

---

## Free Will in Context: A Contemporary Philosophical Perspective

Patrick Grim, B.Phil., Ph.D.\*

---

**Philosophical work on free will is inevitably framed by the problem of free will and determinism. This paper offers an overview of the current state of the philosophical art. Early sections focus on quantum indeterminism, an outline of the most influential logical argument for incompatibilism between free will and determinism, and telling problems that face incompatibilism. A major portion of the paper focuses on the compatibilist alternative, favored by many working philosophers. The conditional account of free will offered by classical compatibilism can be shown to be inadequate. A number of compatibilist options remain open, however, and seem promising for future research. These include 'hierarchical' or 'mesh' accounts of free will, normative perspectives and an approach to free will in terms of an emphasis on context. Final sections draw out the implications of contemporary compatibilism for the brain and behavioral sciences and for the law. Copyright © 2007 John Wiley & Sons, Ltd.**

Philosophical work on free will, contemporary as well as historical, is inevitably framed by the problem of free will and determinism. One of my goals in what follows is to give a feel for the main lines of this debate in philosophy today.

I will also be outlining a particular perspective on free will. Many working philosophers consider themselves compatibilists; the perspective outlined, building on a number of arguments in the recent literature, is a contemporary form of such a view. It cannot, however, claim to be *the* contemporary philosophical perspective. There is no such thing. Against a background of the perennial problem of free will and determinism, through ongoing argument and debate, philosophers continue to try to work toward an understanding of precisely what it means for an action to be free.

---

\*Correspondence to: Patrick Grim, B. Phil., Ph.D., Group for Logic and Formal Semantics, Department of Philosophy, State University of New York at Stony Brook, Stony Brook, NY 11794-3750, U.S.A.  
E-mail: pgrim@notes.cc.sunysb.edu

## THE DILEMMA OF DETERMINISM

The problem of free will and determinism has much of the impact that it does because it is *not* a technical problem; it is a question that almost everyone raises in his or her own thinking at one point or another. Let us start, then, by phrasing the problem as simply as possible.

We think of ourselves as facing alternative courses of possible action, deliberating and making choices as to what we shall do. With our choices comes responsibility; our choices are morally right or morally wrong, and we are praiseworthy or blameworthy for those choices. A conception of free will is essential to our conceptions of ourselves, is fundamental to our ethical view of the world, and is central to the system of law, adjudication, and punishment that inherits much of the structure of our moral views.

However, we also think of the physical universe as governed by natural law. Why does a particular event occur? Because of previous events and because the laws of nature are what they are. The history of the universe is a complex chain or net of events in which later events are the natural consequence of earlier events in accordance with natural law.<sup>1</sup>

Both of these conceptions are part of our everyday thinking. Indeed it might be impossible to live the lives we do without both (a) a notion of a comprehensible universe governed by natural law and (b) a conception of free choice and responsibility. And yet these two conceptions appear to collide quite directly in the philosophical problem of free will and determinism.

We are creatures in the physical universe. Our actions are therefore events in the history of the universe, and so are produced by earlier events in accord with natural law. Given previous events and the laws of nature, it appears, the events that are our actions were therefore inevitable. So what sense does it make to say we could do otherwise, that we have alternatives, or that we can make genuine choices? If the things we do are simply the result of natural forces, how can we be held morally responsible for them? And if we cannot be held morally responsible, what possible justification can there be for a system of law constructed on conceptions of blame and punishment?

A number of philosophers see the problem of free will and determinism as reflecting an irresolvable clash between two different perspectives we take regarding ourselves and our actions—two incompatible perspectives, neither of which we could live without (Blackburn, 1999; Bok, 1998; Nagel, 1986). One perspective is a subjective or personal view—the view we have to take of ourselves as acting agents, with alternatives before us and choices to make. The other is the objective or impersonal perspective—what Thomas Nagel calls “the view from nowhere”—in which we see ourselves as parts of the natural world and our brain states and actions as part of an impersonal stream of physical events.

The problem of free will and determinism has a long and fascinating history. But it also has an active present: Sophisticated philosophical work on the problem has flourished over the last 50 years, without final resolution but with important bits of

---

<sup>1</sup>The role of natural law in the ordinary phrasing of the problem is one of physical necessity: that earlier events necessitate later events in accordance with natural law. Statistical laws are given their due at a later point, but pure regularity views are beyond the scope of the present discussion.

progress, with agreement among many on at least a general line of approach, and with new ideas that show significant promise.

The following section touches briefly on the physics of determinism. The next section lays out the logical argument for incompatibilism between free will and determinism that continues to command the most attention, but crucial problems facing incompatibilism are outlined in the fourth section. The fifth to 10th sections concentrate on the compatibilist alternative—its classical form and classical difficulties, the influential Frankfurt account, and two lines of compatibilist thought that appear promising for further work. One of these is an emphasis on normativity, clear in a range of current work (Fischer and Ravizza, 1998; Wallace, 1994; Watson, 1996; Wolf, 1990). The other, which focuses on the crucial contextuality of attributions of free will, has been developed only recently. The 11th and 12th sections follow the implications of contemporary compatibilism for brain and behavioral sciences and for the law.

## WHY PHYSICS DOES NOT SETTLE THE ISSUE

As a scientific theory, determinism has been dead since the 1930s. Why, then, can the problem of free will and determinism not be safely laid to rest?

The basic issue is often introduced, as we have introduced it above, by evoking a universe that operates in terms of previous events and deterministic natural laws. That was Newton's universe, but it is no longer ours. The lesson of quantum mechanics, on standard interpretations, is that fundamental physical law is statistical rather than deterministic. Natural law dictates that a certain percentage of uranium atoms will decay over a given time, for example—the familiar half-life of uranium—but it does not dictate that a particular uranium atom will decay at a particular time. Quantum mechanics is perhaps the best confirmed physical theory in the history of science, and it appears to be a direct consequence of the theory that some events happen without a cause. This is not merely a matter of our ignorance, it should be noted, nor of an explanation that quantum mechanics simply fails to provide. One of the fascinations of quantum mechanics is that it seems to entail that there are events for which there *can* be no cause.

Because of the evidence for quantum mechanics, therefore, we have evidence of the highest order that the universe simply does not function in the way envisaged by the classical determinist—as a stream of events, each of which is an inevitable consequence of events earlier in the stream. Quantum events occur with predictable regularity but without causal precedents, interrupting that stream.

Although quantum indeterminacy effectively sinks classical determinism, however, it is far from clear that it leaves any room for free will. Speculative roles for quantum mechanics in understanding brain function have been proposed by Stapp (1999) and Penrose (1989, 1994, 1997), with predecessors in Eccles and Popper (1977) and Eccles (1994).<sup>2</sup> The general verdict in the philosophical community, however, is that quantum indeterminacy is simply irrelevant to the basic conceptual tension at issue. Why?

---

<sup>2</sup>Non-algorithmic mathematical thought and consciousness take central stage in these accounts; free will tends to come in from the wings.

There is, first of all, a technical issue of scale. Although indeterministic on the micro scale, quantum mechanics results in a fairly traditional determinism on the macro scale—the scale of neurons, brains, and people.<sup>3</sup>

Even were this not the case, however, the indeterminacy of quantum physics would be irrelevant to the basic conceptual problem at issue. The speculation that appears in Penrose and others is that my decisions might be the results not of previous events but of indeterministic quantum phenomena, but that simply replaces one thing over which I have no control—deterministic events in a deterministic universe—with another: random quantum events in my brain. On such an approach I would no longer be a captive of a deterministic universe, but I would remain the victim of random quantum events. How could I be held responsible for events that are the result of pure chance? Despite a full incorporation of quantum indeterminacy, such a proposal still leaves no apparent place for genuinely free choices or for moral responsibility (see Dennett, 1984, 2003; Honderich, 1993).

All crucial problems remain, then, if we redefine determinism for present purposes as the claim that all events are the necessary consequence of previous events in accord with the laws of nature, or of previous events and quantum randomness in accord with the laws of nature.

## THE INCOMPATIBILIST ARGUMENT

If determinism is true, with or without quantum randomness, there can be no free will. This is the incompatibilist claim—that determinism and free will are incompatible. The most influential incompatibilist argument over the past 35 years has been a formalization in modal logic due to Peter van Inwagen (1983, 2002).

One should not let the symbolism be intimidating here: The idea is simply to make everything explicit. The core ideas actually lie very close to the formal surface.

The argument uses a concept of necessity, central to modal logic, represented by  $\Box$ . Contingent truths are propositions that are in fact true but could have been otherwise—like the proposition that 33% of institutionalized schizophrenics have widened sulci, for example, or that John Roberts was nominated as Chief Justice. These are true in one “possible world”—one way things could have come out (and did), but are not true in other “possible worlds”—other ways that things could have come out. Necessary truths, in contrast, are true propositions that could *not* have been otherwise. Mathematical and logical truths are prime examples, though there may be others. No matter how the physical universe came out,  $2 + 2$  would be 4 and no contradiction could be true; necessary truths are thought of as true “in all possible worlds.” In the formalization of the argument, “ $\Box p$ ” represents the claim that proposition  $p$  is a necessary truth.<sup>4</sup>

The argument also uses a symbol “N”, which is van Inwagen’s own.  $Np$  is meant to represent “ $p$ , and no one has or ever has had any choice about whether  $p$ .” Where  $p$

<sup>3</sup>Penrose (1994), following Fröhlich (1975), argues for the possibility of cellular structures that could scale up micro-level quantum phenomena to a macro level, but the jury is still out as to whether this is possible.

<sup>4</sup>A good introduction to the philosophical notion of “possible worlds” is the work of Lewis (1973). For an introduction to modal logic see Chellas (1980).

represents the proposition that rust is an oxide, for example,  $Np$  should be read as “Rust is an oxide, and no one has or ever has had any choice about whether rust is an oxide.”

Two rules of inference are used in the argument, phrased in terms of the symbol “ $p \vdash q$ ” for “ $p$  entails  $q$ ” or “ $q$  is deducible from  $p$ .” The first rule of inference states that from the fact that a proposition is necessarily true one can deduce that no one has or ever has had any choice regarding its truth:

$$\text{Rule } \alpha \quad \Box p \vdash Nq$$

The second rule of inference states that from the proposition that no one has any choice about  $p$ , together with the proposition that no one has any choice about whether if  $p$  then  $q$ , one can deduce that no one has any choice about  $q$ :

$$\text{Rule } \beta \quad Np, N(p \supset q) \vdash Nq$$

Here “ $\supset$ ” is the “if... then” of material implication, familiar from truth tables.

Van Inwagen formulates the crucial deterministic premise as follows: We let  $L$  represent an immense conjunction of all the natural laws of the universe. Let  $P_0$  represent the state of the world at some time in the remote past, and let  $P$  represent any true proposition about the present—the proposition that I take the bus rather than drive to work today, for example. The deterministic thesis is the first step of the argument:

$$1. \quad \Box((P_0 \ \& \ L) \supset P) \quad \text{Premise}$$

From here the argument proceeds with justifications at each step:

- |                                      |                                     |
|--------------------------------------|-------------------------------------|
| 2. $\Box(P_0 \supset (L \supset P))$ | A standard logical transform from 1 |
| 3. $N(P_0 \supset (L \supset P))$    | From 2 and rule $\alpha$            |
| 4. $NP_0$                            | Premise                             |
| 5. $N(L \supset P)$                  | From 3 and 4 by rule $\beta$        |
| 6. $NL$                              | Premise                             |
| 7. $NP$                              | From 5 and 6 by rule $\beta$        |

The premises are essentially these: That we have no choice about the state of the world at some point in the remote past (step 4), we have no choice about the laws of nature (step 6), and that anything that happens now is a necessary consequence of those two things (step 1). If this determinist claim is true, we can deduce using rules of inference  $\alpha$  and  $\beta$  that we have no choice about anything that happens now.

## PROBLEMS FOR INCOMPATIBILISM

Van Inwagen’s formal argument has been a major piece of ammunition on the incompatibilist side of the argument. Few incompatibilists hold that determinism is in fact true, and thus that free will must be an illusion (Smilansky, 2000). Most are libertarians (not to be confused with libertarians in political theory): They hold that free will is clearly real, that it is incompatible with determinism, and thus that determinism, despite its appeal, must be false.

There are, however, two immediate and major problems that face incompatibilism. The first is that van Inwagen’s core argument is flawed.

Philosophical arguments, particularly when they involve counter-examples, should perhaps come with a warning. Philosophers seem to delight in trivial or even silly examples. What is at issue even in these silly examples, however, is the attempt to find universal claims regarding the logical structure and relationships of our concepts—in this case, important concepts of free will, necessity, and determinism. A universal claim is falsified by a single contrary instance—hence the importance of counter-examples. In the attempt to evaluate conceptual claims, moreover, it is often a good idea to test the limiting cases or the polar extremes—hence the extravagant hypotheticals of many philosophical examples. However silly or apparently trivial, the philosopher's counter-example is designed to make an important logical point.

The van Inwagen argument relies crucially on  $\beta$  as a rule of inference: That if no one has a choice about whether  $p$  occurs, and no one has a choice about whether if  $p$  then  $q$ , then no one has a choice about whether  $q$ . Despite its initial plausibility, this rule of inference turns out to be invalid.

If  $\beta$  were valid, using standard rules of logic, this would be valid as well:

$$Np, Nq \vdash N(p \ \& \ q)$$

If I have no control over  $p$ , and no control over  $q$ , I have no control over whether  $p$  and  $q$  occur together. A simple trivial example, however, is enough to show that this principle, derivable from  $\beta$ , simply does not hold (McKay & Johnson, 1996).

It was entirely in my power yesterday to flip a coin, or to choose not to. One thing that was not in my power, however, was to make the coin come up heads rather than tails in a free toss. Another thing that was not in my power was to make it come up tails rather than heads. For any particular toss, then, I had no control over this:

It did not come up heads.

or this:

It did not come up tails.

By the principle above, it would follow that I had no control over this:

It did not come up heads and it did not come up tails.

But of course I did have control over that—it was within my power to keep the coin from coming up either heads or tails simply by refusing to toss it at all. This principle, then, does not always hold true:

$$Np, Nq \vdash N(p \ \& \ q)$$

If the rule of inference  $\beta$  were valid, this principle *would* always hold true.  $\beta$  is therefore an invalid rule of inference, and so is the central argument that relies on it.

Van Inwagen recognizes the force of the counter-example (van Inwagen, 2002), concedes that this form of the argument fails, but attempts to salvage the argument in terms of a tighter specification for “N”. With that tighter specification, however, the argument loses much of its punch and much of its contact with the informal argument with which we began: “N” no longer has the intuitive simplicity of the concept “no one has any choice whether. . .”.

The history of van Inwagen's argument also offers a general philosophical lesson: that arguments that appear both clear and intuitively forceful, such as the modal

argument above, can nonetheless fail. The explicitness that comes with formalization is often useful, and is always worth a try. But there may also be tricky conceptual shifts that formalization disguises rather than reveals.

A second problem would remain for incompatibilism, at least of the libertarian stripe, even if the van Inwagen argument were sound. Libertarians hold that there is indeed free will, that determinism is incompatible with that fact, and thus that determinism must be false. On such a view it is not true that all events in the universe are determined by previous events and physical law. Free choices are exceptions, somehow real and causally efficacious but nonetheless distinct from the stream of the physical universe.

The challenge is to make such a notion of libertarian free will intelligible. In such a view, there are events that might or might not occur, given all physical laws and the entire past of the physical universe, but these sound as much a matter of pure chance as do purely random events, and with the same disheartening conclusion for the concept of free will. How can I be held responsible for a pure matter of chance, and how can I think of such events as genuinely mine?

The problem of intelligibility leads van Inwagen himself to conclude that free will is incompatible with both determinism and indeterminism, and must thus remain a mystery (van Inwagen, 2002).<sup>5</sup>

## CLASSICAL COMPATIBILISM

If defined as above so as to include quantum randomness, most contemporary philosophers hold that determinism of some form is true. Most philosophers also seem to hold, however, that we make genuinely free choices and must bear moral responsibility for these choices. The explicit conjunction of those claims is compatibilism. When we properly understand free will, the compatibilist maintains, we will see that it does not conflict with the idea of a deterministic universe. The task for the compatibilist is to make good on this central claim—in particular, to show us how free will *is* to be properly understood.

What compatibilists have traditionally offered is a conditional account of free will. To be free, on such an account, is simply to have the power and ability to act in accord with one's wishes or desires. To be unfree is to be unable to further one's ends, due to denial of opportunity, physical restraint, coercion or duress. Freedom, on the conditional account, is simply a matter of whether I do as I wish. I wanted to take the road to the left, and did so. To say that I was also free to take the road to the right is simply to say that I would have taken that road if I had wanted to.

Compatibilist accounts of freedom are clear in the work of both Hobbes in the 16th century and Hume in the 17th:

For he is free to do a thing, that may do it if he have the will to do it, and may forbear, if he have the will to forbear (Hobbes, 1654/1962, p. 240).

---

<sup>5</sup>There have been attempts to provide a more direct libertarian answer to the intelligibility problem. See especially the section on "Libertarian perspectives on free agency and free will" in the work of Kane (2002). The notion of agent causation is crucial here (Chisholm, 1976, 2003). Although space does not allow for full discussion, agent causation may also find a place within the normative compatibilist approach outlined in a later section.

By liberty, then, we can only mean a power of acting or not acting according to the determinations of the will; that is, if we chose to remain at rest, we may; if we choose to move, we also may. Now this hypothetical liberty is universally allowed to belong to everyone who is not a prisoner and in chains (Hume, 1748/1955, p. 104).

Conditional accounts reappear in John Stuart Mill in the 19th century, and in 20th century work from G. E. Moore to Donald Davidson.

If free will is understood along the lines of a conditional account, conflict between free will and determinism disappears. Yesterday, after some deliberation, I took a later plane than the one I was originally scheduled for. What determinism demands is that it was because of earlier events and the laws of nature that I took the plane I eventually did, and indeed that it was because of earlier events and the laws of nature that I deliberated and made the decision I did. Determinism is a thesis regarding the “because” of actual events. Free will, on a conditional account, is something very different: a matter of a hypothetical, or a “what if”. What free will demands, on such an account, is simply that my actions would have been different if my desires had been different; that I would have done otherwise if I had wanted to. The fact that I did one thing under the circumstances that in fact obtained need not conflict with the claim that I would have done something different under slightly different circumstances. Determinism, therefore, need not conflict with free will.

At first glance, at least, the conditional account fits nicely with our basic moral and legal intuitions regarding freedom. We work with a presumption that people act as they have chosen to act and that they can therefore be held responsible for their actions. That presumption, however, is defeasible; we retract a presumption of responsibility if we find that desires and actions do not fit the standard pattern. On the “mental” side, we do not hold a person criminally responsible if they are unable to form a genuine intention, in violation of *mens rea*, or if what is at issue should be construed as something that happened to them rather than something they did, in violation of *actus reus*.<sup>6</sup> On the “external” side, responsibility is similarly reduced in cases of duress and coercion. He could not have done what the law demanded, we say, even if he had wanted to.

It is also a point in favor of a philosophical position if does not merely contradict an opposing position but can explain why someone might be tempted to such a view. This aspect of the conditional strategy has been a constant part of Daniel Dennett’s strategy (Dennett, 1984, 1995, 2003). It can be proposed, for example, that the conceptual mistake that determinists make is to confuse causality with coercion. Coercion is freedom-undermining: If I am coerced to do something, my freedom *is* undermined, and with it my responsibility for what I do. Mere causality, on the other hand, is *not* freedom-undermining: My choice to marry is a free one despite the fact that it may be explained in terms of causal precedents and the flow of hormones. Despite some similarity and overlap, causality and coercion are essentially different concepts. The appeal of determinism, so the argument runs, relies on a subtle confusion between the two.

---

<sup>6</sup>The insanity defense, in its various incarnations from M’Naghten to the present, builds on similar convictions regarding cases in which people should not be held responsible for their actions. It is tempting to say “from M’Naghten to the present and back again”, since in Federal law the historical trail of 100 years through the Smith Criterion (1929), *Durham* (1954), *Bravner* (1972), and the Insanity Defense Reform Act of 1984 lead essentially from M’Naghten back to M’Naghten.



## PROBLEMS FOR COMPATIBILISM

Despite the broad appeal of compatibilism, a simple conditional account of freedom turns out to be inadequate. It cannot be maintained that someone acts freely, or can be held responsible for their actions, if and only if he or she would have done otherwise had he or she so chosen.

An immediate and obvious problem is posed by cases in which it is clear that the agent could not have *chosen* otherwise. In cases of addiction and psychological compulsion, an agent may not be acting freely despite the fact that he or she would have acted differently if he or she had so chosen. The problem is not that the agent's actions do not accord with his or her desires; the problem is that the desires themselves are compelled by addiction or disease. With the standard warning regarding philosophical examples, one can envisage a hypothetical case in which a mad neurophysiologist manipulates someone precisely by stimulating the person's brain in ways that manipulate desires. On a conditional account it would still be true that the person does what he or she wishes, but the person is enslaved nonetheless—a clear strike against any simple conditional account.

The requirements of a conditional account may be satisfied, therefore, although the person does not act freely and should not be held responsible for his or her actions: The conditional account fails to give us sufficient conditions for freedom and responsibility. The conditional account of free will characteristic of classic compatibilism has proven a failure. That does not mean, however, that we are forced to an incompatibilist account instead. Much of what is most interesting in contemporary philosophical work is the attempt to fill out richer and more sensitive compatibilist accounts of free will.

## DOES FREEDOM DEMAND ALTERNATIVE POSSIBILITIES?

A person can act freely, and can be held morally responsible for his or her actions, only if he or she could have done otherwise. This *principle of alternative possibilities* seems both a plausible component of our moral thinking and one of the prime points at which the determinist can attack. If ours is a classically determinist universe, with a future necessitated by past events and the laws of nature, it appears there must be only one possible future—and so no genuine alternative possibilities between which I can choose. If ours is a determinist world in the wider sense outlined above, with a future necessitated by past events, the laws of nature, and quantum randomness, its alternative possibilities are still not *mine*; there are still no genuine alternatives between which I can choose.

Perhaps the principle of alternative possibilities is not as obvious as it seems, however. If free and moral action does *not* require alternative possibilities, a determinist denial of these possibilities need not entail a denial of free will. Here Frankfurt-style examples (Frankfurt, 1969) have been the focus of a great deal of attention. Despite their apparent silliness, they carry an important logical punch:

After months of deliberation, John W. Oswald decides to shoot presidential candidate Keneagan, purchases the gun and bullets, drives to a hidden location, aims at Keneagan

and pulls the trigger. John W. Oswald is also a recent patient of the mad neurophysiologist Professor Moriarty, however, who has implanted a device in Oswald's brain. As long as Oswald maintains his desire to kill Keneagan, the device remains inert. Should Oswald change his mind, however, Moriarty's device would trigger a chain of brain states in Oswald's head that would produce the same effect. Consider the case in which Oswald maintains his desire to kill Keneagan, completes his plan, and Moriarty's device is never activated. In such a case it seems clear that Oswald has acted freely in shooting Keneagan and should be held responsible for his actions. Had he chosen not to kill Keneagan, however, the Moriarty device would have been activated with precisely the same result. Oswald should be held responsible for his actions, in other words, even though he *could not have done otherwise*.<sup>7</sup> Free and moral action, on such an argument, does not require alternative possibilities: The principle of alternative possibilities is not as obvious as it might seem.

There are less outrageous cases, of a different kind, that make the same point. Dennett (1984) uses the example of Luther's pronouncement before the Diet of Worms: "Here I stand. I can do no other." What Luther meant was that his conscience made it *impossible* for him to recant. Even if that were literally true, Dennett argues, we certainly would not exclude his actions from moral evaluation or praise. In this and other cases of character, it may be precisely the fact that someone could not have done otherwise that impresses us as a particular moral strength.<sup>8</sup> A range of other subtle and intriguing cases of responsibility without control appear in the work of Sher (2006).

## THE FRANKFURT ACCOUNT OF FREE WILL

Frankfurt uses examples of the John W. Oswald sort not only to argue against the principle of alternative possibilities but to motivate a "hierarchical" or "mesh" account of free action (Frankfurt, 1969, 1988, 1999; see also Dworkin, 1988; Bok, 1998). Hierarchical accounts are some of the most fully developed outlines of free will to date, and have interesting points of contact with one contemporary approach to consciousness.

Frankfurt distinguishes between first-order desires and second-order desires: desires to have or not to have certain first-order desires. The reflective self-evaluation involved in the formation of second-order desires is characteristic of humans as opposed to other animals, Frankfurt maintains, and is central to what we mean by free action. On his account, the first-order desires that move us to action are *free* only if they mesh with our higher-order desires: only if they accord with the desires that (on a higher level) we want to have.

<sup>7</sup>Although they have had an immense impact on the literature, it should be noted that Frankfurt examples remain an item of philosophical debate. See the work of Widerker and McKenna (2003). It is also clear that Frankfurt examples rely on casual overdetermination. There has been significant progress on causality in recent years, using notions of Bayesian nets, that may also have applications here. See especially the work of Pearl (2000) and Williamson (2005).

<sup>8</sup>Susan Wolf argues for an intriguing asymmetry thesis regarding such cases: That 'could not have done otherwise' does not disqualify a good action from responsibility, though it does disqualify a bad action. See her book (Wolf, 1990).

Hierarchical accounts employ the same central strategy as the simple conditional account: Actions may be free in this sense even though the desires behind them are themselves determined. A hierarchical account does not immediately fall victim to the example of addiction, however: The first-order actions of an addict are not free precisely because they are *not* in accord with desires he wants to have. Such an account also avoids a range of John W. Oswald cases. What makes Oswald's action free is that it is in fact in accord with his higher-order desires. On a hierarchical account it is Oswald's actual desires that are important; that Moriarty's device would have prevented him from doing otherwise simply becomes irrelevant.

In all these regards the hierarchical approach seems promising. It is also intriguing that a similar account can be offered for consciousness. According to David Rosenthal (1997, 2005), our conscious states are not those that have a certain qualitative glow but simply those that we are conscious *of*. States for which there is a higher-order state of awareness are *conscious* on Rosenthal's account, therefore, in much the same way that desires for which there is a higher-order state of desire are *free* on Frankfurt's account. This convergence of research in two different but perhaps related areas suggests possibilities worthy of further work.

Hierarchical or mesh accounts are some of the richest philosophical accounts of free will yet developed. It is not clear, however, that they are adequate either to our concepts of moral responsibility or to the challenge of determinism. Consider, for example, the case of an addict who initially qualifies as unfree on Frankfurt's outline—although his first-order desires are to take the drug, these are not the desires he wants to have. In the long run, however, the drug comes to impact his higher-order desires as well: In time he comes to *want* to have the cravings that he does. That, it would seem, is a case of an even more pervasive and invidious addiction. On Frankfurt's hierarchical account, however, we would be forced to say that the long-term addict has now become free (McKenna, 2006).

This difficulty reflects not merely a single counter-example but a pervasive problem for hierarchical accounts. If free choice is understood solely in terms of a mesh between internal levels of desire, it will not matter how these levels are produced. That seems clearly off the mark: It would appear that desire production at any level by brainwashing or the interventions of a mad neurophysiologist would still count as coercive in ways that would defeat a claim of free choice.<sup>9</sup>

## REASONS AND CAUSES: THE NORMATIVITY OF FREEDOM

What the compatibilist owes us is an account, compatible with determinism, of how free choice should be understood. A conditional account is inadequate to the task. Despite points of clear interest, accounts in the Frankfurt tradition face crucial difficulties as well. There are, however, a number of options that remain open. In this and the following section I want to sketch two strands of contemporary compatibilism worthy of further work.

Whether a choice qualifies as free is treated in much of the debate as if this were a straightforward descriptive question. The unstated assumption is that there will be

---

<sup>9</sup>Recent literature on manipulation and free choice includes the work of Pereboom (2001) and Mele (2005).

some physical or metaphysical character of actions or choices that qualifies them as free. An alternative possibility, evident in a number of contemporary attempts, is that the question of free choice is not a descriptive but rather an essentially normative one (see also Wolf, 1981; Wallace, 1994).

Consider a comparison. To judge an outcome as deserved, a decision as just, or an exchange as fair is not merely to describe the cases but to evaluate them. The normative proposal regarding free will is that the locus of concepts that are important to us with regard to choice and action, including concepts of alternative choice and the ability to do otherwise, are evaluative as well, and are evaluative to their very core. If so, the attempt to treat free will as a matter of descriptive metaphysics, as the determinist does, will simply be off the mark.

A number of those pursuing a normative compatibilism take Peter Strawson (1962) as a precursor (Wallace, 1994; Watson, 1996; Wolf, 1990). Strawson emphasizes the role of “reactive attitudes” in our ascription of moral responsibility. The complex of normative attitudes we form in response to others’ actions, as Gary Watson characterizes the view, is, “as natural and primitive in human life as friendship and animosity, sympathy and antipathy. It rests on needs and concerns that are not so much to be justified as acknowledged” (Watson, 1987, p. 259, quoted by Haji, 2002). Here an important term is “primitive”; the concepts at issue are taken to be as irreducibly basic as the primitive concepts of logic. In such a view, the concepts involved in judgments regarding free choice are not merely of importance for secondary normative judgments, but are *themselves* normative all the way down.

For one of the most fully developed and most plausible normative outlines—the “reason-responsive” outline of Fischer (1987, 1994) and Fischer and Ravizza (1998)—it is Alasdair MacIntyre (1957) rather than Peter Strawson that is the clearest precursor. MacIntyre defines free behavior as behavior that can be influenced or inhibited by relevant reasons. He is clear that the concept of reasons is an essentially normative one, not to be identified with causes. Judgments of free choice and free will are therefore essentially normative as well. In the development of the view by Fischer and Ravizza, the paradigm of free action is action that the agent would not have taken had he or she had reason not to.<sup>10</sup> Free action, on such an account, is action that reliably tracks the agent’s values or reasons for action. Fischer and Ravizza attempt to make the requirements weak enough that one can also be responsible for occasional irrational actions—what is required is that the mechanism of choice be one that generally tracks reasons, even if it is occasionally unreliable. They attempt to avoid Frankfurt-style examples through careful attention to the conditions under which an agent would have acted otherwise had he or she had reason to.

What are of most importance for future work are not the details of such an account but the general emphasis on normativity. If freedom of choice is not a metaphysical property but an evaluative status, metaphysical determinist arguments will be simply off target. The “could not have done otherwise” concept that appears in determinist arguments is a descriptive property of events. The “could have done otherwise” concept crucial to our moral judgments, in contrast, may be something very different: part of a locus of concepts to be understood not in terms of the metaphysics of causality but the normative evaluation of action.

---

<sup>10</sup>For present purposes I put aside Fischer’s distinction between guidance and regulative control.

## NEW OPTIONS FOR COMPATIBILISM: FREEDOM IN CONTEXT

There is a more recent strand in work on determinism that also offers new options for compatibilism. Just as free choice is often treated in the literature as if it were a descriptive property rather than a normative status, it is almost always treated as an absolute property of a choice—something that the choice simply has or lacks. The alternative possibility is that our notion of freedom is importantly *contextual* instead: that it is only against a particular background of assumed alternatives, or an assumed context, that a choice is spoken of as free (Feldman, 2004; Hawthorne, 2001).

Consider, for comparison, the concept “large”. Whether something is appropriately spoken of as large is a matter of comparison class: A large flea is smaller than a tiny elephant. When I am fumbling for a wrench and ask you to hand me “the large one”, moreover, the context of choices is crucial to understanding what is meant. “Large” is not an absolute term, but one tied for its sense to both context and comparison class (see Grim, 2006).

Consider the concept “flat” (Lewis, 1996; Unger, 1975). We are pleased to find a flat area in which to land the plane. We have to keep sanding the table top until it is flat. But the sense in which the landing area is flat would clearly not suffice for the table top. “Flat” is not an absolute; in different contexts, the term calls forth different standards.

At least many of our uses of “free” are contextual in the same way. Mike is free to choose whatever he wants for lunch, we say. The context, however, may be one in which Mike is an inmate choosing from a limited number of options on a prison menu. When a man has a gun to your head, you have no choice but to hand over your wallet. In a larger sense, of course, you do have a choice: You can choose to keep your wallet and have your head blown off. What examples like this seem to tell us is that neither “free choice” nor “could have chosen otherwise” are absolute properties. The core concepts are contextually sensitive, and the moral judgments that ride on them can be expected to be contextually sensitive in much the same way.

What is particularly interesting is that the contextuality of free choice, once noticed, is evident at a number of crucial points in the determinism debate. In the Frankfurt examples, it is said that John W. Oswald “could not have done otherwise” than shoot presidential candidate Keneagan—had he changed his mind, Moriarty’s device would have produced the same result. However, there is a context in which it is clear that John W. Oswald *could* have done otherwise—the imaginative context in which we consider Oswald apart from the device.

In an earlier section we outlined the counter-example by McKay and Johnson (1996) to a rule of inference crucial to van Inwagen’s formal argument for incompatibilism. Here it is worth attending to that example again.

Where N means “no one had control over whether. . .”, what McKay and Johnson argue is that this is not true:

$$Np, Nq \vdash N(p \ \& \ q)$$

They have us consider the possibility that it was entirely in my power yesterday to flip a coin or not. It was not in my power, however, to make it come up heads and not

tails. It was not in my power to make it come up tails rather than heads. I therefore had no control over this:

It did not come up heads.

nor this:

It did not come up tails.

By the principle above, it would follow that I had no control over this:

It did not come up heads and it did not come up tails.

But of course I did have control over that—I could have made it come up neither heads nor tails simply by choosing not to flip it in the first place.

Or so the argument goes. Once we become attuned to the possibility that “free choice”, “could have done otherwise”, and “have control over. . .” may be context sensitive, however, cases of this sort are no longer so clear. It is in the context of already supposing a coin flipped that it is true to say I had no control over whether it came up heads. But in a broader context, I clearly *did* have control over that. I could have prevented the coin coming up heads simply by refusing to flip it. The plausibility of the counter-example by McKay and Johnson, in other words, relies on a barely perceptible shift of context behind the concept of control.

This observation would vitiate examples such as that by McKay and Johnson, but it would neither save van Inwagen’s argument nor vindicate determinism in any sense. In van Inwagen’s argument and throughout the determinism debate, concepts of freedom, power, control and “could have done otherwise” are treated as if they are absolute properties. If they are not absolute—if they are contextually sensitive properties instead, contextually applied in normal cases of moral evaluation—the determinist argument is again off the mark.<sup>11</sup>

The possibility that concepts of choice, freedom, and control are contextually sensitive puts them in good company: Many of our concepts are of this sort, including concepts of knowledge and the possibility of error. David Lewis uses this fact to produce an interesting and plausible reply to skepticism (Lewis, 1996). There is much that we know—what penguins eat, the characteristic brain waves of REM sleep, the winner of the 2005 World Series. Our common notions of knowledge, however, presuppose appropriate contextual standards of immunity to error. What the skeptic does is to force an unrealistic standard, appropriate to philosophical discussion alone. Is it not possible that some evil demon is deceiving us, and so we *might* be wrong? Then do we *really* know?

If notions of choice, freedom, and control are contextually sensitive in a similar way, the determinist’s gambit corresponds quite closely to the skeptic’s. In normal contexts we use notions of free choice and moral responsibility that carry with them a notion of control, of alternatives, and “could have done otherwise” appropriate to context. What the determinist does is to raise the ante to a level appropriate to philosophical discussion but only to philosophical discussion. Are you not part of the natural universe, with actions caused by previous events? How then can you say you are *really* in control?

<sup>11</sup>With regard to both normative and contextualist approaches, it might be argued that the central concepts at issue are hybrids: that they include both metaphysical and normative elements, or both absolute and contextualist elements. The implications of that possibility for the debate may depend on the precise form of hybrid proposed.

On Lewis's analysis of skepticism, there is a context in which the skeptic is right, and we have no knowledge—but only an artificial context, produced by rhetorical appeal to unreal possibilities. On a contextualist account of freedom, there might be a context in which the determinist is right as well, and we have no free choice—but only in the artificial context produced by rhetorical appeal to all chains of causes and the perspective of the universe as a whole.

Here we can return to more humble comparisons. The fact that no sense can be made of “*absolutely large*,” regardless of comparison class, or “*absolutely flat*,” beyond all contextual standards, does not entail that things are not properly and usefully spoken of as large and flat in normal contexts. The skeptic's unrealistic context does not impugn genuine knowledge in normal contexts. The cosmic perspective forced by the determinist similarly fails to impugn the genuinely free choices we make in normal contexts, choices that are free in a full and real sense.

A full compatibilist account would give a full outline of what free action means, from which it would be clear that freedom of action is fully compatible with the causal flow of the physical universe. This task cannot be said to be complete—normative and contextualist approaches, like other compatibilist approaches before them, are merely outlines of what form a complete account might take. What these approaches suggest, moreover, is that it may be wrong to hope for any simple definition of free will. It is a complex of concepts regarding power, control, alternative possibilities, “could have done otherwise”, free choice and moral responsibility that is at issue, and these concepts are both intricate and subtle.

## **IMPLICATIONS FOR BRAIN AND BEHAVIORAL SCIENCES**

New strands of compatibilism, if correct, have clear implications for the scientific study of free will. The most direct implication is straightforward, simple, and strongly negative: that it is a mistake to try to find some locus of free will in the brain.

In normative, contextualist, and even some hierarchical accounts of free will, our notions of free will and moral responsibility are closely tied to forms of social evaluation in general. In any of these accounts, freedom is properly thought of not as an absolute property that some choice or brain state has or lacks, but as a category of evaluation that we use in judging and regulating our own and others' actions.

Here the concept of rationality can serve as an illuminating parallel. In order for a decision to qualify as rational, it has to mesh in certain ways with internal desires of the agent but also with external canons of logic—canons that the agent may or may not have incorporated. We should not, therefore, expect to see rationality as an inherent characteristic of a belief or brain state. It is only in context, relative to other characteristics of the agent, and moreover in terms of logical standards external to the agent, that a belief or brain state is properly classified as rational.

The lessons of a contemporary compatibilism are that much the same will be true in judging whether a choice is a free one. We must consider freedom in context, relative to desires and other characteristics of the agent, and moreover in terms of external standards of responsibility. The determinist's mistake, on such accounts, is to think that freedom will be a single property of a choice or action, absolute in the

sense of either being present or absent. The hope is that the behavioral or brain scientist will not make the same mistake.

Benjamin Libet's work on volition and willed action serves as a useful example, and has been much discussed in the philosophical as well as the scientific literature (Libet, 2002; for criticisms see O'Connor, 2005). What Libet's work seems to show is that a specific readiness potential in the brain, productive of action, begins several milliseconds before human subjects are conscious of any intention to act. One way of expressing such a view is to say that our actions are not the consequence of our conscious decisions—our 'decisions' are merely an epiphenomenon of causal connections that run directly from readiness potentials to the actions themselves.<sup>12</sup>

Libet characterizes his own work as an experimental approach to the question of whether we have free will. From the perspective of contemporary compatibilism, however, his work serves as a useful example of how *not* to pursue any such question. The picture of free will that Libet's work would challenge, if any, is a picture in which an action is free if and only if it is produced and preceded by an explicit conscious decision. This is a picture of free will that falls in the general category of absolute and descriptive characterizations, and it is a picture that is clearly off the mark. Many of the actions we regard as free and responsible are *not* preceded by conscious decision. Do you deliberate and consciously decide on every action involved in driving responsibly? If so, you do not drive responsibly, and you probably do not get anywhere.

Despite the undeniable interest of Libet's work, therefore, it is mischaracterized by both its author and others. Free will in the sense that it is important to us is more complex than any caricature of prior decision in the inner theater, and is inevitably more entwined with our modes of social action.

This negative conclusion should not be overstated, however. It is clear that there are brain and behavioral capabilities that are physically necessary for the exercise of free will, even if there is no set of descriptive capabilities conceptually necessary and sufficient for defining it. Scientific work regarding these necessary capacities will continue to be of importance for our moral and legal deliberations. What we should not expect from brain and behavioral science is any discovery of a location for responsibility or a particular brain process that *is* free will. If contemporary philosophical approaches are any guide, free will and moral responsibility are simply not this kind of thing.

## IMPLICATIONS FOR THE LAW

The implications of contemporary compatibilism for the brain and behavioral sciences are importantly negative: a warning against tempting mistakes regarding the nature of free will. The implications of contemporary compatibilism for the law, on the other hand, are almost entirely positive: that traditional attempts to handle issues of free will in the legal sphere are on precisely the right track.

The law is one thing, its rightness or wrongness another. Despite this basic principle of legal positivism, our systems of law are inevitably shaped by what we hold to be morally right. If a law is unjust, we have a clear reason to change it. If a

---

<sup>12</sup>Libet's own view is somewhat more subtle: He notes that a conscious decision may nonetheless cancel an action already in progress (Libet, 2002).



procedure is unfair, we have an ethical reason to do things differently. The fact that an agent should not be held morally responsible for his or her act is, *prima facie*, a reason to think he or she should not be held legally responsible either. Although the law is one thing and its rightness or wrongness another, the complex of legal concepts regarding responsibility is clearly built on, and is clearly answerable to, the corresponding complex of our moral concepts.

On a contemporary compatibilist framework, judgments of free choice and responsibility are subtle and socially complex, essentially normative and unavoidably contextual. That means there will be no “free will” test—a legal litmus that will tell us whether the ingredient of freedom necessary for culpability is present or absent. What we should expect instead are a range of broad generalizations that lay out the intersection of basic concepts, a fabric of defeasible presumptions regarding responsibility or the lack of it, standards as to what reasonable individuals can be expected to know, and a rich set of agreed examples. This is, in the end, what our moral knowledge regarding responsibility is like. It is also what our system of law has attempted to develop over the course of centuries.

Consider again the principle of *mens rea*—the core “mentality” requirement for criminal responsibility in common law. The *mens rea* requirement is often glossed by saying, “There is no crime without criminal intent,” but this is in fact misleading. As it appears in American law, what *mens rea* requires is (1) intention, (2) recklessness, or (3) negligence. Of the three, it is only intention that we might be tempted to identify with some internal mental state, and even this may be a temptation that should be resisted. It is clear that recklessness is not purely a matter of the internal state of an agent—it is a matter of what behavior is normatively appropriate in the circumstances, of what care should have been taken but was not. Negligence is a matter of the states of attention that do *not* appear in the agent, but should have. What *mens rea* offers, then, is not a simple characterization of free choice in terms of some internal mental state. What we see instead are precisely the strands of normativity and contextuality that contemporary compatibilism would suggest are appropriate to judgments of responsibility.

H.L.A. Hart made a similar point long ago regarding classic principles of contract law. Among the defenses with which an otherwise valid claim to contract can be met are these:

- A. Defenses which refer to the knowledge possessed by the defendant.
  - i. Fraudulent misrepresentation
  - ii. Innocent misrepresentation
  - iii. Non-disclosure of material facts
- B. Defenses which refer to what may be called the will of the defendant.
  - i. Duress.
  - ii. Undue influence.
- C. Defenses which may cover both knowledge and will.
  - i. Lunacy.
  - ii. Intoxication. (Hart, 1959).

These are also precisely the kinds of consideration that lead us to say that an individual’s action was not freely chosen, and so should not carry moral responsibility. Here again the legal tradition instantiates the web of social and

normative considerations, structured in terms of defeasible presumptions and disqualifying circumstances, that contemporary compatibilism maintains are crucial to our judgments of free will.

One might conclude at this point that the law has understood free choice and responsibility better than the behavioral sciences have. This in turn might serve as a warning to the legal profession not to expect from the sciences any easy answers regarding freedom and responsibility. If strands of contemporary compatibilism are right, answers to these questions will never be any simpler or easier than they are now.

## CONCLUSION

The philosopher inevitably sees questions of free will against the background of determinism. What I have attempted here is a brief survey of the state of the philosophical art regarding the problem of free will and determinism, including formalization of the incompatibilist argument and its difficulties, the failure of conditional accounts and classical compatibilism, and options that remain.

In the latter regard, I have emphasized contemporary normative approaches and have added a recent emphasis on the contextuality of judgments of free will. Contemporary compatibilist accounts of this sort offer warnings for the behavioral and neurosciences; the hope is that investigators in these fields will avoid the mistakes that philosophers have already made and have lived to regret. On the other hand, contemporary compatibilist accounts are very much in line with legal principles and procedures evolved over centuries. In such an approach, the ongoing legal task is very close to the philosophical task that remains, though with a practical rather than a merely theoretical purpose: to codify the complex social considerations, contextually sensitive and essentially normative, relevant to the judgments we have to make regarding free will and responsibility.

## REFERENCES

- Blackburn, S. (1999). *Think*. Oxford: Oxford University Press.
- Bok, H. (1998). *Freedom and responsibility*. Princeton, NJ: Princeton University Press.
- Chellas, B. (1980). *Modal logic: An introduction*. New York: Cambridge University Press.
- Chisholm, R. (1976). *Person and object*. LaSalle, IL: Open Court.
- Chisholm, R. (2003). Human freedom and the self. In G. Watson (Ed.), *Free will* (2nd ed., pp. 26–37). Oxford: Oxford University Press.
- Dennett, D. (1984). *Elbow room*. Cambridge, MA: MIT Press.
- Dennett, D. (1995). *Darwin's dangerous idea*. New York: Touchstone.
- Dennett, D. (2003). *Freedom evolves*. New York: Viking Penguin.
- Dworkin, G. (1988). *The theory and practice of autonomy*. Cambridge: Cambridge University Press.
- Eccles, J. (1994). *How the self controls its brain*. Berlin: Springer.
- Eccles, J. & Popper, K. (1977). *The self and its brain*. New York: Springer.
- Feldman, R. (2004). Freedom and contextualism. In J. Campbell, M. O'Rourke, & D. Shier (Eds), *Topics in contemporary philosophy Vol. II: Freedom and determinism* (pp. 255–276). Cambridge, MA: MIT Press.
- Fischer, J. M. (1987). Responsiveness and moral responsibility. In F. Schoeman (Ed.), *Responsibility, character and emotions* (pp. 81–106). Cambridge: Cambridge University Press.
- Fischer, J. M. (1994). *The metaphysics of free will: A study of control*. Oxford: Blackwell.
- Fischer, J. M., & Ravizza, M., (1998). *Responsibility and control: A theory of moral responsibility*. Cambridge: Cambridge University Press.
- Frankfurt, H. (1969). Alternate possibilities and moral responsibility. *Journal of Philosophy*, 66, 829–839.

- Frankfurt, H. (1988). *The importance of what we care about*. New York: Cambridge University Press.
- Frankfurt, H. (1999). *Necessity, volition and love*. Cambridge: Cambridge University Press.
- Fröhlich, H. (1975). The extraordinary dielectric properties of biological materials and the action of enzymes. *Proceedings of the National Academy of Sciences*, 72(11), 4211–15.
- Grim, P. (2006). The buried quantifier: An account of vagueness and the sorites. *Analysis*, 65, 95–103.
- Haji, I. (2002). Compatibilist views of freedom and responsibility. In R. Kane (Ed.), *The Oxford handbook of free will* (pp. 202–228). Oxford: Oxford University Press.
- Hart, H. L. A. (1959). The ascription of responsibility and rights. In A. Flew (Ed.), *Essays on logic and language (First Series)* (pp. 145–166). Oxford: Blackwell.
- Hawthorne, J. (2001). Freedom in context. *Philosophical Studies*, 104, 63–79.
- Hobbes, T. (1654/1962). Of liberty and necessity. In W. Molesworth (ed.), *The English works of Thomas Hobbes*, Vol. 5. London: Scientia Aalen.
- Honderich, T. (1993). *How free are you?* Oxford: Oxford University Press.
- Hume, D. (1748/1955). *An enquiry concerning human understanding*, L. Selby-Bigge (Ed.) Oxford: Clarendon.
- Kane, R. (2002). *The Oxford handbook of free will*. New York: Oxford University Press.
- Lewis, D. (1973). *Counterfactuals*. Malden, MA: Blackwell.
- Lewis, D. (1996). Elusive knowledge. *Australasian Journal of Philosophy*, 74, 549–567. Reprinted in P. Grim, K. Baynes, G. Mar (Eds.) (1998). *The philosopher's annual XIX* (pp. 139–159). Atascadero: Ridgeview.
- Libet, B. (2002). Do we have free will? In R. Kane (Ed.), *The Oxford handbook of free will* (pp. 551–564). Oxford: Oxford University Press.
- MacIntyre, A. (1957). *Determinism*. *Mind*, 66, 28–41. Reprinted in B. Berofsky (Ed.) (1966). *Free will and determinism* (pp. 240–255). New York: Harper and Row.
- McKay, T., & Johnson, D. (1996). A reconsideration of an argument against compatibilism. *Philosophical Topics*, 24, 113–122.
- McKenna, M. (2006). Compatibilism. Retrieved June 14, 2006, from Stanford Encyclopedia of Philosophy <http://plato.stanford.edu>.
- Mele, A. (2005). A critique of Pereboom's 'four-case argument' for incompatibilism. *Analysis*, 65, 75–80.
- Nagel, T. (1986). *The view from nowhere*. New York: Oxford University Press.
- O'Connor, T. (2005). Freedom with a human face. *Midwest Studies in Philosophy*, 29, 207–227.
- Pearl, J. (2000). *Causality: models, reasoning, and inference*. Cambridge: Cambridge University Press.
- Penrose, R. (1989). *The emperor's new mind*. Oxford: Oxford University Press.
- Penrose, R. (1994). *Shadows of the mind*. Oxford: Oxford University Press.
- Penrose, R. (1997). *The large, the small, and the human mind*. Cambridge: Cambridge University Press.
- Pereboom, D. (2001). *Living without free will*. Cambridge: Cambridge University Press.
- Rosenthal, D. (1997). A theory of consciousness. In N. Block, O. Flanagan, G. Güzelidere (Eds.), *The nature of consciousness: Philosophical debates* (pp. 729–754). Cambridge, MA: MIT Press.
- Rosenthal, D. (2005). *Consciousness and mind*. Oxford: Clarendon.
- Sher, G. (2006). Out of control. *Ethics*, 116, 285–301.
- Smilansky, S. (2000). *Free will and illusion*. Oxford: Clarendon.
- Stapp, H. P. (1999). Attention, intention and will in quantum physics. In B. Libet, A. Freeman, K. Sutherland (Eds.), *The volitional brain: Towards a neuroscience of free will* (pp. 143–164). Thorverton, UK: Imprint.
- Strawson, P. (1962). Freedom and resentment. *Proceedings of the British Academy*, 48, 1–25.
- Unger, P. (1975). *Ignorance*. New York: Oxford University Press.
- van Inwagen, P. (1983). *An essay on free will*. Oxford: Clarendon.
- van Inwagen, P. (2002). Free will remains a mystery. In R. Kane (Ed.), *The Oxford handbook of free will* (pp. 158–177). Oxford: Oxford University Press.
- Wallace, R. J. (1994). *Responsibility and the moral sentiments*. Cambridge, MA: Harvard University Press.
- Watson, G. (1987). Responsibility and the limits of evil: Variations on a Strawsonian theme. In F. Schoeman (Ed.), *Responsibility, character and emotions* (pp. 256–286). Cambridge: Cambridge University Press.
- Watson, G. (1996). Two faces of responsibility. *Philosophical Topics*, 24, 227–248.
- Widerker, D., & McKenna, M. (2003). *Moral responsibility and alternative possibilities*. Burlington, VT: Ashgate.
- Williamson, J. (2005). *Bayesian nets and causality: Philosophical and computational foundations*. Oxford: Oxford University Press.
- Wolf, S. (1981). The importance of free will. *Mind*, 40, 386–405.
- Wolf, S. (1990). *Freedom within reason*. Oxford: Oxford University Press.