispositions and our environment gives us reason to believe that our futures will turn 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 that we lack complete knowledge of our dispositions and environmental conditions. Imagine that someone reasonably believes that he has a disposition that might well be an impediment to realizing

something he hopes to achieve. However, 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 of his will allow him to transcend the impediment. For instance, imagine that someone aspires to become a successful politician, but he is concerned that his fear of public speaking will get in the way. He does not know 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. As a result, he might reasonably hope that he will get over his fear and succeed in his ambition. Given hard incompatibilism, if he in fact does overcome his problem and succeeds in political life, this will not be an achievement of his in as robust a sense as we might naturally suppose, but it will be his achievement in a substantial sense nonetheless.

Still, with Saul Smilansky one might contend that although determinism leaves room for a limited foundation of the sense of self-worth that derives from achievement or virtue, the hard incompatibilist'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, Smilansky thinks that it would be best for us to foster the illusion that we have free will (Smilansky 2000). Now I agree that there is a kind of self-respect that presupposes an incompatibilist foundation, and that it would be undercut if hard incompatibilism were true. I question, however, whether Smilansky is right about how damaging it would be for us to give up this sort of self-respect, and whether his appeal to illusion is required.

First, note 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 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 have their source in our volition. To be sure, we also value efforts that are voluntary in the sense that they are willed by us—in

productive work and altruistic behavior, and indeed, in the formation of moral character. However, does it matter very much 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. It is not implausible that it is formed to a significant degree as a result of upbringing, and moreover, the belief that this is so 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 understand that their good moral character is not their own doing, and that they do not deserve a great deal of 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 nevertheless deserve praise and respect for producing their moral character? I suspect that most would eventually be able to accept the truth without incurring much loss. All of this, I think, 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.

7. EMOTIONS, REACTIVE ATTITUDES, AND PERSONAL RELATIONSHIPS

Peter Strawson (1962) argues that the justification for judgments of blameworthiness and praiseworthiness has its foundation in what he calls the *reactive attitudes*, reactions to how people voluntarily behave—attitudes such as moral resentment, guilt, gratitude, forgiveness, and love. Moreover, because moral responsibility has

this kind 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 could 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 give up the reactive attitudes altogether, but in a limited range of cases we can adopt what he calls the “objective attitude,” a cold and calculating stance toward others, which he describes as follows:

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)

If determinism did imperil the reactive attitudes, and we were able to relinquish them, Strawson suggests that we would face the prospect of adopting this objective attitude toward everyone, as a result of which our interpersonal relationships would be undermined. Since 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, and hence we would never have sufficient reason to give up our reactive attitudes, and thus to stop regarding people as morally responsible.

Clearly, if we persistently maintained an objective attitude toward others, our relationships would be undermined. However, I deny that it would be appropriate to adopt this stance

if we came to believe that hard incompatibilism were true. In my conception, some of the reactive attitudes would in fact be challenged by hard incompatibilism, for 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. But I claim that the reactive attitudes that we would want to retain either are not threatened by hard incompatibilism 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 human relationships.

It is plausible that to a certain degree moral resentment and indignation are beyond our power to affect. Even supposing that a hard incompatibilist is thoroughly committed to morality and rationality, and that she is admirably in control of her emotions, she might still be unable to eliminate these attitudes. Thus as hard incompatibilists we might expect people to be morally resentful in certain circumstances, and we would not regard them as morally responsible for it. But we also have the ability to prevent, temper, and sometimes to dispel moral resentment, and given a belief in hard incompatibilism, we might attempt such measures for the sake of morality and rationality. Modifications of this sort, assisted by a hard incompatibilist conviction, might well be good for interpersonal relationships.

It might be objected 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 a person is wronged in a relationship, she typically has further attitudes not threatened by hard incompatibilism, attitudes whose expression can play the communicative role in question. These attitudes include being alarmed or distressed about what another has done, and moral sadness or concern for him. Moral resentment, then, is not obviously required for effective communication in personal relationships.

Forgiveness might appear to presuppose that the person being forgiven is blameworthy, and if

this is so, this attitude would also be undercut by hard incompatibilism. But certain key features of forgiveness are not endangered by hard incompatibilism, and they are sufficient to sustain the role forgiveness has in relationships. Suppose a friend repeatedly mistreats you, and because of this you decide to end your relationship with him. However, he then apologizes to you, indicating his recognition that his actions were wrong, his wish that he had not mistreated you, and his commitment to refrain from the immoral behavior. Because of this you decide not to end the friendship. In this case, the feature of forgiveness that is consistent with hard incompatibilism is the willingness to cease to regard past immoral behavior as a reason to weaken or end a relationship. The aspect of forgiveness that is undermined by hard incompatibilism is the willingness to disregard the friend's blameworthiness. But since she has given up the belief that we are morally responsible, the hard incompatibilist no longer needs a willingness to disregard blameworthiness in order to foster good relationships.

One might object that hard incompatibilism imperils the self-directed attitudes of guilt and repentance, and that this would be especially bad for relationships. In the absence of guilt and repentance, we would not only be incapable of restoring relationships damaged because we have done wrong, but we would also be kept from restoring our moral integrity. For without the attitudes of guilt and repentance, we would lack the psychological mechanisms that can play these roles. But note first that it is because guilt essentially involves a belief that one is blameworthy that this attitude would be threatened by hard incompatibilism. It is for this reason that repentance would also seem to be (indirectly) threatened, for feeling guilty would appear to be required to motivate repentance. Imagine, however, that you have acted immorally; still because you endorse hard incompatibilism, you deny that you are blameworthy. Instead, you acknowledge that you were the agent of wrongdoing, you feel sad that you have done wrong, and you deeply regret having acted as you did. In addition, 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 to seek out the help of others in sustaining your resolution. None of these measures are jeopardized by hard incompatibilism.

Gratitude would appear to presuppose that the person to whom one is grateful is morally responsible for a beneficial act, whereupon hard incompatibilism would undermine gratitude. But as in the case of forgiveness, certain aspects of this attitude would be unaffected, and these aspects can provide what is needed for good relationships. Gratitude involves, first of all, being thankful toward a person who has acted beneficially. It is true that being thankful toward someone usually involves the belief that she is praiseworthy for some action. Still, one can also be thankful to a small child for some kindness, without believing that she is morally responsible for it. This aspect of thankfulness could be retained even without the presupposition of praiseworthiness. Typically gratitude also involves joy as a response to what someone has done. But no feature of hard incompatibilism undermines being joyful and expressing joy when others are, for example, considerate or generous in one's behalf. Expressing joy can bring about the sense of harmony and goodwill often produced by gratitude, and thus here hard incompatibilism is not at a disadvantage.

Is the kind of love that mature adults have for each other in good relationships imperiled by hard incompatibilism, as Strawson's line of argument suggests? Consider first whether for loving someone it is important that the person who is loved has 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 choices. Moreover, when adults love each other, it is also very seldom, if at all, for these sorts of reasons. The reasons we love others are surely varied and complex. Besides moral character and behavior, features such as intelligence, appearance, style, and resemblance to others in one's personal history all might play a part. Suppose morally admirable qualities are particularly important in occasioning, enriching, and maintaining love.

Even if there is an aspect of love that we conceive as a deserved response to morally admirable qualities, it is unlikely that love would vanish or even be diminished if we came to believe that these qualities are not produced or sustained by freely willed decisions. Such admirable qualities are loveable whether or not we deserve praise for having them.

One might contend that we want to be freely loved by others—to be loved by them as a result of their free will. Against this, the love parents have for their children typically comes about independently of the parents' will altogether, and we don't think that love of this sort is deficient. Kane recognizes this fact about parents' love, and he acknowledges that romantic love is similar in this respect. However, he maintains that there is a kind of love we very much want that would not exist if all love were causally determined by factors beyond our control (Kane 1996: 88). The plausibility of Kane's claim might be enhanced by reflecting on how you would react upon discovering that someone you love was causally determined to love you by, say, a benevolent manipulator.

Setting aside *free* will for a moment, when does the will play any role at all in engendering love? When a relationship is disintegrating, people will at times decide to try to restore the love they once had for one another. When a student finds herself in conflict with a roommate from the outset, she might choose to take steps to improve the relationship. When a marriage is arranged, the partners may decide to do what they can to love each other. In these kinds of circumstances we might want others to make a decision that might produce or maintain love. But this is not to say that we would want that decision to be freely willed in the sense required for moral responsibility. For it is not clear that value would be added by the decision's being free in this sense. Moreover, although in some circumstances we might want others to make decisions of this sort, we would typically prefer love that did not require such decisions. This is so not only for intimate romantic relationships—where it is quite obvious—but also for friendships and relationships between parents and children.

Suppose Kane's view could be defended, and we did want love that is freely willed in the sense required for moral responsibility. If we in fact desired love of this kind, then we would want a kind of love that is impossible if hard incompatibilism is true. Still, the sorts of love not challenged by hard incompatibilism are sufficient for good 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 deeply devoted to each other, and whose friendship became close through their interactions, then the possibility of fulfillment through interpersonal relationships remains intact.

Hard incompatibilism, therefore, does not yield a threat to interpersonal relationships. It might challenge certain attitudes that typically have a role in such relationships. Moral resentment, indignation, and guilt would likely be irrational for a hard incompatibilist, since these attitudes would have presuppositions believed to be false. But these attitudes are either not required for good relationships, or they have analogues that could play their typical role. Moreover, love—the reactive attitude most essential to good interpersonal relationships—does not seem threatened by hard incompatibilism at all. Love of another involves, fundamentally, wishing for the other's good, taking on her aims and desires, and a desire to be together with her, and none of this is endangered by hard incompatibilism.

8. THE GOOD IN HARD INCOMPATIBILISM

Hard incompatibilism also promises substantial benefits for human life. Of all the attitudes associated with the assumption that we are morally responsible, anger seems most closely connected with it. Discussions about moral responsibility most often focus not on how we judge morally exemplary agents, but rather on how we regard those who are morally deficient. Examples designed to elicit a strong intuition that an agent is morally responsible most often feature an especially heinous action, and the intuition usually involves sympathetic anger. It may be,

then, that our attachment to the assumption that we are morally responsible derives to a significant degree from the role anger plays in our emotional lives. Perhaps we feel that giving up the assumption of responsibility is threatening because the rationality of anger would be undercut as a result.

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

Moral anger comprises a significant part of our moral lives as we ordinarily conceive them. It motivates us to resist abuse, discrimination, and oppression. At the same time, expression of moral anger often has harmful effects, failing to contribute to the well-being either of those toward whom it is directed or of those expressing the anger. Often its expression is intended to cause little else than emotional or physical pain. Consequently, it has a tendency to damage relationships, impair 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 gives rise to a strong demand that it be morally justified when it occurs. The demand to morally justify behavior that is harmful is generally a very strong one, and expressions of moral anger are often harmful. This demand is made more urgent by the fact that we are often attached to moral anger, and that we frequently enjoy expressing it. Most commonly we justify expression of moral anger by arguing that wrongdoers deserve it, and we believe that they deserve it because they are morally responsible for what they do. If hard incompatibilism is true, however, justification of this sort is undermined. Yet given the concerns to which expression of moral anger give rise, this may be a good thing.

Accepting hard incompatibilism is not likely to modify our attitudes to the extent that expression of moral anger ceases to be a problem for us. However, moral anger is often sustained and magnified by the belief that its object is morally responsible for immoral behavior. Destructive moral anger in relationships is nurtured in this way by the assumption that the other is blameworthy. The anger that fuels ethnic conflicts, for example, is almost always fostered by the conviction that a group of people deserves blame for past wrongs. Hard incompatibilism advocates giving up such beliefs because they are false. As a result, moral anger might decrease, and its expressions subside.

Would the benefits that would result if moral anger were modified in this way compensate for

the losses that would ensue? Moral anger motivates us to oppose wrongful behavior. Would we lose the motivation to oppose immorality? If for hard incompatibilist reasons the assumption that wrongdoers are blameworthy is withdrawn, the belief that they have in fact behaved immorally would not be threatened. Even if those who commit genocide are not morally responsible, their actions are nonetheless clearly horribly immoral, and a conviction that this is so would remain untouched. This, together with a commitment to oppose wrongdoing, would permit a resolve to resist abuse, discrimination, and oppression. Accepting hard incompatibilism would thus allow us to retain the benefits moral anger can also provide, while at the same time challenging its destructive effects.

REFERENCES

- Chisholm, Roderick (1976). *Person and Object* (La Salle: Open Court).
- Clarke, Randolph (2003). *Libertarian Theories of Free Will* (Oxford: Oxford University Press).
- Ekstrom, Laura W. (2000). *Free Will: A Philosophical Study* (Boulder: Westview).
- Farrell, Daniel M. (1985). "The Justification of General Deterrence," *The Philosophical Review* 104.
- Fischer, John Martin (1994). *The Metaphysics of Free Will* (Oxford: Blackwell).
- Frankfurt, Harry G. (1971). "Freedom of the Will and the Concept of a Person," *Journal of Philosophy* 68, pp. 5–20.
- Honderich, Ted (1988). *A Theory of Determinism* (Oxford: Oxford University Press).
- Hume, David (1739/1978). *A Treatise of Human Nature*; L. A. Selby-Bigge, ed. (Oxford: Oxford University Press).
- Kane, Robert (1996). *The Significance of Free Will* (New York: Oxford University Press).
- Kant, Immanuel (1781/1997). *Critique of Pure Reason*, tr. Paul Guyer and Allen Wood (Cambridge: Cambridge University Press, 1997).
- Lycan, William G. (1997). *Consciousness* (Cambridge: MIT Press).
- Mele, Alfred (2006). *Free Will and Luck* (Oxford: Oxford University Press).
- O'Connor, Timothy (2000). *Persons and Causes* (Oxford: Oxford University Press).
- Pereboom, Derk (2001). *Living Without Free Will* (Cambridge: Cambridge University Press).
- Schoeman, Ferdinand D. (1979). "On Incapacitating the Dangerous," *American Philosophical Quarterly* 16.
- Smilansky, Saul (2000). *Free Will and Illusion* (Oxford: Oxford University Press).
- Spinoza, Baruch (1677/1985). *Ethics*, in *The Collected Works of Spinoza*, ed. and tr. Edwin Curley, Volume 1 (Princeton: Princeton University Press).
- Strawson, Peter F. (1962). "Freedom and Resentment," *Proceedings of the British Academy* 48 pp. 1–25.
- van Inwagen, Peter (1983). *An Essay On Free Will* (Oxford: Oxford University Press).
- Wallace, R. Jay (1994). *Responsibility and the Moral Sentiments* (Cambridge, Harvard University Press).