

What is Human Freedom?

Lynne Rudder Baker

After centuries of reflection, the issue of human freedom remains vital largely because of its connection to moral responsibility. When I ask—What is human freedom?—I mean to be asking what kind of freedom is required for moral responsibility? Questions about moral responsibility are intimately connected to questions about social policy and justice; so, the issue of moral responsibility—of desert, of whether or not anyone is ever really praiseworthy or blameworthy—has practical as well as theoretical significance.

I start with two assumptions: (1) Human persons are a natural part of the natural world, and so are under whatever laws govern the rest of the natural world. (2) Nevertheless, human persons are free in a way that other beings are not. With these two assumptions, I shall first say what human freedom is not; then I shall explain what I think human freedom is. Then I shall formulate sufficient conditions for moral responsibility, and conclude with some thoughts on the pervasiveness of luck.

First, let me make some terminological remarks. Some philosophers think that the freedom needed for moral responsibility is compatible with determinism; such philosophers are called ‘compatibilists.’ Other philosophers think that the freedom needed for moral responsibility is not compatible with determinism; such philosophers are called ‘incompatibilists.’ Incompatibilists divide into two kinds: one kind of incompatibilist thinks that we have the freedom needed for moral responsibility, and hence that determinism is false. Such philosophers are called ‘libertarians’. The other kind of incompatibilist thinks that determinism is true,

and hence that we do not have the freedom needed for moral responsibility. Such philosophers are called ‘hard determinists.’ I’ll argue for a compatibilist position—mainly against libertarians, and at the end briefly against hard determinists.

I. What Human Freedom is Not

Many philosophers believe that the freedom needed for moral responsibility for an action is free will, understood in a particularly stringent way that precludes determinism. Determinism is the thesis that the future is completely determined by the laws of nature together with antecedent conditions. Those who understand freedom this way, and believe that we have such freedom are called ‘libertarians.’ What is such freedom? Here are some recent formulations of free will by prominent libertarians:

- (i) Free will entails “the power of agents to be the ultimate creators (or originators) and sustainers of their own ends or purposes.” (Kane 1998: 4.)
- (ii) Free will with respect to an action entails that the “action is causally brought about by something that (a) is not itself causally brought about by anything over which she has no control, and (b) is related to her in such a way that, in virtue of its causing her action, she determines which action she performs.” Clarke 1993: 203).
- (iii) Free will with respect to an action entails that the agent’s “own intellect and will are the sole ultimate source or first cause of her act.” (Stump 2001: 126.)

(iv) Free will with respect to an action X at time t entails that it is within the agent's power at t both to perform X and to refrain from performing X. (Plantinga 1974: 165-6.)

Whether they regard incompatibility with determinism as a defining condition, or only as a consequence of the nature of free-will, all libertarians agree with (v) and (vi) that an exercise of free will is not causally determined by anything outside the agent. The other formulations suggest two distinct underlying intuitions behind libertarianism. Condition (iv) suggests the Principle of Alternative Possibilities: An agent has free will with respect to an action (and is morally responsible for it) only if the agent could have done otherwise in a nonhypothetical sense. There have been endless discussions over the proper interpretation of 'could have done otherwise.' (Berofsky 2002) Since I think that just about everything that can be said about 'could have done otherwise' has been said, and also that the Principle of Alternative Possibilities has been decisively refuted,¹ I shall put aside the controversy over the Principle of Alternative Possibilities and focus on the other libertarian intuition—one which, if correct, supersedes the Principle of Alternative Possibilities anyway.²

Formulations (i), (ii), and (iii) suggest a [a libertarian intuition different from the Principle of Alternative Possibilities: An agent has free will with respect to an action (and is morally responsible for it) only if the agent has power or control of her actions in an especially weighty way.³ What kind of control is at issue for the libertarian?

In ordinary English, we say that the driver had control of the car (Fischer 1994: 132ff), or that the professor had the power to give all As and the power to refrain from giving all As. Most of us have control over our limbs while we're

awake. These ordinary examples of power or control, however, are clearly compatible with determinism. The kind of power or control that a libertarian demands is much stronger than our ordinary concept yields. The requisite power or control is originative. The libertarian free agent has power or control over the ultimate sources of her action.

Free will, according to libertarians, requires that a free choice or action not have its origin in anything beyond the agent's control. Otherwise, according to prominent libertarian Robert Kane, "the action, or the agent's will to perform it, would have its source in something the agent played no role in producing. Then the arche [sufficient ground or cause or explanation] of the action, or of the agent's will to perform it, would not be 'in the agent,' but in something else." (Kane 1998: 35.) "[T]o will freely, in this traditional sense," Kane says, "is to be the ultimate creator (prime mover, so to speak) of your own purposes." (Kane 1998: 4.) Roderick Chisholm, a proponent of agent-causation, put it this way:

If we are responsible, and if what I have been trying to say is true, then we have a prerogative which some would attribute only to God: each of us, when we act, is a prime mover unmoved. In doing what we do, we cause certain events to happen, and nothing—or no one—causes us to cause those events to happen.⁴

One of the recent advances in the free-will debate has been disentanglement of the Principle of Alternative Possibilities from the condition of originative control endorsed by Kane, Chisholm, Clarke and others. Although compatibilists (i.e., those who believe that human freedom is compatible with determinism) have led the attack on the Principle of Alternative Possibilities, some libertarians have pointed out that one can reject the Principle of Alternative Possibilities, without

rejecting libertarianism. (Hunt 2000; Kane 1998, Ch. 3; Stump 1999; Pereboom 2001, Ch. 1.) The libertarian intuition about ultimate origination is deeper than the one codified in the Principle of Alternative Possibilities. The attacks of the Principle of Alternative Possibilities, led by Frankfurt, do not threaten this deeper libertarian conception of freedom and moral responsibility at all.⁵

The deeper libertarian conception of free will and moral responsibility is thus not a matter of alternative possibilities, nor is it simply a matter of indeterminism. The mere occurrence of an undetermined event in the causal history of an action does not make the action free by anyone's lights. (We need to know what 'the power to do A or to refrain from doing A' consists in if it is not the ordinary power—say, to wear boots or to refrain from wearing boots.) The fact that the causal history of an action contains an event governed by a statistical law rather than a universal law cannot render an agent morally responsible for the action. Nor, on a non-nomological conception of causation, would the fact that an action or decision was caused by a random event render an agent morally responsible for it. An indeterministically caused decision is no more within the agent's control than is a deterministically caused decision. What is needed—as recent work on libertarianism has shown—is the notion of ultimate origination or ultimate control.

But what is ultimate origination? Consider Kane again, as he speaks of “the power of agents to be the ultimate creators (or originators) and sustainers of their own ends or purposes:”

[W]hen we trace the causal or explanatory chains of action back to their sources in the purposes of free agents, these causal chains must come to an end or terminate in the willings (choices, decisions, or efforts) of the agents,

which cause or bring about their purposes. If these willings were in turn caused by something else, so that the explanatory chains could be traced back further to heredity or environment, to God, or fate, then the ultimacy would not lie with the agents but with something else. (Kane 1998: 4.)

This suggests that an agent is the ultimate originator of a choice or action if and only if the sufficient condition for the choice or action ineliminably includes something over which the agent has complete control. Ultimate originators have ultimate control. Of course, agents do not have ultimate control over all the necessary conditions for their actions. However, if there is a sufficient condition for the choice or action over which the agent lacks ultimate control, the agent is not the ultimate originator of it.

Perhaps someone will object that libertarianism does not require that agents of free actions be ultimate originators. I would reply: Not only have I cited a number of libertarians who are explicitly committed to the idea of ultimate origination, but also without ultimate origination, the libertarian would have no better grounding for moral responsibility than does the compatibilist. Beliefs, desires, character, upbringing, genetic make-up and so on are outside the agent's control. If you remove the requirement of ultimate origination and suppose that these causal factors are sufficient for an agent's free action or choice, your view is tantamount to compatibilism since these causal factors are beyond the agent's control. So, ultimate control is required for libertarianism.

Before turning to arguments against libertarianism, let me give an example to suggest that our moral practices are at odds with libertarianism: In the 19th century, the institution of slavery set the prevailing ethos of the American South. Children were typically brought up strictly in an environment dominated by race.

Many 19th-century U.S. Southerners believed that slavery is sanctioned by God (part of the “orders of creation”), and by the U.S. Constitution and by common sense.⁶ In 1863, a cavalry lieutenant from Mississippi reaffirmed his belief that “this country without slave labor would be wholly worthless....We can only live & exist by that species of labor: hence I am willing to continue the fight to the last.” (McPherson 2001: 32.) Given his beliefs and desires—and the fact that there are no competing ones—it would have been irrational for him to have stopped fighting, whether he had libertarian freedom or not and whether he had ultimate control over his decisions or not. Since I assume that the lieutenant was acting rationally on this occasion, even if he could have done otherwise, he would not have done otherwise—not without a change in antecedent conditions.⁷

But our practices hold people who fight for a cause because they decide to without coercion morally responsible for fighting—with no inquiry into whether they have had ultimate control over their decisions. (Indeed, it is difficult to see how such an inquiry could even proceed.) Ultimate control is simply irrelevant to our moral practices. The lieutenant indicates that his beliefs and desires are sufficient for his intention to “continue to fight to the last,” and our ordinary moral practices take them to be sufficient. Otherwise, before holding someone morally responsible for an action, we should inquire into its history to see whether there was some “input” over which the agent had ultimate control. In any case, a libertarian must either deny that the lieutenant was morally responsible for fighting, or else reject the (plausible) description of the case, a description that fully accounts for his action in terms of attitudes and values. There would be a high cost for our moral practices either way. [ADD: Indeed, since ultimate origination seems undetectable and unverifiable, I cannot see how any moral practices could conform to a requirement of ultimate control.]

Let us now turn from pragmatic considerations to direct arguments against libertarianism. Now I shall argue that no human person is ultimate originator of any action, decision or choice.⁸

There are two main kinds of libertarian theories: Agent-causal theories and event-causal theories.⁹ According to agent-causal theories, the causation of “free choices or actions cannot be explained as the causation of events or occurrence by other events or occurrences.” (Kane 1989: 118.) An agent-causal free action “is such that its occurring rather than not here and now, or vice versa, has as its ultimate or final explanation the fact that it is caused by the agent here and now.” (Kane 1989: 120. Emphasis his.) According to event-causal libertarian theories, all causation is causation by events. Event-causal libertarian theories hold that for an agent to cause her free action or choice is for there to be an event involving the agent that causes the action or choice. But both kinds of libertarian theories agree that a free action or choice has its *final or ultimate* explanation in the agent herself: The agent produced it by, e.g., making an effort of will (event-causal theories) or the agent just caused it *simpliciter* (agent-causal theories).¹⁰ There is no further explanation to be had.

Event-causal libertarians like Kane take indeterminism to make room for our being ultimate originators. Agents in an undetermined world have “the power to make choices which can only and finally be explained in terms of their own wills.” (Kane 1989: 129; emphasis his.) Kane takes efforts of will to be indeterministic processes that terminate intentionally in a choice that is undetermined. These indeterministic efforts of will (Should I or shouldn’t I?) are correlated with indeterministic macroprocesses in the brain, which in turn result from amplification of microindeterminacies. (Kane 1989: 129) Kane says that when “the indeterminate effort becomes determinate choice, the agent will make one set

of reasons or the other prevail then and there by deciding.” (Kane 1989: 134; emphasis his.)

But this raises a question of intelligibility: How can an undetermined event that terminates an indeterministic physical process be produced by an agent’s deciding? It is wholly obscure how an agent can intentionally bring an indeterministic brain process to a conclusion by making a choice.¹¹ Since Kane’s view is the most detailed formulation of event-causal libertarianism that I know of, I do not believe that event-causal libertarianism will allow an agent to be the ultimate originator of an action.

In general, we have no more control over an event produced by an indeterministic process than we do over an event produced by a deterministic process. Part of the motivation for libertarian theories is that libertarians take determinism to threaten free will as they see it. The present point, not original with me, is that if determinism threatens free will, then so does indeterminism. (See, e.g., Strawson 2000; Double 1991; van Inwagen 1975.)

Agent-causal theories also take agents to be ultimate originators. Clarke echoes Chisholm’s conception of an agent as an “uncaused cause.” (Clarke 1993: 201.) Before giving a general argument against the whole libertarian conception of free agents as ultimate originators, let me note that agent-causal theories also founder on a different shoal. According to agent-causal theories, there is a fundamental causal power that allows agents to be the source (i.e., to cause) their free decisions and actions. The exercise of this agent-causal power cannot be reduced to causation between events. The agent, a substance, is the cause of an event—not in virtue of any event involving the agent, but simply by exercise of agent-causal power. In this way, the agent can initiate new causal chains—causal

chains with a first event caused directly by the agent. The problems with agent-causal theories are well-known: The theories are obscure; they simply postulate a power that meets libertarian conditions for free will without making plausible how we could have such an amazing power. Although I am sympathetic to the claim that there are emergent properties, appeal to emergent properties in this context does not explain (to me) how there could be libertarian agent-causation. (See O'Connor 1995; O'Connor 2000.) The other problem is that agent-causation seems at odds with the dominant worldview that is broadly “naturalistic.”

One of my explicit initial assumptions is that human persons are part of nature. Although there are many versions of “naturalism” with which I have little sympathy, I differ from some traditional metaphysicians who want to insulate metaphysical theories from developments of science.¹² Without going so far as to regard science as the arbiter of reality, the successes of the sciences give us good reason to take the sciences to be the source of knowledge of physical mechanisms. Although the brain, a physical organ, still seems largely a mystery (especially in relevant areas of, say, complex motivation), enough is known about the brain in broad outline to make agent-causation wholly implausible, or seemingly magical.¹³ If the ‘agent-causation’ hypothesis were true, then there should be neural gaps—brain events with no causal (event-)antecedents. In order to be useful to proponents of agent causation, such neural gaps must be different from the ordinary quantum indeterminacy found in inanimate objects. But despite burgeoning neuroscientific research on free will, there is no evidence of relevant neural gaps.¹⁴ Therefore, we should reject agent-causal libertarian theories.¹⁵

The libertarian holds that if an action is free, it is not explainable by the sum of the kinds of causal conditions that psychologists and social scientists appeal to—such as how an agent sees her situation, what she takes her options to be, her

other beliefs, her desires, her character, her experience, her genetic inheritance, her environment. Taken together, these cannot be sufficient for a libertarian free action. Libertarian free action awaits exercise of an ability that no natural being has: the ability to rise above the complex mix of causes (heredity, environment, beliefs, desires, etc.) and interject an unexplainable X factor, over which the agent has ultimate control, and which renders theretofore-insufficient causes sufficient for the choice or action. An ultimate originator seems not to be a part of the natural order.

In short: Ultimate origination is an illusion. No one has it. But without ultimate origination or ultimate control, libertarianism would be on a par with compatibilism vis-à-vis moral responsibility. So, we are justified in rejecting libertarianism.

II. What Human Freedom Is

Since one of my explicit initial assumptions was that we human persons are free in a way that other creatures are not, we should look for the source of our freedom in what distinguishes us from all other natural creatures. Indeed, what sets us apart from other creatures makes freedom a kind of default for our choices and actions. As I have argued in Persons and Bodies, persons are unique in having first-person perspectives. A first-person perspective is the ability to think of oneself from the “inside,” so to speak.

On my view, a first-person perspective is the defining characteristic of persons.¹⁶ A first-person perspective is the ability to entertain a certain kind of thought about oneself—not just a thought about someone who happens to be oneself, but about oneself as oneself: the person *per se*, without recourse to any name or description. Oedipus’s realization that he himself was the killer of Laius,

as opposed to his realization that a man at the crossroads was the killer of Laius, required a first-person perspective. As the example of Oedipus indicates, a first-person perspective cannot be eliminated or reduced to a third-person perspective without cognitive loss.

First-person perspectives enable us to reflect on and evaluate our desires, and on possible courses of action. We can resist temptation (sometimes). Dogs and other higher mammals simply act on their desires: there's an attractive potential mate; go for her. They do not deliberate and then decide to turn away. By contrast, we are able to formulate the thought: I don't want to be the kind of person who does such-and-such. We can be motivated by such thoughts. A first-person perspective is what is distinctive about persons.¹⁷ We not only make choices; we evaluate our choices--regardless of whether our evaluations are themselves caused by forces behind our control or not. First-person perspectives are also the source of our freedom and moral responsibility. Here is what I take human freedom to be:

(HF) Human freedom is the unique ability, made possible by a first-person perspective, to reflect on and evaluate our desires and to choose one course of action over another.

Much of the interest in the question of free will, as I mentioned at the outset, stems from concern about moral responsibility. Some philosophers believe that moral responsibility requires libertarianism. In light of the considerations just given, I believe that such philosophers should conclude that no one is ever morally responsible for anything.¹⁸ But logically speaking, such a dire conclusion is not forced. We are equally entitled to conclude that libertarianism is not necessary for moral responsibility.

I shall argue for moral responsibility without libertarianism. To do this, I shall set out and defend nonlibertarian conditions for moral responsibility in terms of reflective endorsement. Libertarians (and other incompatibilists) often charge that compatibilist conditions for moral responsibility are only necessary, not sufficient. I'll try to meet this charge by supplying sufficient conditions for moral responsibility that are compatible both with determinism and with indeterminism.¹⁹

The close tie between human freedom and moral responsibility will be apparent: anyone who is morally responsible for a choice or act is free with respect to that choice or act. I am not arguing from (HF) to moral responsibility. I take (HF) to be what human freedom is—an ability. Anyone who satisfies the sufficient conditions for moral responsibility with respect to a choice or action ipso facto exercises the ability specified in (HF) on the occasion of choosing or acting.

The conditions that I hold to be sufficient for moral responsibility are based on those of Harry Frankfurt. Frankfurt offers a hierarchical view of the will: We not only have first-order desires; we also have second-order desires. (Frankfurt, 1971) We not only want to do certain things; we also want to want certain things, to be moved by certain desires. A person's will is the desire which actually motivates him: To will X is to have an effective desire to choose or do X, a desire that moves one to choose or to do X. If one has the will that one wants to have (if, that is, one is moved by the desire that one wants to move one), then on Frankfurt's view, one has free will.²⁰

One important feature of Frankfurt's view of the hierarchical will that has gone largely unremarked is that it requires that the agent have a first-person perspective. A person must be able to conceive of her desires *as her own*—from the first-person—if she is to desire to have a certain desire. If Sally wants to be

moved by a certain desire, then she wants that she herself (considered from the first-person) be moved by that desire. To want that someone who just happens to be herself be moved by that desire is not enough. If I can make the distinction between wanting that I (myself) be moved by a certain desire and wanting that S (where 'S' is a third-person name for me) be moved by that desire, then I have a first-person perspective.²¹ As we shall see, only beings with first-person perspectives can be morally responsible.

The importance of the first-person perspective for moral responsibility is that it gives us limited control over our desires. A first-person perspective makes it possible for us, unlike the other animals, to discover what goals we have, to evaluate them and to try to change them. Animals without first-person perspectives do not have this control: they act on their desires, but they cannot set about changing them. But persons can even interfere with biological goals of survival and reproduction. To a significant extent, a person can know what desires she has: she approves of some and is repelled by others. We have partial control, albeit limited by the kinds of people we already are, over our desires. Our partial control is manifest in our (sometimes successful) efforts to change. Yet, such effort—which is limited by our heredity, environment, experience, even by our insight and imagination—is itself a product of factors beyond our ultimate control.

This partial control over our desires has the consequence that, even if determinism is true, we are not mere conduits for causes beyond our control. Because of our first-person perspectives, we have a hand in what causes our actions: we can modify our first-order desires that produce our intentions and in this way to help shape the causes of what we do. We can decide to try to be one kind of person rather than another, to be generous rather than stingy, say. For incompatibilists—both libertarians and hard determinists—control is all or

nothing: either we have ultimate control or we are mere puppets. Attention to the first-person perspective should make us reject this dichotomy: we are neither puppets nor ultimate originators. We can try to become the people that we want to be—even though our wanting to become a certain kind of person, along with the trying itself, is caused by factors beyond our control.

The first-person perspective cannot be acquired by neural manipulation—any more than a disposition to be honest or the ability to read Italian or any other intentional disposition can be acquired by neural manipulation. I am supposing that a person at *t* (when he is not actually reading Italian) may have the ability to read Italian without there being anything in his brain at *t* that makes it the case that a person can read Italian. I doubt that putting the brain in a given state is ever sufficient for having an intentional disposition. Someone can have a first-person perspective only if he has consciousness and has had many kinds of intentional states. Moreover, distinct first-person perspectives may be qualitatively similar.

Frankfurt's conditions for moral responsibility clearly require a first-person perspective: Frankfurt says: "Suppose that a person has done what he wanted to do, that he did it because he wanted to do it, and that the will by which he was moved when he did it was his will because it was the will he wanted." (Frankfurt 1971): 19.) Then, if one "identifies" (Frankfurt's word) oneself with a first-order desire that motivates one to do *X*, then, at least *prima facie*, one is responsible for doing *X*. Every piece of this account requires a first-person perspective. Using a '*' to mark our attribution to the agent of a the first-person perspective: *S*'s doing what he* wanted to do, *S*'s doing it because he* wanted to, *S*'s having the will that he* wanted to have, *S*'s identifying himself* (not equivalent to *S*'s identifying *S*) with a first-order desire that motivated him—all these are manifestations of a first-

person perspective. So, Frankfurt's conditions for moral responsibility require that the agent have a first-person perspective.

Suppose that in addition to satisfying Frankfurt's conditions, the agent endorsed her desires, knowing that they resulted from things that happened before she was born; suppose that she identified herself with the factors beyond her control that contributed to her desires and beliefs. Suppose that the agent grew up in a martial family, and that her family was the source of her desire to be a combat soldier: If she had had a different kind of family, she would not have wanted to go into combat. Of course, she had no control over what kind of family she was born into. She was nonetheless morally responsible for becoming a combat soldier. To handle such a case, I'll add to Frankfurt's condition this one: "S would still have wanted to will X even if she* had known the provenance of her wanting to will X."

²² Now we have compatibilist conditions for free will, which, I submit, become sufficient conditions for moral responsibility. I'll call this view the Reflective-Endorsement view:

(RE) A person S is morally responsible for a choice or action X if X occurs and:

- (i) S wills X and S's willing X causally contributes to the occurrence of X,
- (ii) S wants that she* will X [i.e., S wants to will X],
- (iii) S wills X because she* wants to will X, and
- (iv) S would still have wanted to will X even if she had known the provenance of her* wanting to will X.

Conditions (ii) – (iv) of the Reflective Endorsement view each requires that S have a first-person perspective.²³ Moreover, these conditions are compatible with the

truth of determinism (as well as with the truth of indeterminism). The causal histories of S's desires and willings may be traceable to factors outside of S's control. Even if S could have done otherwise, she would not have done otherwise. Since "the will that moved her when she acted was her will because she wanted it to be, she cannot claim that her will was forced upon her or that she was a passive bystander to its constitution."²⁴ If S satisfies conditions (i) – (iv), she is morally responsible for her choice or action X and her choice or action is therefore free. Conditions sufficient for being morally responsible with respect to a particular choice or action are also sufficient for being free with respect to that choice or action.

As an illustration of the Reflective-Endorsement view, consider the real-life case of Bobby Frank Cherry, who was recently convicted in the bombing that killed four African-American Sunday-School girls in a church in Birmingham, Alabama, in 1963. Suppose that Cherry (i) willed to participate in the bombing. As a convinced white supremacist (who apparently bragged of his participation), he (ii) wanted to will to participate, and he (iii) willed to participate because he wanted to. He (iv) would still be proud of his participation, and would participate again, even though he knew that his wanting to will to participate in the bombing had been caused by his racist upbringing. ("Damn right," he might have said, "and I'm bringing up my boys the same way.") It seems to me obvious that he was morally responsible for his participation in the bombing. He was moved by the desire (to bomb the black church) that he wanted to be moved by. Cherry satisfies (i) – (iv) of (RE), and (I believe that) we properly find him morally responsible.

My line of thought is this: Human freedom is an ability that one must have in order generally to be morally responsible for a choice or action. To be morally responsible for a particular decision or action X, it is sufficient (but not necessary)

that one satisfy the four conditions of the Reflective-Endorsement view for the choice or action X. If one is morally responsible for some choice or action X, then ipso facto one is free with respect to X.

Freedom of action is the default with respect to ordinary action: When we do something intentionally with relevant knowledge of what we're doing and without interference or compulsion, we do it freely. We are usually morally responsible for what we do freely, but the voluminous literature has shown the difficulty in formulating compatibilist sufficient conditions for moral responsibility in terms of freedom of action. Hence, my conditions—which I think are sufficient for moral responsibility—are in terms of Reflective-Endorsement.

A basic conviction—shared by both libertarian and compatibilist—is that moral responsibility concerns the agent in a deep way. An agent is morally responsible for an action if he endorses the beliefs and desires on which he acts: When he affirms them as his own (and makes no factual errors about the circumstances, etc.), he is morally responsible for acting on them. Whereas the libertarian is concerned that a morally responsible agent have (impossible) control over factors that contribute to what he wills, the compatibilist is concerned that a morally responsible agent endorse and align himself* with what he wills. Doing what we want to do, with reflective endorsement, is all the control that we have, and that's enough.

If I can say, “These desires reflect who I am, and this is the kind of person that I want to be,” then I am morally responsible for acting on those desires—whether determinism is true or not. I have (nonultimate) control of what I do if, in the absence of psychopathology like kleptomania, I deliberate and make an uncoerced decision to do it.²⁵ Unlike the libertarian, the compatibilist does not

think that external causes of decisions or actions per se threaten moral responsibility; moral responsibility is precluded only by certain kinds of external causes—those that by-pass the agent’s own psychological contribution involving his first-person perspective. By focusing on the first-person perspective, a compatibilist can agree with the libertarian to this extent: Being morally responsible has to do with the agent in a deep way, not just with an external series of causes. Reflective Endorsement offers a via media between a (physically dubious) libertarianism, and a (morally hopeless) hard determinism that precludes moral responsibility.

III. Fending Off the Mad Neuroscientist

This discussion also indicates how a proponent of the Reflective-Endorsement View can respond to anti-compatibilist arguments invoking the infamous mad neuroscientist. The skeleton argument against compatibilism is this: A mad neuroscientist manipulates an agent’s brain, so that the agent wills to do X (or whatever would satisfy the compatibilist conditions on offer). Then, when the agent does X simply as a result of the neuroscientist’s manipulations, he satisfies the compatibilist conditions for moral responsibility; but he is not really morally responsible for doing X. So, it is claimed, compatibilism is false.

The Reflective-Endorsement view does not succumb to this argument. In the first place, in order to be a candidate for being morally responsible for doing something, a being must have a first-person perspective; and a neuroscientist cannot provide a being with a first-person perspective merely by manipulating his brain (any more than he could provide a being with an ability to speak Italian by manipulating his brain).

In the second place, if the neuroscientist works his will on a being who already has a first-person perspective, then that person either will or will not satisfy condition (iv) of (RE). On the one hand, if the agent whose brain is manipulated by the mad neuroscientist does not satisfy (iv)—if knowing of the manipulation, he would be appalled by his wanting to will X—then the Reflective-Endorsement view does not hold him morally responsible. On the other hand, if the agent does satisfy (iv)—if knowing of the manipulation, he would still want to will X—then the Reflective-Endorsement view holds him morally responsible, as indeed it should.

An anti-compatibilist may press on: What if satisfaction of condition (iv)—an agent's being such that he would still want to will X even if he knew the provenance of his wanting to will X—were itself produced by the mad neuroscientist? In this case, the mad neuroscientist would make the agent still endorse his willing X even if the agent knew that the endorsement and the willing were caused by a mad neuroscientist. This possibility prompts me to interpret 'knowing the provenance of the agent's endorsement of his willing X' as including a completeness clause: "There is no further knowledge of the circumstances of the agent's endorsement of his willing X that would lead the agent to repudiate his endorsement of his willing X." Now the agent who would have repudiated his endorsement of his willing X if the mad scientist had not prevented his repudiation does not satisfy (iv) after all, and hence is not deemed morally responsible by (RE). However, an agent who does satisfy (i)-(iv), where (iv) is interpreted as including this completeness clause, would not have repudiated his endorsement of willing X even if the mad neuroscientist not had prevented his repudiation. Such an agent is morally responsible for X.

So, an agent under the control of a mad scientist either satisfies all the conditions of Reflective-Endorsement or not. In neither case do we have a situation in which an agent satisfies (RE) without being morally responsible.²⁶ So, I believe that the Reflective-Endorsement view provides sufficient conditions for moral responsibility that escape the threat from the mad neuroscientist.

The Pervasiveness of Luck

Many things beyond our control affect us for good or ill, regardless of what we deserve. These things may be called ‘luck’ (which suggests randomness and purposelessness) or ‘fate’ or ‘Providence’ (which suggests inevitability and purposefulness).²⁷ I want to hold in abeyance questions about purpose or meaning, and just focus on the bare phenomenon of pervasive arbitrariness: goods and ills—including those that produce our characters and circumstances—are typically not apportioned according to desert. I’ll use words like ‘luck’ and ‘arbitrariness’ for this phenomenon (considered apart from any explanation of it).

Does the pervasiveness of luck preclude moral responsibility? I fully agree with hard determinists about the pervasiveness of arbitrariness—if the sun had not been blazing on Meursault in Camus’s L’Étranger or if he had been in a different mood, or if the Arab on the beach that day had stayed home, Meursault would not have murdered the Arab. Yes, I agree that luck is everywhere. But life—with arbitrary differences in natural endowment and environmental advantages—is unfair, from anyone’s point of view. Luck “goes all the way through...[E]verything ends up being ultimately a matter of luck,” insists Saul Smilansky. (Smilanski 2003: 275.) I agree. Amazingly, Smilansky, a hard determinist, writes as if this were news to compatibilists. He thinks that compatibilists fail to acknowledge the pervasiveness of luck, and that that failure

makes compatibilism—though necessary for treating people with dignity—shallow. But compatibilists are as likely as anyone else to be aware of the “Lucky Gene Club”, whose members are those born without crippling defects into loving and financially secure families. It does not take much empathy or insight to realize that even your ability to put to good use your lucky genes is itself a matter of luck. Star athletes who are not even especially reflective routinely refer to their talent as a gift and their opportunities for proper training as good fortune. Arbitrariness is not hidden from the view of compatibilists or of anyone else; it smacks you in the face.²⁸

Acknowledgment of luck and arbitrariness should not lead us to incompatibilism. We can (and should) try to ameliorate gross arbitrary disparities, but we cannot eliminate them altogether; nor, in the face of arbitrariness, should we give up morality.²⁹ Morality requires that there be a morally significant difference (which, incidentally, also entails a first-person perspective) between “I’m sorry that I did it” and “I’m sorry that it happened.” We need to invoke moral responsibility to make sense of such a difference. So, the pervasiveness of luck should not induce us to abandon moral responsibility. Indeed, the view that moral responsibility requires sovereignty over luck itself is a libertarian notion that should not survive the refutation of libertarianism.

Rather, the ubiquity of luck should lead us—compatibilists and incompatibilists alike—to temper judgments of moral responsibility with pity (e.g., pity for the child molester who was himself molested as a child). But luck should not drive us away from moral judgments altogether.³⁰ Recognizing the unfairness of luck, it is a mistake to infer (as Smilansky does) that it is “unfair to blame a person for something not ultimately under her control.”³¹ The unfairness of luck does not transfer to unfairness of everything that has luck somewhere in the causal

background. To give up moral responsibility in the face of the defeat of libertarianism would be to deny that we are moral beings. But it is as indubitable that we are moral beings as it is that nothing is under our ultimate control.

Religious thinkers have seen clearly the two-fold situation—that nothing is under our ultimate control, and yet we are moral beings.³² They use the term ‘God’s will’ or ‘particular Providence’ to apply to exactly the same phenomena that Smilansky and I have used ‘luck’ to refer to: the uneven distribution of goods and ills that are beyond our ultimate control. Augustine, Calvin and Jonathan Edwards³³ are all what we now call compatibilists; none denied that we are morally responsible for what we do willingly—even though they insisted that we have no ultimate control. According to such thinkers, God’s will is ultimately arbitrary in the older sense that what happens is what God rules, what He arbitrates, and God does not arbitrate on the basis of any principle (at least not on the basis of any principle knowable by us). So, the effects of God’s will are arbitrary in exactly the way that luck is.

In sum, everyone should recognize that luck (or God’s will) is pervasive, regardless of any position on freedom and responsibility. Proponents of the Reflective-Endorsement view can reject hard determinism while acknowledging ultimate arbitrariness: Reality is tragic; we might wish that it were otherwise, but what’s shallow is such wishful thinking. Compatibilism-cum-pity is an appropriate and profound response to the way things really are.³⁴

Conclusion

Human freedom is an ability. It is the unique ability, made possible by a first-person perspective, to reflect on and evaluate our desires and to choose one course of action over another. Only those with such an ability can be morally

responsible for what they do. After arguing that libertarian accounts of freedom are untenable, I have offered the Reflective-Endorsement view as compatibilist conditions sufficient for moral responsibility. The importance of moral responsibility—in our lives as moral beings, in our social practices—makes compatibilism desirable. The availability of sufficient conditions for moral responsibility makes compatibilism feasible. So, I urge acceptance of the Reflective-Endorsement view.³⁵

References

- Baker, Lynne Rudder (2004) “On Being One’s Own Person,” in Reasons of One’s Own, Maureen Sie, Bert van Den Brink, Marc Slors, eds. Ashgate Limited Publishing: 129-149.
- _____ (2003a) “Why Christians Should Not Be Libertarians: An Augustinian Challenge,” Faith and Philosophy 20: 460-478.
- _____ (2003b) “First-Person Knowledge,” The Nature and Limits of Human Understanding: The Gifford Lectures, University of Glasgow, 2001, Anthony J. Sanford, ed., London: T.&T. Clark, 165-184.
- _____ (2003c) “The Difference that Self-Consciousness Makes,” On Human Persons, Klaus Petrus, ed. (Frankfurt/London: Ontos-Verlag, 23-40.
- _____ (2000) Persons and Bodies: A Constitution View, Cambridge: Cambridge University Press.
- _____ (1998) “The First-Person Perspective: A Test for Naturalism,” American Philosophical Quarterly 35 (1998): 327-348.

- _____ (1981) "Why Computers Can't Act," American Philosophical Quarterly 18 (1981): 157-163.
- Berofsky, Bernard (2002) "Ifs, Cans, and Free Will: The Issues," The Oxford Handbook of Free Will, Robert Kane, ed., New York: Oxford University Press, 181-201.
- Chisholm, Roderick (1966) "Freedom and Action," Freedom and Determinism, Keith Lehrer, ec. (New York: Random House, 11-44.
- Clarke, Randolph (1993) "Toward a Credible Agent-Causal Account of Free Will," Noûs 27: 191-203. Reprinted in Agents, Causes, and Events: Essays on Indeterminism and Free Will, Timothy O'Connor, ed. (Oxford: Oxford University Press, 1995): 201-218. Page references in text to the latter.
- Compton, John J. (2001) "The Persistence of the Problem of Freedom," The Review of Metaphysics 55: 95-115.
- Double, Richard (1991) The Non-Reality of Free Will Oxford: Oxford University Press.
- Ekstrom, Laura Waddell (2002) "Libertarianism and Frankfurt-style Cases," The Oxford Handbook of Free Will, Robert Kane, ed. New York: Oxford University Press, 309-322.
- Fischer, John Martin (1994) The Metaphysics of Free Will Oxford: Blackwell Publishing Company.
- _____ (2002) "Frankfurt-type Examples and Semi-Compatibilism," The Oxford Handbook of Free Will, Robert Kane, ed., New York: Oxford University Press, 281-308.

- _____ and Mark Ravizza (1998) Responsibility and Control: A Theory of Moral Responsibility, Cambridge: Cambridge University Press.
- Frankfurt, Harry (1969) “Alternate Possibilities and Moral Responsibility,” Journal of Philosophy 66: 829-839.
- _____ (1971) “Freedom of the Will and the Concept of a Person,” Journal of Philosophy 68 5-20.
- Ginet, Carl (1990) On Action (Cambridge: Cambridge University Press).
- Goetz, Stewart (1997) “Libertarian Choice,” *Faith and Philosophy* 14: 195-211.
- Haji, Ishtiyaque (1998) Moral Appraisability, New York: Oxford University Press.
- Hunt, David (2000) “Moral Responsibility and Unavoidable Action,” Philosophical Studies 97: 195-227.
- Kane, Robert (1989) “Two Kinds of Incompatibilism,” Philosophy and Phenomenological Research 50: 219-254. Reprinted in Agents, Causes, and Events: Essays on Indeterminism and Free Will, Timothy O’Connor, ed. (Oxford: Oxford University Press, 1995): 115-150. Page references in the text are to the latter.
- _____ (1998) The Significance of Free Will, New York: Oxford University Press.
- _____, ed. (2002) The Oxford Handbook of Free Will, New York: Oxford University Press.
- Libet, Benjamin (2002) “Do We Have Free Will?” in The Oxford Handbook of Free Will, Robert, Kane, ed. (Oxford: Oxford University Press): 551-564

- McPherson, James M. (2001) "Southern Comfort," April 12, 2001, New York Review of Books.
- O'Connor, Timothy (1995) "Agent Causation," Agents, Causes, and Events: Essays on Indeterminism and Free Will, Timothy O'Connor, ed. Oxford: Oxford University Press, 173-200.
- _____, (2000) Persons and Causes: The Metaphysics of Free Will, Oxford: Oxford University Press.
- Pereboom, Derk (2001) Living Without Free Will, Cambridge: Cambridge University Press.
- Perry, John (1979) "The Problem of the Essential Indexical," Noûs 13: 3-22.
- Plantinga, Alvin (1974) The Nature of Necessity, Oxford: Clarendon Press.
- Rubin, Elizabeth (2002) [No title] New York Times Magazine, June 30, 2002, 6.
- Smilanski, Saul (2003) "Compatibilism: The Argument from Shallowness," Philosophical Studies 115: 257-282.
- Strawson, Galen (2000) "The Unhelpfulness of Indeterminism," Philosophy and Phenomenological Research 60: 149-156.
- Stump, Eleonore (1999) "Alternative Possibilities and Moral Responsibility: The Flicker of Freedom," Journal of Ethics 3: 299-324.
- _____(2001) "Augustine on Free Will," The Cambridge Companion to Augustine, Eleonore Stump and Norman Kretzmann, eds., Cambridge: Cambridge University Press, 124-147.

Van Inwagen, Peter (1975) "The Incompatibility of Free Will and Determinism," Philosophical Studies 27: 185-199. Reprinted in Free Will, Derk Pereboom, ed., Indianapolis: Hackett Publishing Company, 1997: 184-198. [Page references in the text are to the latter.]

_____ (1978) "Ability and Responsibility," Philosophical Review 87: 201-224.

_____ (2000) "Free Will Remains a Mystery," Philosophical Perspectives 14, Action and Freedom, James E. Tomberlin, ed., Oxford: Blackwell Publishers, 1-19. Reprinted in The Oxford Handbook of Free Will, Robert Kane, ed., New York: Oxford University Press, 158-177.

Walter, Henrik (2001) Neurophilosophy of Free Will: From Libertarian Illusions to a Concept of Natural Autonomy, Cynthia Klohr, trans., Cambridge MA: MIT/Bradford.

Widerker, David (2002) "Responsibility and Frankfurt-type Examples," The Oxford Handbook of Free Will, Robert Kane, ed. (New York: Oxford University Press, 323-334).

_____ (2002) "Farewell to the Direct Argument," Journal of Philosophy 99: 316-324.

Zagzebski, Linda (2000) "Does Libertarian Freedom Require Alternate Possibilities?" Philosophical Perspectives 14: 231-248.

Notes

¹ For an important argument—one that spawned an industry of responses—against the Principle of Alternative Possibilities, see Frankfurt 1969. I believe that Pereboom’s “new Frankfurt-style scenario” cinches the case against the Principle of Alternative Possibilities. Pereboom 2001: 18-28. Nevertheless, I admit that the Principle of Alternative Possibilities remains controversial—though I do not think that it should.

² In any case, David Widerker argues that the notion of avoidability remains central to the incompatibilist’s claim that determinism excludes moral responsibility. Widerker 2002.

³ Randolph Clarke calls the condition in (ii) the “condition of production.” Clarke 1993: 203.

⁴ Chisholm 1966: 23. Emphasis mine. Timothy O’Connor, another agent-causalist mentions that some have used the term ‘unmoved movers,’ but he prefers ‘not wholly moved movers.’ (O’Connor 1995: 174). O’Connor wants to allow causal relations between an agent’s reason and resulting behavior. On O’Connor’s view, agents have a property that can “make possible the direct, purposive bringing about of an effect by the agent who bears it.” (p. 177) Emphasis his.

⁵ E.g, Linda Zagzebski, an incompatibilist, says: “The presence of alternate possibilities may be a reliable sign of the presence of the agency needed for responsibility, but it is not necessary for it.” Zagzebski 2000: 245. Some philosophers agree that Frankfurt-style examples cast doubt on the principle, but they separate blameworthiness from wrongness. E.g., Haji 1998 argues that if determinism is true, one may be blameworthy, but there are no wrong actions.

⁶ Mississippi’s secession commissioner to Maryland insisted that “slavery was ordained by God and sanctioned by humanity.” McPherson, 2001: 30.

⁷ Unlike Kant, I do not think that it is always irrational to be immoral.

⁸ For theological arguments against the libertarian conception of freedom, see Baker 2003a.

⁹ These labels are Pereboom's. Kane calls the two types 'Agent Cause Theories' and 'Teleological Intelligibility Theories.' Kane 1995. There is another brand that might be called 'agent-noncausal libertarianism,' according to which "a choice is essentially an uncaused purposeful mental action." On this view, agents have "ultimate and irreducible" mental powers, and "[t]o choose is to exercise the mental power to choose." Goetz 1997: 196. Also see Ginet 1990. Since on these accounts, agents are said to have ultimate mental powers to make uncaused choices, I think that it's fair to say that noncausal libertarian accounts also conceive of free agents as ultimate originators.

¹⁰ Noncausal libertarians would not say that the agent caused the choice, but rather that the choice was the product of the agent's uncaused exercise of the power to choose.

¹¹ Although this point has been made numerous times, I believe that Pereboom's careful arguments against event-causal libertarianism are decisive. For more detailed arguments against all varieties of libertarianism, see Pereboom 2001.

¹² There are extremely complex methodological issues here that, unfortunately, I cannot pursue in this work.

¹³ The reasons to reject agent-causal libertarianism apply equally to agent-noncausal libertarianism.

¹⁴ In light of the *prima facie* implausibility of neural gaps, it would take us off the point to discuss the absence of evidence for them in detail. For some recent results in neuroscience, see Libet 2002 and Walter 2001. For detailed empirical objections to agent-causal libertarianism, see Pereboom 2001: Ch. 3.

¹⁵ Or at least wait for neuroscientific evidence of the relevant gaps. In general, I do not think that philosophy must simply defer to science. However, when a philosophical thesis has empirical implications (as agent-causation does), and the empirical implications are *prima facie* implausible (as relevant neural gaps are), then I think that the absence of scientific evidence in favor of the empirical implications justifies rejection of the philosophical thesis. Nevertheless, if there were empirical evidence of relevant neural gaps, I would reconsider my opinion.

¹⁶ On my view, the first-person perspective is not just a pragmatic feature of language that can be understood just in terms of indexicals. (Baker 2000). It cannot be assimilated to Perry's important work on the essential indexical. (Perry 1979.)

¹⁷ I have discussed the first-person perspective at length elsewhere. In addition to Baker 2000, see Baker 2003b, Baker 2003c, and Baker 2004.

¹⁸ Derk Pereboom does conclude that no one is morally responsible for anything.

¹⁹ A full account of moral responsibility would provide necessary as well as sufficient conditions. However, since the charge against compatibilism has been that compatibilist conditions are not sufficient for moral responsibility, the Reflective-Endorsement account is a contribution to the debate—even though it does not provide necessary conditions for free decisions or actions. As I mentioned, freedom is the default for intentional action. Freedom is defeasible with respect to intentional action in that the burden of proof is on anyone who denies freedom to an intentional action.

²⁰ Frankfurt himself holds that freedom of the will, although compatible with determinism, is not required for moral responsibility. What is at stake in this discussion, however, are sufficient conditions for moral responsibility.

²¹ For a general account of the first-person perspective, see Baker 2000. For an application of the account to the issue of moral agency, see Baker 2004.

²² More precisely, the sentence in the text abbreviates this:

(a) S has the capacity to consider the sources of her* desires.

(b) If (i) S had known that her* wanting to will X had causal antecedents that traced back to factors beyond her* control, and

(ii) S had known of the causal antecedents that traced back to factors beyond her* control that they were in the causal history of her* wanting to will X and that they were beyond her* control, then:

S still would have: willed X and wanted to will X and willed X because she* wanted to will X.

²³ Depending on how the choice or action X is instantiated, condition (i) may require that S have a first-person perspective as well. E.g., S's willing that she* be a suicide-bomber would presuppose a first-person perspective.

²⁴ Frankfurt 1971: 19. I changed the pronouns to feminine in order to make the sentence grammatically parallel to (i) – (iv).

²⁵ The conditions of the Reflective Endorsement view may be regarded as ways to accommodate Aristotelian defeating conditions of voluntariness: ignorance, compulsion, reduced capacity and the like. One reason not to focus on Aristotle in the debate over freedom is that his view has been claimed by both compatibilists and libertarians.

²⁶ Since (RE) is only a sufficient condition, there may be cases in which an agent who does not satisfy (i) – (iv) is nevertheless morally responsible.

²⁷ I do not intend to disagree here with thinkers like Augustine, Boethius, and Aquinas, who take luck and Providence to be different categories. However, for purposes here, luck and Providence can be treated together inasmuch as they both concern matters beyond an agent's control that are distributed in a way disproportionate to desert.

²⁸ Smilansky advocates combining compatibilism (at the level of social institutions) with hard determinism (at the ultimate level). I am skeptical of his use of "levels". And the compatibilist can, and should, recognize luck as well as the hard determinist.

²⁹ Pereboom argues that many of the purposes served by morality can still be served without moral responsibility. Removing moral responsibility seems to me to eviscerate morality.

³⁰ Nor should luck drive us to the pretense of making moral judgments that we do not really endorse for the sake of a decent society. This sort of pretense, associated with certain political philosophers, I find appalling.

³¹ If, as I am confident, this intuition is mistaken, then hard determinism is irrelevant to the fact of ultimate arbitrariness.. Smilansky 2003: 268. (Emphasis his.)

³² In the Christian tradition, Augustine, Calvin and Jonathan Edwards come to mind in this regard. I suspect that there are comparable figures in the Jewish and Muslim traditions as well.

³³ I would also include Aristotle and Thomas Aquinas, but since libertarians also claim them, their inclusion would be more controversial.

³⁴ I think that a person's bad luck is a good reason to mitigate punishment without supposing that the person lacks moral responsibility. So, with respect to punishment, I would fall between Jonathan Edwards, who would not count bad luck as a mitigating factor in the context of punishment, and Derk Pereboom, who holds that, because of the pervasiveness of luck, no one ever deserves any punishment at all.

³⁵ I benefited from comments on several versions of this paper by Derk Pereboom, Peter van Inwagen, Katherine Sonderegger, Hilary Kornblith, Gareth B. Matthews, and Maureen Sie. Also, thanks are due to Ishtiyaque Haji and Stefaan E. Cuypers, who commented on an earlier version of this paper at Erasmus University in Rotterdam.

University of Massachusetts Amherst