P. F. Strawson's Free Will Naturalism

Joe Campbell
Washington State University, United States
josephc@wsu.edu

Abstract

This is an explication and defense of P. F. Strawson's naturalist theory of free will and moral responsibility. I respond to a set of criticisms of the view by free will skeptics, compatibilists, and libertarians who adopt the *core assumption*: Strawson thinks that our reactive attitudes provide the basis for a rational justification of our blaming and praising practices. My primary aim is to explain and defend Strawson's naturalism in light of criticisms based on the core assumption. Strawson's critiques of incompatibilism and free will skepticism are not intended to provide rational justifications for either compatibilism or the claim that some persons have free will. Hence, the charge that Strawson's "arguments" are faulty is misplaced. The core assumption resting behind such critiques is mistaken.

1 Introduction

This is an explication and defense of P. F. Strawson's theory of free will and moral responsibility (1962, 1985), aka *Strawson's naturalism.* I respond to criticisms leveled by free will skeptics (Smilansky 2001; Pereboom 2001, 2014; Sommers 2007; G. Strawson 2010), compatibilists (Watson 1987; Scanlon 1988; Russell 1992; Wallace 1994; McKenna 1998, 2005, 2012), and libertarians (Wiggins 2003). These criticisms adopt the *core assumption*: Strawson thinks that our reactive attitudes provide a rational justification for our blaming and

praising practices.[1] On this interpretation, Strawson gives arguments for both free willism and compatibilism. *Free willism* is the claim that some persons have free will.[2] *Compatibilism* is the thesis that free willism is consistent with *determinism*, the view that "the past *determines* a unique future" (van Inwagen 1983: 2).

An example of a *core assumption criticism* of Strawson's naturalism is Paul Russell's claim that Strawson adopts both rationalistic and naturalistic strategies to the challenges of incompatibilism and free will skepticism (1992: 289ff.). Other critics join Russell in adopting the core assumption. Saul Smilansky writes that Strawson "attempts to offer, as we have seen, a defence of our common attitudes and practices in this predicament" (2001: 79). Similarly, Galen Strawson claims that, according to Strawson, "the concepts of, and belief in, moral responsibility and freedom are *in some sense* shown to have application, and to be justified, respectively, by the mere fact of the existence of our commitment to the personal-reactive attitudes and practices" (2010: 74). Lastly, Derk Pereboom writes: "To secure his case for compatibilism, Strawson, like Hume, argues first for a psychological thesis, that the reactive attitudes *cannot* be affected by a belief in universal determinism, and then for a normative thesis, that they *should not* be affected by this belief" (2001: 92).

Russell is correct that the rational and naturalistic strategies are in conflict (1992: 291), but on my interpretation it is wrong to attribute both strategies to Strawson precisely because his naturalism precludes the adoption of a rationalistic strategy. Strawson's comments against *incompatibilism* and *free will skepticism*—respectively, the denials of compatibilism and free willism—are not intended to provide the basis for some rational justification. Therefore, the charge that Strawson's "arguments" for compatibilism and free willism are faulty or inconclusive, raised by Russell and others, is misplaced. The core assumption behind these criticisms is incorrect.

After some preliminary comments (Section 2), we address Strawson's reconciliation project (1962: 45–46) and its relation to the core assumption (Section 3). Next, Strawson's naturalism is expanded and some opposing views are handled (Section 4). In the remaining sections, additional core assumption criticisms are considered, and I respond to some objections (Sections 5–7). We get a clearer, more consistent understanding of Strawson once we reject the core assumption and fully appreciate his own variety of naturalism about free will and moral responsibility.

---

1 A *rational justification* for a proposition *p* provides a reason to believe that *p* (see Section 4).

2 The name "free willism" was suggested to me by Brandon Warmke.

2   Preliminaries

Free will is the most fundamental freedom-relevant condition necessary for moral responsibility (see Fischer and Ravizza 1993b: 8; Mele 2006: 17; Pereboom 2014: 4). *Libertarians* hold that free willism is incompatible with determinism, free willism is true, and determinism is false. Free will skepticism is a metaphysical thesis, not an epistemological thesis. In other words, it is the stronger claim that no one has free will rather than the weaker claim that no one knows whether anyone has free will. This is important since Strawson's naturalism is intended to provide comfort in the face of various epistemological problems like skepticism about the external world, the problem of induction, and knowledge of other minds as well as worries about free will and moral responsibility.[3]

The kind of moral responsibility under consideration in this essay is *robust moral responsibility*. Tamler Sommers (2007: 322) calls it "robust (desert-entailing) moral responsibility" and Pereboom calls it "moral responsibility in the basic desert sense" (Pereboom 2014: 2). As these philosophers suggest, robust moral responsibility requires basic desert (Strawson 1962: 64; 1980: 261–262; 2011c: 149). A person is morally responsible for a wrong action if and only if she is *blameworthy*, that is, deserving of blame for the action. The free will skeptic is related to Strawson's "genuine moral skeptic," someone who believes that "the notions of moral guilt, of blame, of moral responsibility are inherently confused and that we can see this to be so if we consider the consequences either of the truth of determinism or of its falsity" (1962: 45).[4] For this reason, I often slip and slide between talk about free will and talk about moral responsibility.

This essay critiques a popular and influential view, e.g., that Strawson adopts the core assumption. *Core assumption criticisms* are based on a reading of Strawson (1962) that tends to deemphasize his later work (especially 1985). In his subsequent work, Strawson makes it clear that he is not providing a rational justification for our ordinary claims about free will and moral responsibility, since these claims cannot be grounded in that way. Strawson resists the kind

---

3   The category of metaphysical skepticism is contentious. Even Strawson (1985: 2) has concerns, but that is a topic for another occasion.

4   See Pereboom (2014: 4) for a brief overview of the varieties of free will skepticism. Many free will skeptics adopt a weaker view than Strawson's genuine moral skeptic, since not all free will skeptics reject moral obligation (see Pereboom 2014: 138–146). Thanks to Caouette for pointing this out. Further, many free will skeptics adopt the stronger thesis of *impossibilism*: the concept of *free will* is incoherent (see Vihvelin 2008). Examples of impossibilists include Smilansky (2001), Sommers (2007), and their ideological mentor: Galen Strawson (2010).

of reductive accounts of our judgments about blame and praise suggested by the core assumption.

Throughout the paper I move easily between discussions of incompatibilism and free will skepticism. Why is that?[5] Here is a thesis that all free will skeptics should endorse:

> *The anti-compatibilism thesis*
> No compatibilist account provides an adequate set of sufficient conditions for free will and thus moral responsibility.

One reason that many free will skeptics endorse this claim is that they acknowledge that persons often satisfy the conditions for free will and moral responsibility offered in compatibilist accounts. This is a large part of Pereboom's project, since he believes that such compatibilist conditions help ground deliberation, meaning, etc. Yet, on Pereboom's view, compatibilist conditions are not enough for genuine free will (see below). Compatibilist conditions are necessary but not sufficient. Bruce Waller writes: "when those who reject moral responsibility deny the existence of the requisite miraculous powers, they are not denying the existence of moderate (non miraculous) levels of rationality, self-control, and moral commitment" (2006: 87).

Of course, free will skeptics argue that no one has provided sufficient conditions for free will, neither compatibilists nor libertarians (cf. Strawson 2011c: 151). If we start with incompatibilism, then one merely needs to add that there is no satisfactory libertarian account of free will in order to establish free will skepticism. There are two ways to go from here. First, one might present the argument for free will skepticism as a constructive dilemma, with two separate premises: one for incompatibism and another against libertarianism (Campbell 2011). Second, one might combine the reasons against both kinds of free will into a single general premise or a set of premises. In the second option, the argument for free will skepticism need not include incompatibilism as an explicit premise. Hence, the view that all arguments for free will skepticism require the thesis of incompatibilism is contentious.

This is not a reason for the free will skeptic to reject the anti-compatibilist thesis. Either the argument for free will skepticism fits the form of a constructive dilemma or it provides a streamlined argument with various premises that entail the falsity of both compatibilist and libertarian theories of free will. In either case, the free will skeptic is also committed to the denial of compatibilism.

---

5 Thanks to a referee from the *International Journal for the Study of Skepticism* for raising this question.

Perhaps there is a general premise that collects together reasons for rejecting free will that have nothing to do with determinism *per se*, or perhaps there are a set of reasons that jointly rule out both compatibilist and libertarian theories of free will. It doesn't matter. For the general premise, or set of premises, of any argument for free will skepticism will entail incompatibilism. Otherwise, the argument is invalid since it keeps the compatibilist option of free will open. Free will skeptics are committed to the anti-compatibilism thesis and, thus, to incompatibilism.

Of course, one is free to define terms differently and resist the argument of the previous paragraph (see Waller 2003, Honderich 2004, Vihvelin 2008, Levy 2011). What matters most is whether Strawson would accept my argument, and given his understanding of the relevant terms and debate, he would. It is interesting that Waller's (2003) view bears a strong resemblance to Strawson's, for both attempt 'reconciliations' between compatibilism and incompatibilism. There is more to be said but this should be enough to establish a link between incompatibilism and free will skepticism.

A more thorough explanation of Strawson's naturalism is provided below but some initial points are worth noting. On my interpretation, Strawson (1985) seeks to clarify and expand on the view given in his previous work (1962). In between these essays, he contrasts his naturalism with that of W.V.O. Quine. Strawson writes:

> Each standpoint could be called a form of naturalism—one with a pronounced bias towards science, the other of a more humanistic or, as I called it, 'liberal' variety. But to those—not few—in whom the drive for completely unified philosophical explanation is strong this tolerance may well seem unacceptable. (2011e: 243)

As I show in the next and subsequent sections, Strawson (1985: Ch. 1) explicitly aligns himself with *nonreductive naturalism*, a view he takes to be shared by David Hume (1975) and Ludwig Wittgenstein (1969). It might be odd to think of Hume or Wittgenstein as naturalists but here is a framework to make it easier.

Suppose a three-fold division of responses to philosophical problems, such as the meaning of life: transcendentalism, skepticism, and naturalism. *Transcendentalists* adopt theories that require transcendence of the natural order. For instance, one might think that the meaning of life requires the existence of God or immortality, and that neither can be explained by appeal to "the facts as we know them" (Strawson 1962: 46). Similarly, the incompatibilist might contend that free will requires "deep openness," "agent causation," or "soul power" (Mele 2014: 107), and in doing so she goes beyond the natural order in an effort to explain the concept of *free will*.

Of course, empirically viable accounts of libertarianism are prevalent (Kane 1996, Balaguer 2010), but keep in mind that the purpose here is to make a plausible case for Strawson's naturalism and its connection with Hume and Wittgenstein, and these more empirically viable views of libertarianism were not available when Strawson was writing. At that time, most libertarian accounts were versions of the agent causation view, which is a kind of transcendentalism. I have more to say about naturalist accounts of libertarianism toward the end of this section.

Skeptics share with transcendentalists the view that securing the relevant philosophical token—meaning, free will, etc.—requires that it have some transcendental property but they reject the transcendentalist's claim that the property is ever manifested; some skeptics believe that the property can never be manifested. *Naturalists*, on this three-fold model, seek an understanding of the relevant philosophical concept that does not appeal to transcendental properties. As we've seen, there are both reductive and nonreductive naturalists.

This is a very rough sketch as it neglects the subtleties of the various transcendentalists, skeptics, and naturalists. I am not defending it so much as offering it to explain Strawson's project. In this light, Strawson's thesis has four main parts. First, people accept various *hinge commitments*,[6] some of which entail that we have robust moral responsibility. Second, our natural attachment to such hinge commitments renders idle both skeptical arguments against them as well as philosophical arguments in their favor. Third, acceptance of hinge commitments is nonrational. Fourth, and in conclusion, there is no need to require the "obscure and panicky metaphysics of libertarianism" (Strawson 1962: 66) in an attempt to save robust moral responsibility, for we have these commitments independent of and regardless of our commitments to determinism.

Understood in this way, Strawson's response applies to naturalistic libertarian views as well as transcendental views. The need for incompatibilism and free will skepticism is lifted once we fully grasp Strawson's naturalism. We can have

6   The expression "hinge commitment" is from Duncan Pritchard (2015). In the Hume literature there is the related expression "natural belief" (Kemp Smith 1966, Campbell 2015), and in the Wittgenstein literature "hinge proposition" is common. As I explain below, commitments are attitudes but not all attitudes are beliefs (cf. Levy 2014). Indeed, it is worth noting that Strawson, citing Hume and Wittgenstein, questions "the appropriateness of the ordinary concepts of 'belief' and 'proposition' in this connection" (1998e: 370).

robust moral responsibility without transcendental, libertarian free will. This aligns Strawson with Hume and Wittgenstein, for neither is a transcendentalist and, while both make concessions to skepticism, neither is a skeptic. Both dabbled with a rather strong reductionism, but both eventually became aware of its limitations (Campbell 2015).

3   Strawson's Reconciliation Project and the Core Assumption

In "Freedom and Resentment," Strawson attempts to reconcile two opposing views: *optimism* and *pessimism* (1962: 45–46). The optimist is a compatibilist; the pessimist an incompatibilist. Thus, what they are optimistic or pessimistic about are the prospects for free will given determinism. Strawson does not plan to reconcile incompatibilism with compatibilism, for that is impossible. Rather, Strawson hopes that the pessimist will make a "formal withdrawal in return for a vital concession" from the optimist (1962: 46).

According to pessimism, the concept of *moral responsibility* has "no application, and the practices of punishing and blaming, of expressing moral condemnation and approval, are really unjustified" given determinism (1962: 45). *Ergo*, determinism entails free will skepticism. Pessimism is incompatibilism, though Strawson aims for the libertarian, the incompatibilist who accepts free willism. Optimism is compatibilism, but Strawson's target is one who points to "the efficacy of the practices of punishment, and of moral condemnation and approval, in regulating behaviour in socially desirable ways" (1962: 46). This refers to a group of views endorsed by philosophers like Moritz Schlick (1939: Ch. 7), P. H. Nowell-Smith (1948), and J. J. C. Smart (1961). T. M. Scanlon calls these "influencability theories," for they claim that "the purpose of moral praise and blame is to influence people's behavior" (Scanlon 1988: 357).

A well-known influencability theory is the *social regulation view*, which "holds that blaming and praising judgments and acts are to be understood, and justified, as forms of social regulation" (Watson 1987: 166; see also Fischer and Ravizza 1993b: 11–12; Bennett 2008). Two key features of influencability theories are that (*i*) they provide at best pragmatic justifications of blaming and praising behavior, and (*ii*) they are closely aligned with consequentialist theories of punishment (Nowell-Smith 1948: 56). Strawson's worry about influencability theories, a worry he shares with the pessimist, is that they fail to provide robust moral responsibility. He was right to have this worry since the social regulation would not count as a version of compatibilism given the requirement of robust moral responsibility (cf. Pereboom 2014: 2).

Strawson wants the optimist to concede that influencability theories are "inadequate" and that they leave out something "vital."[7] In return, he hopes the pessimist will withdraw her metaphysical commitments, especially her belief that moral responsibility "can be secure only if, beyond the facts as we know them, there is the further fact that determinism is false" (1962: 46). Strawson envisages a theory of moral responsibility that is practical enough to satisfy the influencability theorist, allows for the robust moral responsibility that the pessimist seeks, yet avoids "panicky metaphysics." Seen in this light, Strawson's theory is a compatibilist view (G. Strawson 2010: 84; Smilansky 2001: 72; cf. Waller 2003).

Most parties to the debate accept the story up to this point. But where does Strawson go from here, and what does he take himself to be doing along the way? The majority opinion is that in telling the story as he does, Strawson attempts to provide a rational justification of our moral practices, similar to the alleged justification provided by influencability theorists. This is the core assumption.

Central to both the core assumption and Strawson's criticism of free will skepticism is his concept of the *objective attitude*. Strawson writes:

> To adopt the objective attitude to another human being is to see him, perhaps, as an object of social policy; as a subject for what, in a wide range of sense, might be called treatment; as something certainly to be taken account, perhaps precautionary account, of; to be managed or handled or cured or trained; perhaps simply to be avoided... (1962: 52)

The objective attitude is contrasted with the *participant attitude*, that is, "the range of reactive feelings and attitudes which belong to involvement or participation with others in inter-personal human relationships" (1962: 52). The participant and objective attitudes are primarily distinguished by the appropriateness of various *reactive attitudes*, emotions like "gratitude, resentment, forgiveness, love, and hurt feelings" (1962: 48).

According to the core assumption, Strawson believes that reactive attitudes serve to ground or justify our blaming and praising practices. Derk Pereboom writes:

> the basis of moral responsibility [for Strawson] is to be found in reactive attitudes such as indignation, moral resentment, guilt, and gratitude. For example, the fact that agents are typically resented for certain kinds of immoral actions is what constitutes their being blameworthy for performing

---

7   In correspondence, Michael McKenna contends that, in addition, the optimist is required to offer an alternative account that is vital and robust, in contrast to influencability theories.

them. The key idea is that justification for claims of blameworthiness and praiseworthiness ends in the system of human reactive attitudes.

<div align="center">2014: 72; see also WALLACE 1994: 8FF., WIGGINS 2003: 117–120</div>

The core assumption comes in degrees. At one end of the spectrum is the claim that Strawson adopts the *constitutive view*, which "holds that it is these reactive attitudes and practices themselves that are constitutive of responsibility" and that responsibility is, thus, "grounded in nothing more than our adopting these attitudes toward one another" (Fischer and Ravizza 1993b: 16). According to John Martin Fischer and Mark Ravizza, Strawson's view is "that our questions about the propriety of the reactive attitudes should be taken as essentially practical questions, not theoretical ones" (1993b: 16). Gary Watson writes: "Strawson's radical claim is that these 'reactive attitudes' (as he calls them) are *constitutive* of moral responsibility; to regard oneself or another as responsible just is the proneness to react to them in these kinds of ways under certain considerations" (1987: 257).

To say that the reactive attitudes are *constitutive* of moral responsibility is to say that claims about moral responsibility reduce to claims about our propensities to blame and praise. This view is open to a clear and decisive criticism, for as Fischer and Ravizza note, "there seems to be a difference between being *held* responsible and actually *being* responsible" (1993b: 18). It does not follow from the fact that a collection of people blame someone else for an action that she is indeed blameworthy, no matter how numerous the collection is. If Strawson adopts the constitutive view, he is mistaken.

At the other end of the spectrum is the claim that Strawson accepts a more moderate *holding-being view*, which weakens the connection between holding responsible and being responsible. As Michael McKenna puts it, "being morally responsible must be understood by reference to the nature of holding morally responsible" (2012: 30; see also Scanlon 1988: 358ff.). This theory sheds the semantic reductionism of the constitutive view. Not all versions of the being-holding view require the kind of epistemological reductionism associated with the core assumption but some do. What all core assumption criticisms have in common is the attempt to provide a rational justification of our practices of holding persons blameworthy or praiseworthy for their actions. The whole point behind Strawson's naturalism, as I see it, is that we should resist this kind of reductionism.[8]

---

8 Thanks to a referee from the *International Journal for the Study of Skepticism* for pressing me on this and other points in this paragraph.

4   Strawson's Naturalism, the Constitutive View, and Pragmatism

Strawson's response to the free will skeptic is part of a broader naturalism directed at skeptical worries in general. Among the forms of skepticism that Strawson considers are threats to "the existence of the external world, i.e. of physical objects or bodies; our knowledge of other minds; the justification of induction; the reality of the past" (1985: 2–3). Thus, some comments are in order before we discuss Strawson's response to the free will skeptic. Along the way we'll see how Strawson's naturalism applies to the constitutive view and to pragmatism.

There are two kinds of naturalism, according to Strawson: "strict or reductive naturalism" and "catholic or liberal naturalism" (1985: 1; 2011d: 170; 2011e: 243). He also uses the terms "hard" and "soft," respectively, to contrast these two forms of naturalism (1985: 1). Other names include, for the former, "austere" (1998a: 168) and, for the latter, "nonreductive" (1985: 41) and "humanistic" (1998a: 168). I refer to the former as "reductive naturalism" and to the latter as "Strawson's naturalism" or "nonreductive naturalism."

Reductive naturalism does not allow for the existence of anything "which is not ultimately reducible to or explicable in terms of the natural sciences" (1998a: 168). Strawson writes: "An exponent of some subvariety of reductive naturalism in some particular area of debate may sometimes be seen, or represented, as a kind of skeptic in that area: say, a moral skeptic or a skeptic about the mental or about abstract entities or about what are called 'intensions'" (1985: 2). He continues:

> It is reductive naturalism which holds that the naturalistic or objective view of human beings and human behavior undermines the validity of moral attitudes and reactions and displays moral judgment as no more than a vehicle of illusion (1985: 43).

Reductive naturalists are aligned with Strawson's genuine moral skeptic, for they reject whatever cannot be reduced. Not surprisingly, Strawson favors his own naturalism.

Strawson's naturalism "provides for a richer conception of the real, making room, for example, for morality and moral responsibility, for sensible qualities as genuinely characterizing physical things, for determinate meanings, meaning-rules, and universals—all as we ordinarily conceive them" (1998a: 168). According to Strawson's naturalism, skeptical criticisms "are not to be met with argument" but "are simply to be neglected" (1985: 13). Hence:

> To try to meet the skeptic's challenge, in whatever way, by whatever style
> of argument, is to try to go further back. If one is to begin at the begin-
> ning, one must refuse the challenge as our naturalist refuses it. (1985:
> 24–25)

Implicit in this nonresponse to the skeptic is the suggestion that there are
no adequate, rational replies available. Strawson's view is a kind of *concessive*
response to the skeptical challenge (DeRose 1999: 19–22). Strawson writes:
"the point has been, not to offer a logical justification of the belief in exter-
nal objects and other minds or the practice of induction, but to represent
skeptical arguments and rational counter-arguments as equally idle—not
senseless, but idle—since what we have here are original, natural, inescap-
able commitments which we neither choose nor could give up" (1985: 28; cf.
1998b: 242).

Strawson's concession does not result in a victory for the skeptic since skep-
tical arguments have no lasting impact. We continue to act as if those beliefs
and attitudes that are part of our hinge commitments are true independent of
our inability to respond to skeptical criticisms. Skeptical arguments are, thus,
"idle" and have no effect on our beliefs. Yet it isn't just skeptical arguments that
are idle, for "arguments on both sides are idle" (1985: 29). Reason is ineffec-
tive in responding to the skeptic and it is equally ineffective in grounding our
hinge commitments. Thus, there are two sides of the idleness of hinge commit-
ments. First, there is no adequate response to the skeptic. Call this "skeptical
idleness." Second, hinge commitments are not supported by arguments. Call
this "grounding idleness." These two forms of idleness are explained in more
detail below.

I've been careful not to classify hinge commitments as *beliefs*. They are at-
titudes, some of which may be manifested as beliefs. Consider the issue of im-
plicit bias. Neil Levy (2014) argues that some implicit biases are not beliefs,
for they are not robust or "thick" enough to qualify. This has a bearing on our
judgments of blameworthiness for actions that are the result of implicit biases.
Whether Levy is right or not is important but not for the purposes of this paper.
I use this only to help make a point. Some hinge commitments are beliefs. Oth-
ers are more like Levy's implicit biases in the sense that we don't have enough
conscious awareness of them for them to qualify as beliefs. Regardless, com-
mitments may play a role in our behavior similar to beliefs.

To get a better idea of the grounding idleness of hinge commitments, con-
sider a short version of G. E. Moore's proof of the external world (Moore 1959;
cf. Wright 2004).

> Premise: Here is a hand.
> Conclusion: Therefore, the external world exists.[9]

Something is a *logical justification* for a belief *p* only if it attempts to offer a reason to think that *p* is likely to be true. Moore takes a shot at providing a logical justification for the belief that the external world exists. The problem with this approach is that when hinge commitments are manifested as beliefs they usually take the form of a *Moorean fact*, "one of those things that we know better than we know the premises of any philosophical argument to the contrary" (Lewis 1999: 220)—or, one might add, that we know better than the premises of any philosophical argument in its favor. How can my belief that I have a hand support my claim that the external world exists, since the former seems to presuppose the latter? What *could* support my belief that I have a hand?

On the other hand, a *pragmatic justification* for *p* gives a reason for believing that *p* without providing a reason to think that *p* is likely to be true. Logical and pragmatic justifications are both *rational justifications* since they provide reasons for believing that *p*. Pascal's wager is a good example of a pragmatic justification for belief in the existence of God. Pascal offers reasons for believing in the existence of God—e.g., it is better for you if you do than if you don't—but those reasons do not render it likely that God exists. Pascal is a skeptic about religious belief. It might be that each of us would be better off believing in the existence of God. Perhaps the benefits of belief far outweigh the dangers of disbelief. Yet that does not render it more probable than not that God exists.

Examples of hinge commitments include: that the external world exists, that other minds exist, that the future will be like the past, that the past exists or at least did exist, that we have free will, that we are morally responsible for some of our actions, and so on. My claim is that Strawson is not providing a rational justification—neither a logical nor a pragmatic justification—for our hinge commitments. Rather, he tries to explain the appropriateness of those commitments. His claim is not that we *should* accept them but that we *do* accept them and that we do so independent of reasoning or argument. In addition, hinge commitments are appropriate because of the fundamental role that they play in our overall system of beliefs and attitudes. They are the given in rational discourse.

---

9 At least one critic (Sosa 1998: 366–367) interprets Strawson as providing a similar justification for our belief in the external world. See below and Strawson (1998e) for a reply to this interpretation.

Strawson's comments against free will skepticism do not make it more probable than not that some persons have free will. His response to the free will skeptic is part of a more general strategy. Strawson writes:

> The correct way with the professional skeptical doubt is not to attempt to rebut it with argument, but to point out that it is idle, unreal, a pretense; and then the rebutting arguments will appear as equally idle; the reasons produced in those arguments to justify induction or belief in the existence of body are not, and do not become, *our* reasons for these beliefs; there is no such thing as *the reasons for which we hold* these beliefs. We simply cannot help accepting them as defining the areas within which the questions come up of what beliefs we should rationally hold on such-and-such a matter. (1985: 19–20)

In light of this observation, Strawson's naturalism is not a version of the constitutive theory, for the constitutive theory is a kind of reductive naturalism, a view against which Strawson distinguishes his own naturalism. Nor is it a logical justification of our attitudes and practices, for Strawson explicitly denies that it is.

Does Strawson provide a pragmatic justification for compatibilism or free willism? Remember, there are two sides to the idleness of hinge commitments. Hinge commitments are idle because skeptical arguments have no lasting impact. I go out the door, not the window even if I have no rational response to the problem of induction (Hume 1947: Part I). In addition, they are idle because it is unclear exactly what could serve the role as a reason to accept a hinge commitment. Arguments in their favor are even odder than the skeptical arguments against them. Wittgenstein asks: "Does my telephone call to New York strengthen my conviction that the earth exists?" (1969: §210) Once you seek a rational response, seek to provide a rational justification for a hinge commitment, you fall into the skeptic's trap. What you seek cannot be found. On the field of reason, the skeptic wins. This is Strawson's concession to skepticism.

The concession to skepticism does not count for much since we retain our hinge commitments, including many ordinary beliefs, in spite of the skeptical objections against them. "The nonreductive naturalist's point," Strawson writes, "is that there can only be a *lack* where there is a *need*" (1985: 41). We believe that we have free will and that we are morally responsible for some of our actions. That we have these beliefs is confirmed by our behavior, especially our behavior as manifested by our reactive attitudes. The reactive attitudes serve

to establish, not justify, our acceptance of hinge commitments related to free will and moral responsibility. Strawson shows that these commitments remain even after we recognize that there are powerful arguments against them. All of this goes against the claim that Strawson provides a rational justification for compatibilism or free willism, whether these claims are supported by pragmatism, the constitutive theory, or whatever.[10]

5   McKenna's Four Arguments and the Argument from Excuses

There are various counterproposals to my interpretation of Strawson. R. Jay Wallace (1994) contends that there are at least three arguments for compatibilism in Strawson (1962): the naturalist argument, the pragmatist argument, and the internalist justification argument. McKenna (2005) does Wallace one better and claims that Strawson (1962; 1985) offers four arguments for compatibilism (see below). Waller (2006) suggests that Strawson gives the excuse-extensionism argument and Pereboom (2014) presents at least two arguments. Note that what is contended is that Strawson offers several arguments for compatibilism. But, if sound, they could easily be turned into arguments for free willism, as well, since the reason determinism is irrelevant to free will is that we have a hinge commitment to free willism that is more basic than and independent of general philosophical theses like determinism.

McKenna provides the most helpful taxonomy of Strawson's arguments since each of the others noted above is debatably a version of one of his arguments. Here are McKenna's four arguments (2005: 166–168), along with the other references:

- the practical rationality argument (Wallace 2006: 99–103);
- the argument from excuses (Waller 2006: 81–83);
- the internalist justification argument (Wallace 2006: 125–127; Pereboom 2014: 153–154, 179);
- the psychological impossibility argument (Wallace 2006: 28–38; Pereboom 2014: 178ff.).

---

10   Galen Strawson understands Strawson's naturalism better than others, for he notes that it "stresses our commitment to certain *attitudes* and *practices* which appear to presuppose belief in true responsibility" (2010: 72). Yet he also suggests that, according to Strawson, this commitment justifies our belief in moral responsibility (2010: 74).

The practical rationality argument is just a form of pragmatism, discussed in the previous section, so our interest in the remainder of the essay is with the last three arguments.

We begin with this question: should arguments for incompatibilism or free will skepticism make us look at "the normal...in the objective way" (1962: 54), meaning, should our philosophical conclusions lead us to adopt the objective attitude to all persons, including "the normal"? Strawson poses two responses to this question:

> The first is that we cannot, as we are, seriously envisage ourselves adopting a thoroughgoing objectivity of attitude to others as a result of theoretical conviction of the truth of determinism; and the second is that when we do in fact adopt such an attitude in a particular case, our doing so is not the consequence of a theoretical conviction which might be expressed as 'Determinism in this case', but is a consequence of our abandoning, for different reasons in different cases, the ordinary interpersonal attitudes. (1962: 55; cf. 1998c: 259–260)

Strawson's first point is that, though we can adopt the objective attitude in particular cases, "Being human, we cannot, in the normal case, do this for long, or altogether" (1962: 52–53), nor can we do it on the basis of adopting a general thesis like determinism. This is plausibly a version of the psychological impossibility argument, which claims that the acceptance of free will skepticism—a consequence of the joint acceptance of incompatibilism and determinism— is psychologically impossible. The psychological impossibility argument is discussed in detail below.

Strawson's second point is that the assumption of determinism never plays a role when we adopt the objective attitude in particular cases. This is plausibly a version of the argument from excuses. The argument from excuses is best understood as a response to arguments for incompatibilism that incorporate the *generalization strategy*. Wallace writes:

> the incompatibilist might argue, we think it unfair to hold those responsible who are physically constrained, or insane, or under the influence of a drug or hypnotic treatment that has been forcibly administered, and these judgments rely at least implicitly on moral principles that specify what it is about these kinds of conditions that makes it unfair to hold people responsible when they obtain. ... Incompatibilists have traditionally supposed that the excuses and exemptions we acknowledge in practice commit us to principles ... to the effect that we should not hold

people morally responsible unless they have the ability to do otherwise.
(1994: 115)

A proponent of the generalization strategy starts by observing circumstances
in which we are prone to mitigate moral responsibility. The hope is that we
discover that, say, the lack of the ability to do otherwise is a common feature
in cases of mitigation. On the basis of this feature, one can formulate a general
thesis, like the principle of alternative possibilities (PAP), which connects the
ability to do otherwise with moral responsibility. The proponent of the gener-
alization strategy then argues that such freedom-making properties are never
instantiated in deterministic worlds, so incompatibilism is true. As I under-
stand it, generalization strategies are not committed to PAP and might incor-
porate other claims. Strawson's comments are at least responses to arguments
for incompatibilism but are they also arguments for compatibilism, as Wallace
and McKenna take them to be?

Strawson's list of mitigating factors is precise and extensive (Watson 1987).
Strawson considers various "occasions for resentment" (Strawson 1962: 50–52).
He then makes a division and two additional subdivisions, one for each divi-
sion. First, he distinguishes between *exemptions* and *excuses*. In the case of
exemptions, the agent is not a player in the desert game, either for reasons that
are *transitory* ("She wasn't herself") or *enduring* ("She is only a child"). Excuses
are more complicated and philosophical for they establish a failure of either
the *epistemological* ("She didn't realize") or the *freedom-relevant* ("She couldn't
help it") condition for moral responsibility (Campbell 2005: 401–403).

Is Strawson's taxonomy part of an attempt to argue for compatibilism or free
willism? It is difficult to draw this conclusion, for Strawson seems to be up to
something else. On my view, he is reacting and responding to arguments for
incompatibilism, rather than giving an argument for compatibilism. Accord-
ing to Strawson's taxonomy, there is nothing about exemptions and/or excuses
that can be generalized to formulate some PAP-like principle to be used in an
argument for incompatibilism or free will skepticism, or related claims about
moral responsibility. This is merely a criticism of arguments for incompatibil-
ism that adopt the generalization strategy, not an argument for compatibilism.

6 The Psychological Impossibility Argument and the Argument from Excuses Revisited

The psychological impossibility argument is complex. First, it is claimed that
Strawson is wrong to suppose that "an exclusive objectivity of attitude" is

"psychologically impossible" or "practically inconceivable" (Sommers 2007). According to Russell, Strawson admits that we may adopt the objective attitude in cases where we consider an agent to be, say, incapacitated. But, says Russell, the free will skeptic's view is just that *everyone* is similarly afflicted. He continues: "Obviously, it is not inconceivable or self-contradictory to suggest that there could be a world, or things might develop, such that everyone is or becomes incapacitated" (1992: 299).

Russell believes that Strawson adopts both a naturalistic and a rationalistic strategy in response to the related challenges of incompatibilism and free will skepticism. The rationalistic strategy holds that "considerations of determinism, however they are interpreted, do not, as such, provide us with any reason to modify or suspend our reactive attitudes. ... We have, accordingly, no reason whatsoever to suspend or abandon our reactive attitudes entirely even if the thesis of determinism is true" (1992: 290; cf. Smilansky 2001: 73).

The following quote appears to be an instance of Strawson's application of the rationalistic strategy, as Russell understands it:

> The human commitment to participation in ordinary interpersonal relationships is, I think, too thoroughgoing and deeply rooted for us to take seriously the thought that a general theoretical conviction might so change the world that, it is, there were no longer any such things as interpersonal relationships as we normally understand them... (1985: 54)

The claim by Russell is that Strawson's comments about the steadfastness of our reactive attitudes are best understood as a rational justification for compatibilism, perhaps some combination of the psychological impossibility argument and the argument from excuses. Russell writes: "I am concerned with Strawson's specific arguments purporting to show that the truth of the thesis of determinism cannot lead to the conclusion that global excusing considerations apply to everyone" (1992: 298). He continues:

> Strawson...must establish, against the Pessimist, that determinism does not (or cannot) imply that everyone is 'abnormal.' Failing this, the rationalistic strategy would collapse. Nevertheless, the argument which Strawson puts forward is wholly inadequate. (1992: 299)

Russell notes that the rationalistic strategy is in direct conflict with the naturalistic strategy.

According to my interpretation, Strawson does not attempt to provide a rational justification for hinge commitments. This is not the place to discuss whether or not naturalistic responses to skeptical arguments are satisfactory.

What is important, for our purposes, is that Strawson's naturalism is not intended as a rational justification for hinge commitments. Rational justification requires a need and, according to Strawson, there is no need in the case of hinge commitments. Therefore, Russell's claim that Strawson adopts inconsistent strategies is incorrect.

On Pereboom's reading of Strawson, moral responsibility has its foundation in the reactive attitudes, which Strawson requires for the kinds of interpersonal and self-directed commitments that make our lives meaningful. Judgments of blameworthiness and praiseworthiness are ultimately grounded in our blaming and praising practices, particularly the corresponding reactive attitudes. For this reason, general claims like determinism cannot be used to undermine compatibilism, free will, or moral responsibility (2014: 178ff.). Pereboom writes in response to this view: "But I maintain that expressions of these reactive attitudes are suboptimal as modes of communication in relationships relative to alternative attitudes available to us" (2014: 179). In other words, we have reason to resist the reactive attitudes especially if it turns out that determinism is true.

The second stage of this criticism takes things even further and suggests that adopting the objective attitude is not merely possible but is preferable. Watson writes:

> Some have aspired to rid themselves of the readiness to limit goodwill and to acquiesce in the suffering of others not in order to relieve the strains of involvement, nor out of a conviction in determinism, but out of a certain ideal of human relationships, which they see as poisoned by the retributive sentiments. It is an ideal of human fellowship or love which embodies values that are arguably as historically important or our civilization as the notion of moral responsibility itself. The question here is not whether this aspiration is finally commendable, but whether it is compatible with holding one another morally responsible. (1987: 285–286)

Watson concludes: "holding one another responsible is at odds with one historically important ideal of love" (1987: 286). Similarly, Galen Strawson claims that the objective attitude is connected in important ways to the Buddhist doctrine of satkayadrsti and the attainment of nirvana (2010: 101–103).

Free will skeptics are split about the objective attitude. Commenting on the relationship between Albert Einstein's free will skepticism and his "sense of distance" and "need for solitude," Watson writes:

> The 'distance' of which Einstein speaks is just an aspect of the 'detachment' Strawson thinks characterizes the objective stance. At its extremes, it takes the form of human isolation. (1987: 285)

Watson then adds that some folks, like Mahatma Gandhi and Martin Luther King, who embrace the ideal of love noted above, "do not seem to adopt an objective attitude in Strawson's sense" (1987: 286). Galen Strawson, on the other hand, holds that "whatever nirvana is supposed to be like, it is clear that adoption of the objective attitude is in no way incompatible with compassion" (2010: 103).

Galen Strawson identifies the objective attitude with the mere acceptance of free will skepticism whereas Watson thinks of it as characterized by a more fine-grained set of psychological attitudes.[11] Where they both agree is in thinking that the adoption of free will skepticism, and the belief that no one is morally responsible for anything, is not only possible but marks the achievement of a human ideal. For this reason, Watson describes one of Strawson's passages as "troubling" (1987: 285). Russell echoes this sentiment: "What is particularly disturbing about Strawson's naturalistic strategy, expressed in more general terms, is that it casts doubt on our ability or capacity to curb or control our emotional life according to the dictates of reason" (1992: 297; see also Pereboom 2014; Wallace 1994: 28–33).

As powerful as this set of criticisms is, I still think that it is misplaced. To begin, free will skeptics often admit that "an *exclusive* objectivity of attitude" is impossible. At the end of his essay, Sommers writes:

> I concede that we are deeply committed to seeing others, and especially ourselves, as morally responsible agents, and as appropriate candidates for attitudes like resentment. I concede further that it is hard to imagine what it would be like to give up this belief entirely. But is an almost total embrace of the objective attitude *impossible*? (2007: 337)

At one point, Galen Strawson weakens his criticism and asks whether or not "one might be able to engineer…partial but not total erosions" of our commitment to free will (2010: 101)? Pereboom writes: "A certain measure of indignation and resentment is likely to be beyond our power to affect, and thus even supposing that one is committed to doing what is right and rational, one would still be unable to eradicate these attitudes" (2001: 200). In a more recent work, he adds:

> Thus I'm not committed to the view that we can generally succeed in overcoming moral resentment and indignation, but rather to the

11   Watson's view on this issue is shared by Pereboom (2001: 199–200; 2014), and G. Strawson's view is shared by Sommers (2007), among others.

> proposal that we can rest such attitudes and limit their expressions with
> some success, and that we can oppose actions and polices justified on
> the basis of the beliefs about basic desert that accompany such attitudes.
> (2014: 182)

It is not clear that Strawson has much to say about the degree to which we
could or should *limit* our reactive attitudes. The issue for him is whether we
could give them up entirely. Strawson need not disagree with free will skeptics
on the points expressed above.

Consider other forms of skepticism, like skepticism about the external
world or skepticism about induction. According to Strawson, skeptical threats
of this kind are idle and have no effect on our beliefs. A similar point is made
by Hume, who in *Dialogues Concerning Natural Religion*, has the empiricist
Cleanthes say to the skeptic Philo:

> Whether your scepticism be as absolute and sincere as you pretend, we
> shall learn by and by, when the company breaks up: we shall then see,
> whether you go out at the door or the window; and whether you really
> doubt if your body has gravity, or can be injured by its fall; according to
> popular opinion, derived from our fallacious senses, and more fallacious
> experience. (1947: Part I)

Wittgenstein echoes similar sentiments when he writes: "My life shews that I
know or am certain that there is a chair over there, or a door, and so on" (1969:
§7).

On my reading of Strawson, his comments about the steadfastness of our be-
lief in free will, even in light of compelling arguments against it, are merely to
indicate that we do in fact have such hinge commitments; his intent is to iden-
tify those commitments, not to justify them. Hinge commitments are revealed
by our actions in spite of, and often against, our philosophical proclamations.
Strawson is not denying that there are people who profess that no one is mor-
ally responsible for their actions, just as he wouldn't deny that there are people
who claim to be solipsists. Yet at the end of the day, all of them go out the door
and not the window. Many claim to be free will skeptics but Strawson's view
is that they cannot always act in accordance with their free will skepticism
given their hinge commitments. In as much as the free will skeptic exhibits the
practices of blaming and praising others, her actions suggest underlying com-
mitments that go against her skeptical convictions. Perhaps we can stop our
blaming and praising practices?

Part of Strawson's reply to the free will skeptic rests on the analogy between free will skepticism and kinds of epistemological skepticism. In response to skepticism about the past, Strawson writes:

> belief in the reality and determinateness of the past is as much part of that general framework of beliefs to which we are inescapably committed as is the belief in the existence of physical objects and the practice of inductive belief-formation. … All form part of our mutually supportive natural metaphysics. (1985: 29; see also 2011b: 105)

According to Strawson's naturalism, our beliefs in the existence of the past, the reality of the external world, and our inductive practices are part of our *natural metaphysics*, which is based on "an original non-rational commitment which sets the bounds within which, or the stage upon which, reason can effectively operate, and within which the question of the rationality or irrationality, justification or lack of justification, of this or that particular judgment or belief can come up" (1985: 39). Wittgenstein calls these commitments "the inherited background against which I distinguish between true and false" (1969: § 94) and "the substratum of all my inquiring and asserting" (1969: § 162; see also Strawson 1985: 15).

Some commentators suggest that the analogy between epistemological skepticism and free will skepticism is faulty. Sommers writes:

> Perhaps we cannot prove that it is true that we have a body, but we have no reason to think that it is false. By contrast, there are valid arguments with true—or at least plausible—premises which conclude that there is no such thing as robust moral responsibility. (2007: 337)

Pereboom also finds fault with the free will/empirical skepticism analogy (2014: 154–155). Pereboom agrees that something like the internalist justification argument applies in the case of induction, because internalist justifications of induction are plausibly either circular or question begging. For this reason, arguments against induction are idle. On the other hand, arguments for incompatibilism and free will skepticism have more bite. Pereboom adopts a generalization strategy for incompatibilism that is not dependent on PAP, one that incorporates a version of the manipulation argument (2014: Ch. 4). He also gives compelling criticisms against plausible libertarian views (2014: Chs. 2–3) in order to establish free will skepticism. This is discussed more fully in the final section.

Next there is the claim that the adoption of a thoroughgoing, or exclusive, objectivity of attitude is a human ideal. It is one thing to suggest that no one is blameworthy for his actions and quite another to suggest that no one is morally responsible for anything. Socrates, for instance, held that no one is blameworthy since wrong deeds are done out of ignorance. Socrates' position leaves open the possibility of free willism. Of course, some of these issues hinge on one's thoughts about the epistemic condition for moral responsibility (see Fischer and Ravizza 1993b).

Free will skeptics rarely discuss praiseworthiness. A life without praise is not in any sense an ideal moral life. It is ironic that the views of Buddha, Gandhi, and King are used to support the possibility of complete objectivity of attitude yet each has led a praiseworthy life. If some of us are praiseworthy, then free will skepticism is false.

Is Strawson committed to the psychological impossibility of endorsing free will skepticism? Or to put the question another way, *can* we form judgments that go against our hinge commitments? In a sense the answer is 'yes.' and in a sense the answer is 'no.' Suppose the question is: As we stand (holding every fact of the moment fixed), can we form judgments that go against our hinge commitments? In other words, do we have *specific abilities* to alter our hinge commitments? That is doubtful. On the other hand, it is likely that we have *general abilities* to alter our hinge commitments over time (cf. Campbell 2013).

As an example of a hinge commitment, Wittgenstein mentions "no one has ever been to the moon" (1969: §106). This is no longer a hinge commitment since it is generally denied but at one time it plausibly was a hinge commitment. Strawson's claim is that, as we stand, we cannot go against our hinge commitments. This means that we lack the specific ability to alter them and does not preclude that hinge commitments, like most things, are subject to change and the influence of our general abilities, as many free will skeptics admit.

7   The Internal Justification Argument and Concluding Remarks

The internal justification argument, offered by McKenna (2012) and Pereboom (2014), goes beyond the core assumption, for it's based in part on Strawson's later views and takes seriously my interpretation of Strawson's naturalism. It is worth comparing Pereboom's philosophical project with Strawson's before concluding.

One goal of Pereboom (2014) is to show that free will skepticism is not "devastating to our conceptions of agency, morality, and meaning in life" (2014: 4).

Free will skepticism is compatible with "practically viable notions of morality and moral responsibility" (2014: 4; see also Ch. 6) such as "our practice of holding responsible" (2014: 154). It is also compatible with rational deliberation (2014: 4, Ch. 5) as well as models of the criminal justice system that depend on forward-looking, deterrence, utilitarian, or consequentialist theories of punishment (2014: 7, Ch. 7). What the free will skeptic cannot have is robust moral responsibility and anything that depends on it, such as retributivist justifications of criminal justice (2014: 7, 157–160).

Strawson's view runs counter to Pereboom's because the former believes that criticisms from the free will skeptic are idle. As Pereboom notes, for Strawson "the truth of determinism is irrelevant to whether we have the sort of free will required for moral responsibility" (2014: 72). Pereboom holds that one of Strawson's arguments for the core assumption is the internalist justification argument:

> the rules for excusing and exempting internal to the practice of holding responsible will not license a critique of this practice based on universal determinism, for such a critique of this practice will be based on factors external to the practice, and will therefore be illegitimate. (2014: 72)

As Pereboom understands Strawson, "the entire practice of holding morally responsible is insulated from general metaphysical and scientific objections, for the reason that these challenges are external and thus illegitimate" (2014: 153; see also 154, 179).

It is difficult to access the upper hand in the debate between Strawson and Pereboom but that is not my job. I need only to show that my interpretation of Strawson is a horse in the race, a viable point of view. Strawson would accuse Pereboom of making the same mistake as the agent causation theorist and the empirically minded libertarian, the mistake of supposing that free will requires some special accommodation. Strawson writes: "it is quite false that these requirements of justice can only be met by the satisfaction of some condition of *ultimate* responsibility which can no more be coherently stated than can the libertarian's conception of free will" (1998a: 170).

A hinge commitment is "internal to the structure of all thinking, so that the attempt to question it, which is tantamount to an attempt to reject our conceptual scheme in its entirety, leaves us without the resources for any coherent thought at all" (1998d: 291). One might protest that there should be something more to say in defense of our attitudes and practices than that we have them. Is it just that we are stubborn? Wouldn't the same defense work for deniers of evolution or climate change?

As I understand Strawson, this leads us back into the skeptic's trap. Not every passionate belief counts as a hinge commitment. We've established the two sides of Strawson's idleness criticism and we can use that as a basis to distinguish hinge commitments from other merely obstinate attitudes and beliefs. On the one hand, we need to ask whether skeptical arguments against the commitment are compelling otherwise the non-reductive naturalist is on a par with the obstinate. Of course, should a compelling argument be given against a commitment it would cease to be a hinge commitment anyway, for it would no longer be idle. On the other hand, we may ask whether an argument for the commitment sounds odd. This marks a difference between claims like there is an external world and claims like evolution is false. In the latter case, we can envisage an argument that might prove us wrong but not so in the former case.

Beyond this we can't satisfy the itch. The trick is that we have hinge commitments, some of which are manifested as beliefs. That does not justify them but we have them nonetheless and they play an important role in our overall system of beliefs. I close with two quotes from Wittgenstein (1969):

· When we first begin to *believe* anything, what we believe is not a single proposition, it is a whole system of propositions. (Light dawns gradually over the whole.) (§141, quoted in Strawson 2011b: 108)
· At the foundation of well-founded belief lies belief that is not founded (§253).[12]

References

Balaguer, M. (2010). *Free Will as an Open Scientific Problem*. Cambridge, MA: MIT Press.

Bennett, J. (2008). "Accountability (II)." In M. McKenna and P. Russell (eds.), *Free Will and Reactive Attitudes: Perspectives on P.F. Strawson's "Freedom and Resentment"*, 47–68. Burlington, VT: Ashgate Publishing Company.

Campbell, J. (2005). "Comptatibilist Alternatives," *Canadian Journal of Philosophy* 35: 387–406.

Campbell, J. (2011). *Free Will*. Polity Press.

Campbell, J. (2013). "Two Problems for Classical Incompatibilism." In I. Haji and J. Caouette (eds.), *Free Will and Moral Responsibility*, 26–42. Newcastle upon Tyne: Cambridge Scholars Publishing.

Campbell, J. (2015). "Critical Notice: Paul Russell's *The Riddle of Hume's Treatise: Skepticism, Naturalism, and Irreligion*," *Canadian Journal of Philosophy* 45: 127–137.

DeRose, K. (1999). "Introduction: Responding to Skepticism." In K. DeRose and T. A. Warfield (eds.), *Skepticism: A Contemporary Reader*, 1–24. New York and Oxford: Oxford University Press.

Fischer, J. M. and M. Ravizza. (eds.). (1993a). *Perspectives on Moral Responsibility*. Ithaca and London: Cornell University Press.

Fischer, J. M. and M. Ravizza. (eds.). (1993b). "Introduction." In Fischer and Ravizza (1993a), 1–41.

Hahn, L. E. (ed.). (1998). *The Philosophy of P.F. Strawson*. Chicago and LaSalle, IL: Open Court.

Honderich, T. (2004). "After Compatibibilism and Incompatibilism." In J. Campbell, M. O'Rourke, and D. Shier (eds.), *Freedom and Determinism*, 305–321. Cambridge, MA: MIT Press.

Hume, D. (1947). *Dialogues Concerning Natural Religion*. Indianapolis: Bobbs-Merrill.

Hume, D. (1975). *An Enquiry Concerning Human Understanding*. New York and Oxford: Oxford University Press.

Kane, R. (1996). *The Significance of Free Will*. New York and Oxford: Oxford University Press.

Kemp Smith, N. (1966). *The Philosophy of David Hume: A Critical Study of Its Origins*. New York: St. Martin's Press.

Lewis, D. (1999). "Elusive Knowledge." In DeRose and Warfield (1999), 220–239. (Originally published in 1996.)

Levy, N. (2011). *Hard Luck: How Luck Undermines Free Will and Moral Responsibility*. New York and Oxford: Oxford University Press.

Levy, N. (2014). "Consciouness, Implicit Attitudes and Moral Responsibility," *Noûs* 48: 21–40.

McKenna, M. (1998). "The Limits of Evil and the Role of Moral Address: A Defense of Strawsonian Compatibilism," *The Journal of Ethics* 2: 123–142.

McKenna, M. (2005). "Where Strawson and Frankfurt Meet," *Midwest Studies in Philosophy* 29: 163–180.

McKenna, M. (2012). *Conversation and Responsibility*. New York and Oxford: Oxford University Press.

Mele, A. R. (2006). *Free Will and Luck*. New York and Oxford: Oxford University Press.

Mele, A. R. (2014). *A Dialogue on Free Will and Science*. New York and Oxford: Oxford University Press.

Moore, G. E. (1959). *Philosophical Papers*. London: George Allen and Unwin.

Nowell-Smith, P. (1948). "Freewill and Moral Responsibility," *Mind* 57: 45–61.

Pereboom, D. (2001). *Living Without Free Will*. Cambridge: Cambridge University Press.

Pereboom, D. (2014). *Free Will, Agency, and Meaning in Life*. New York and Oxford: Oxford University Press.

Pritchard, D. (2015). *Epistemic Angst: Radical Scepticism and the Groundlessness of Our Believing*. Princeton, NJ: Princeton University Press.

Russell, P. (1992). "Strawson's Way of Naturalizing Responsibility," *Ethics*: 102: 287–302.

Scanlon, T. M. (1988). "The Significance of Choice." In S. M. McMurrin (ed.), *The Tanner Lectures on Human Values*, 151–177. Salt Lake City, UT: University of Utah Press. Reprinted in Watson (2003), 352–371; page numbers are to this latter edition.

Schlick, M. (1939). *The Problems of Ethics*. Translated by D. Rynin. New York: Prentice Hall.

Smart, J. J. C. (1961). "Freewill, Praise and Blame," *Mind*: 70: 291–306.

Smilansky, S. (2001). "Free Will: From Nature to Illusion," *Proceedings of the Aristotelian Society* 101: 71–95.

Sommers, T. (2007). "The Objective Attitude," *The Philosophical Quarterly* 57: 321–341.

Sosa, E. (1998). "P.F. Strawson's Epistemological Naturalism." In Hahn (1998), 361–369.

Strawson, G. (2010). *Freedom and Belief*, revised edition. New York and Oxford: Oxford University Press. (Originally published in 1986.)

Strawson, P. F. (1962). "Freedom and Resentment." *Proceedings of the British Academy* 48: 1–25. Reprinted in Fischer and Ravizza (1993a), 45–66; page numbers are to this latter edition.

Strawson, P. F. (1980). "P. F. Strawson Replies." In van Straaten (1980), 260–296.

Strawson, P. F. (1985). *Skepticism and Naturalism: Some Varieties*. London: Methuen.

Strawson, P. F. (1998a). "Reply to Simon Blackburn." In Hahn (1998), 168–173.

Strawson, P. F. (1998b). "Reply to Andrew G. Black." In Hahn (1998), 242–244.

Strawson, P. F. (1998c). "Reply to David Pears." In Hahn (1998), 259–262.

Strawson, P. F. (1998d). "Reply to Hillary Putnam." In Hahn (1998), 288–292.

Strawson, P. F. (1998e). "Reply to Ernest Sosa." In Hahn (1998), 370–372.

Strawson, P. F. (2011a). *Philosophical Writings*. Edited by GalenStrawson and Michelle Montague. New York and Oxford: Oxford University Press.

Strawson, P. F. (2011b). "Does Knowledge Have Foundations." In Strawson (2011a), 100–108. (Originally published in 1974.)

Strawson, P. F. (2011c). "Liberty and Necessity." In Strawson (2011a), 146–156. (Originally published in 1983.)

Strawson, P. F. (2011d). "Two Conceptions of Philosophy." In Strawson (2011a), 166–177. (Originally published in 1990.)

Strawson, P. F. (2011e). "Intellectual Autobiography." In Strawson (2011a), 227–247. (Originally published in 1998.)

van Inwagen, P. (1983). *An Essay on Free Will*. Oxford: Clarendon Press.

Van Straaten, Z. (ed.). (1980). *Philosophical Subjects*. Oxford: Clarendon Press.

Vihvelin, K. (2008). "Compatibilism, Incompatibilism, and Impossibilism." In T. Sider, J. Hawthorne, and D. W. Zimmerman (eds.), *Contemporary Debates in Metaphysics*, 303–318. Oxford: Blackwell Publishing.

Wallace, R. J. (1994). *Responsibility and the Moral Sentiments*. Cambridge, MA: Harvard University Press.

Waller, B. (2003). "A Metacompatibilist Account of Free Will: Making Compatibilists and Incompatibilists More Compatible," *Philosophical Studies* 112: 209–224.

Waller, B. (2006). "Denying Moral Responsibility Without Making Excuses," *American Philosophical Quarterly* 43: 81–90.

Watson, G. (1987). "Responsibility and the Limits of Evil: Variations on a Strawsonian Theme." In F. Schoeman (ed.), *Responsibility, Character and the Emotions*, 256–286. Cambridge: Cambridge University Press. Reprinted in Fischer and Ravizza (1993a), 119–148; page numbers are to this latter edition.

Watson, G. (ed.). (2003). *Free Will*, 2nd Edition. New York and Oxford: Oxford University Press.

Wiggins, D. (2003). "Towards a Reasonable Libertarianism." In Watson (2003), 94–121.

Wittgenstein, L. (1969). *On Certainty*. New York: Harper & Row Publishers.

Wright, C. (2004). "Wittgensteinian Certainties." In Denis McManus (ed.), *Wittgenstein and Scepticism*, 22–55. London and New York: Routledge.