Free Will and Reactive Attitudes

Perspectives on P.F. Strawson's "Freedom and Resentment"

Edited by

MICHAEL McKENNA
Florida State University, USA

PAUL RUSSELL
University of British Columbia, Canada

ASHGATE

Contents

List of Contributors *vii*
Acknowledgements *xi*

Introduction: Perspectives on P.F. Strawson's "Freedom and Resentment"
Michael McKenna and Paul Russell 1

1 Freedom and Resentment
Peter Strawson 19

2 Free Will and Rationality
A.J. Ayer 37

3 Accountability (II)
Jonathan Bennett 47

4 The Importance of Free Will
Susan Wolf 69

5 On "Freedom and Resentment"
Galen Strawson 85

6 Responsibility and the Limits of Evil:
Variations on a Strawsonian Theme
Gary Watson 115

7 Strawson's Way of Naturalizing Responsibility
Paul Russell 143

8 Emotions, Expectations and Responsibility
R. Jay Wallace 157

9 Blaming, Understanding and Justification: A Defence of Strawson's
Naturalism about Moral Responsibility
Kevin Magill 187

10 The Limits of Evil and the Role of Moral Address:
A Defense of Strawsonian Compatibilism
Michael McKenna 201

11 Revising the Reactive Attitudes
Derk Pereboom 219

12 Free Will: From Nature to Illusion
Saul Smilansky 235

13 Thinking with your Hypothalamus:
Reflections on a Cognitive Role for the Reactive Emotions
David Zimmerman 255

14 Doing without Desert
Erin Kelly 273

15 Responsibility and the Aims of Theory: Strawson and Revisionism
Manuel Vargas 297

Suggested Further Readings *319*
Index *323*

Introduction

Perspectives on P.F. Strawson's "Freedom and Resentment"

Michael McKenna and Paul Russell

> We are naturally social beings; and given with our natural commitment to social existence is a natural commitment to that whole web or structure of human personal and moral attitudes and feelings, and judgments of which I spoke. Our natural disposition to such attitudes and judgments is naturally secured against arguments suggesting they are in principle unwarranted or unjustified ...
>
> – P.F. Strawson, Skepticism and Naturalism

During the past half century the free will problem has received a considerable amount of philosophical attention. This activity has resulted in a significant advance in the level of sophistication and subtlety in the various theories on offer. P.F. Strawson's "Freedom and Resentment", which was first published in 1962, is among the most important and influential of the contributions produced during this period.[1] In order to appreciate the significance of Strawson's contribution fully his paper must be viewed from two directions. On one side, it must be viewed from the perspective of the "classical" free will debate as it was generally understood around the middle of the twentieth century. On the other side, we need to consider "Freedom and Resentment" in terms of the critical responses and debates that it has generated. The aim of this introduction is to provide the reader with a general framework for understanding the significance of Strawson's contribution from both these perspectives.

I. The Classical Debate and the Dilemma of Determinism

As Strawson presents the issue, the problems that primarily concern the classical free will debate can be described in terms of the dilemma of determinism. The thesis of determinism, as generally understood within this debate, is the claim that everything that happens in the world—including all human thought and action—is subject to causal laws and that this involves the necessitation of effects by antecedent causal conditions.[2]

---

1 Page references for "Freedom and Resentment" are to pages 19–36 of this volume, followed by (after the slash) the corresponding pages of Gary Watson (ed.), *Free Will* (1982). (Thus, 'FR, 19/59' refers to this volume, p. 19/and *Free Will*, p. 59.)

2 On the standard empiricist account associated with classical compatibilism these laws are themselves analysed in terms of regularities or constant conjunctions of resembling events. However, this account of causal relations and causal laws is not essential to classical compatibilism.

The initial difficulty we are presented with is that free will and moral responsibility seem to be impossible if our actions are causally necessitated. At the same time, free will and moral responsibility also seem to be impossible if our actions are not causally necessitated. If both these claims are correct then free will and moral responsibility are impossible *whatever* view we take concerning the truth of determinism.

Why should the causal necessitation of action be viewed as incompatible with freedom and responsibility? One argument that classical incompatibilists have put forward is that moral responsibility requires free will and this requires that the agent could have acted otherwise or had alternative possibilities. However, if our actions are causally necessitated then the agent could not have done otherwise or had no open alternative possibilities. In these circumstances, the argument concludes, the agent is neither free nor responsible. On the other hand, when we move over to the side of indeterminism, we encounter a different set of problems. Classical compatibilists have argued that there is a difference between an action being free and an action being simply uncaused. If an action lacks any cause, they maintain, then it is merely a chance event that cannot be attributed to any agent. (For example, the random "swerves" of an Epicurean atom do not constitute freedom of the kind that is required for moral responsibility.) Clearly, then, denying the thesis of determinism serves only to present us with a different set of problems—those that are associated with the horn of chance.

The two main parties in the classical free will debate both attempt to avoid the skeptical conclusion reached by this argument. Classical compatibilism argues that it is a mistake to suppose that causal necessity is incompatible with freedom and responsibility. One core argument of this compatibilist position is that it is mere confusion to suppose that freedom implies the absence of causal necessitation. What freedom requires is only the absence of compulsion and coercion. Conduct is compelled or coerced when it is caused *in a certain way*: namely when it is produced without or against the agent's will and desires. However, when an action is caused by the agent's desires and willings then it is attributable to the agent and the agent can be said to act freely in the sense required for moral responsibility (Ayer, 1954; Schlick, 1930; Hobart, 1934; Nowell-Smith, 1948; Smart, 1963).[3] Responsible agents, compatibilists have argued, are those who can be appropriately influenced by the incentives of praise and blame and rewards and punishments. In circumstances where action is caused by the agent's own desires and willings we can influence the agent's future conduct by means of these incentives. The only sort of freedom required for moral responsibility, therefore, is the freedom to act according to the determination of our own will. From these arguments we can conclude, according to classical compatibilists, that the thesis of determinism poses no threat to moral responsibility.[4]

---

3 More recently, of course, compatibilists have introduced a number of fine-grained distinctions and more precise definitions concerning freedom, free will, acting freely on so on. However, for the purpose of describing classical compatibilist commitments these distinctions are not needed.

4 Many classical compatibilists also made it a central fixture of their positions to argue directly against the incompatibilist thesis that if determinism if true, no one can do otherwise. They argue that "could have done otherwise" claims should be analysed *conditionally* in terms of what the agent would have done *if* he had chosen or willed differently. (See, among others, Moore, 1912; Nowell-Smith, 1948; Ayer; 1954; and Smart, 1963). It is a curious fact about

Classical incompatibilists have objected to this strategy on several different grounds. One point of particular importance is that they reject the forward-looking, utilitarian understanding of moral responsibility provided by the compatibilist. We know that responsibility is not just a matter of our conduct being capable of being changed or influenced by rewards and punishments because it is evident that both children and animals can be influenced in these ways and they are, nevertheless, paradigmatically not responsible agents (Campbell, 1951). Clearly, then, something is missing from this classical compatibilist account of freedom and moral responsibility. What, then, is missing? According to incompatibilists what is missing is a sufficiently *deep* account of moral responsibility. More specifically, we must be able to account for the importance of moral *desert* and this requires that the agent is, in some relevant way, the *ultimate originator* of her conduct. For this to be possible, the incompatibilst continues, the agent must be able to choose between genuinely open alternatives and this requires the falsity of determinism. Some incompatibilists, known as *libertarians*, have argued that these conditions can be met (Campbell, 1951; Chisholm, 1964, 1967; and Taylor, 1954, 1974). Others, known as *hard determinists*, claim that this demand cannot be satisfied (Edwards, 1958; and Hospers, 1957).

The difficulty that libertarians have faced is to provide some account of free will that extends beyond the simple negative claim that our actions are not causally necessitated. The problem here is to give coherent content to some notion of free will that does not collapse into mere chance or capriciousness. According to classical compatibilists, when libertarians set about to meet this challenge their metaphysical commitments tend to take the form of speculative, anti-naturalistic accounts that are both obscure in themselves and difficult to integrate with the natural order of events in the world (Schlick, 1930; Hobart, 1934). It may be suggested, for example, that free will requires agents who are entirely distinct from their (given) character and desires and who are capable of a form of causality that is different in kind from (efficient) causation in nature.[5] From the classical compatibilist perspective, libertarian theories of this kind are hopelessly obscure and unintelligible.

The dynamics of the classical free will debate, as formulated along the lines of the dilemma of determinism, move us sharply in the direction of the conclusion that moral responsibility is *impossible*—since no plausible interpretation of free will can be provided by either of the two non-skeptical parties in this debate. The problem with this radical skeptical conclusion is that it seems both intellectually incredible and humanly impossible for us to accept or live with.

"Freedom and Resentment" that Strawson seemed not to concern himself with this central classical compatibilist topic. (For a reading of "Freedom and Resentment" that suggests otherwise, see McKenna, 2005; Scanlon, 1998, also suggests that Strawson did indeed have in his sights some treatment of "could have done otherwise", though Scanlon does not develop that point.)

5 One view of this kind is that agents are substances of a distinct kind or non-empirical rational selves who somehow directly bring about those events that are their actions. (See, among others, Campbell, 1951, Taylor, 1958; and Chisholm, 1964.)

II. Strawson's "Reconciling Project" and the Naturalistic Turn

Strawson presents his basic objective in "Freedom and Resentment" as an attempt to "reconcile" the two main parties in the free will dispute. For the purpose of his discussion he adopts his own labels for the various parties he is concerned with. This is both a strength and a weakness of his presentation. Strawson refers to compatibilists as "optimists", those who maintain that our attitudes and practices associated with moral responsibility would in no way be discredited by the thesis of determinism.[6] He refers to incompatibilists as "pessimists", those who believe that the truth of the thesis of determinism would discredit and undermine our commitment to these attitudes and practices. "Pessimists" may, on this account, be either libertarians or moral skeptics. The libertarian, as we have noted, holds that free will requires the falsity of determinism, that free will exists, and that free will serves to support a deep conception of moral responsibility. The moral skeptic is not only a "pessimist" on the assumption that determinism is true but also on the assumption that determinism is false. In other words, the moral skeptic is a *systematic pessimist* who concludes that the dilemma of determinism is intractable.[7] Although there is a sense in which Strawson's labels are confusing, they are nevertheless valuable in so far as they highlight the fact that the free will debate is not merely a matter of abstract, theoretical controversy. On the contrary, this is an issue that is taken to have a relevance to our attitude to life and the human condition, depending on where we stand on this matter.[8] More specifically, the pessimist/incompatibilist takes the view that there is something depressing or dispiriting about the (possible) truth of the thesis of determinism. In contrast with this, the optimist/compatibilist wants to show that there is no basis for worries and anxieties of this kind. Strawson's use of the labels "optimist" and "pessimist" are consistent with his wider objective in "Freedom and Resentment" to explain why the free will debate has the emotional significance for us that it has. It is Strawson's view that the main parties in this dispute have failed to identify what really matters to us in this sphere. His idiosyncratic terminology helps to illuminate and highlight this point.[9]

---

6 Strawson claims to be of "the party of those who do not know what the thesis of determinism is" (FR, 19/59). However, this does not prevent him from going on to characterize the dispute in terms of the general disagreement about the theoretical and practical implications of the truth of determinism.

7 Cp. Galen Strawson (1994).

8 Other great metaphysical questions, such as the existence of God and the immortality of the soul, may also be understood to have deeper significance for our attitude to life and the human condition (i.e. depending on what view we take on these issues). Certainly these are problems that are not merely of abstract, theoretical interest for us.

9 On Strawson's account the free will dispute is produced primarily by the confusions of philosophers. Ordinary life carries on unaffected and unconcerned by these (artificial) problems and difficulties. So considered, the free will dispute is presented as something of a "pseudo-problem", requiring diagnosis and philosophical therapy. Note, however, that this view sits uncomfortably with the view that the free will issue is one that affects our fundamental attitude to life and the human predicament (see the comments in note 8 above).

Strawson's attempt to "reconcile" the optimist and pessimist accepts that the pessimist is correct in holding that the optimist's account of moral responsibility leaves out "something vital". On the other hand, Strawson rejects the suggestion that what is needed is the falsity of determinism and some form of "contra-causal freedom" (FR, 35/79).[10] Although something is indeed missing from the optimist's story this gap is not to be filled using the "obscure and panicky metaphysics of libertarianism" (FR, 36/80). To find what is missing from the optimist and pessimist accounts we require a more radical shift in our philosophical approach to this problem. This involves what may be described as Strawson's "naturalistic turn".

Strawson's strategy in "Freedom and Resentment" involves turning away from conceptual issues about the analysis of "freedom" and "responsibility" and taking a closer look at what actually goes on when we *hold* a person responsible. That is to say, his methodology depends less on conceptual analysis and more on a descriptive account of actual human moral psychology. The place where our investigations ought to begin, Strawson suggests, is with "the very great importance that we attach to the attitudes and intentions towards us of other human beings, and the great extent to which our *personal feelings and reactions* depend upon, or involve, our beliefs about these attitudes and intentions" (FR, 22/62—our emphasis). Where we must begin, therefore, is with "that complicated web of attitudes and feelings which form moral life as we know it" (FR, 34/78). In this way, Strawson's "naturalistic turn" involves, not just a methodological turn away from conceptual analysis to moral psychology, but also an emphasis on the importance of emotion in moral life.[11] What is essential, on this view of things, is that we must not "over-intellectualize" the free will debate—a mistake that is common to both optimist and pessimist strategies (FR, 34–5/78–9).

How does an understanding of our "reactive attitudes and feelings" help us to resolve the free will problem? Strawson's approach is to describe carefully those circumstances in which we consider ourselves required to withdraw or suspend our reactive attitudes towards other individuals. He begins with an examination of the personal reactive attitude of resentment—where we believe that a person has injured or harmed us in some way or failed to show appropriate good will towards us (FR, 23/64).[12] In situations of this kind we are disposed "to modify or mollify" our feeling of resentment when certain relevant excusing considerations are brought to our attention. Strawson distinguishes between two important groups of excuses. The first aims to show that in the particular circumstances the agent's conduct lacked any degree of ill-will or disregard. Although some injury may have occurred, it was accidental, inadvertent or unintentional in some respect or other. Another kind of

10 See Campbell (1951). Campbell's paper is a reply to Nowell-Smith (1948), which Strawson cites as representative of the "optimist" viewpoint. Campbell's views are particularly representative of the (libertarian) "pessimist" outlook that Strawson is concerned with.

11 Strawson's turn away from conceptual analysis is probably one reason why he does not directly concern himself with the debate between his compatibilist and incompatibilist contemporaries over the proper analysis of "could have done otherwise". (See note 4.)

12 Strawson's interest in the case of resentment and its relation to our reactive attitudes is anticipated by Adam Smith in his *Theory of Moral Sentiments* (1759). See also Joseph Butler, "On Resentment" (1726).

excusing consideration goes much further than this and suggests that resentment is not called for on the ground that the agent is somehow an "inappropriate" target of any attitude of resentment because she is "abnormal or immature" and thus not a normal adult who we can reasonably expect to show due care and concern for others.

This analysis of the rationale of excuses as they concern our reactive attitudes suggests that there are two different stances that we can take up in our "human relationships". The first is a "participant" attitude where we believe the person we are dealing with is generally a normal adult who is an appropriate target of our reactive attitudes. The other is the "objective" stance, where we believe that the person we are dealing with is in some way incapacitated for normal adult human relationships and so reactive attitudes are inappropriate when directed at individuals of this kind. In cases of this kind we see the person as "an object of social policy; as a subject for what, in a wide range of sense, might be called treatment … to be managed or handled or cured or trained" (FR, 25/66). In general, however, there is no question of us choosing or needing to justify the fact that we are liable to reactive attitudes and feelings in the normal case. This is simply a "given" of our human nature and without it we would hardly recognize an individual or community as being fully *human*.

Strawson maintains that these observations about human psychology and society are fundamental for understanding what has gone wrong with the classical free will debate. What influence, he asks, would the truth of determinism have upon our personal reactive attitudes such as resentment? More specifically, would the acceptance of this thesis "lead to the repudiation of all such attitudes"? (FR, 25/67). Strawson answers this question at several different levels. In the first place he argues that the truth of determinism in no way serves (theoretically) to discredit our reactive attitudes in any systematic way. For this to be so, he claims, determinism would have to imply that one or other of the two basic forms of excusing considerations hold universally. We have, however, no reason to suppose that this follows. Certainly determinism does not imply that any injury caused by someone is always done accidentally or inadvertently. That is to say, determinism does not imply that no one's conduct ever manifests ill will or lacks proper regard for others. Nor does determinism imply that every agent is somehow abnormal or immature, or in some way incapacitated for adult relationships. We may conclude, therefore, that the truth of the thesis of determinism in no way discredits or theoretically undermines our commitment to reactive attitudes of this kind.

Strawson argues for two further points beyond this. Even if we had some theoretical reason to abandon or suspend these reactive attitudes it would be psychologically impossible for us to do this. To do this would involve "adopting a thoroughgoing objectivity of attitude to others" which is something we are *incapable* of (FR, 27–8/69–70). While it is possible for us to choose to take up the objective stance on occasion—even sometimes when we are dealing with normal, mature adults—the fact is that we cannot do this with regard to all people all of the time *whatever* theoretical view we may hold. Furthermore, even if we were to be given a ("god-like") choice on this matter and could decide for ourselves whether or not to retain "our natural human commitment to ordinary inter-personal attitudes", this choice

must be decided in terms of "the gains and losses to human life, its enrichment or impoverishment" (FR, 28, 31–2/70, 74). Obviously we have little or no reason to choose a life that is emotionally impoverished and de-humanized, so the truth of determinism is itself irrelevant to any choice of this kind.

According to Strawson the very same arguments and considerations apply to our moral reactive attitudes, which are simply "generalized or vicarious analogues of the personal reactive attitudes" (FR, 29/71). These various forms of reactive attitude, Strawson maintains, are all "humanly connected" and subject to the same rational and psychological constraints. Clearly, then, the truth of the thesis of determinism provides us with no theoretical reason to entirely suspend or abandon our commitment to our (moral) reactive attitudes and we have, furthermore, no pragmatic reason to *want* to take an exclusively objective stance toward all other people. More importantly, whatever our theoretical or pragmatic views may be, any effort of this kind is "practically inconceivable", since our commitment to moral reactive attitudes is no less fundamental to human nature than it is in the case of other forms of reactive attitudes (FR, 31–2/74–5).[13]

Strawson draws several important conclusions from these arguments. In the first place, he identifies the respective failings of both the optimist and pessimist positions. The optimist generally attempts to show that the truth of determinism does not prevent rewards and punishments from "regulating behaviour in socially desirable ways".

> The picture painted by the optimist is painted in a style appropriate to a situation envisaged as wholly dominated by objectivity of attitude. The only operative notions involved in this picture are such as those of policy, treatment, control. But a thoroughgoing objectivity of attitude, excluding as it does the moral reactive attitudes, excludes at the same time essential elements in the concepts of *moral* condemnation and *moral* responsibility. (FR, 33/76—Strawson's emphasis)

Strawson argues that the pessimist is right to "recoil" at this picture of things but makes the mistake of concluding that the reactive attitudes cannot be all that we require to "fill the gap in the optimist's account" (FR, 35/79). This leads the pessimist to conclude that "the gap can be filled only if some general metaphysical proposition is repeatedly verified, verified in all cases where it is appropriate to attribute moral responsibility" (FR, 35/79). This proposition, Strawson claims, "is as difficult to state coherently and with intelligible relevance as its determinist contradictory" (FR, 35/79).

Behind their disagreement both the optimist and pessimist are guilty of a shared misunderstanding.

> Both seek, in different ways, to over-intellectualize the facts. Inside the general structure or web of human attitudes and feelings of which I have been speaking, there is endless room for modification, redirection, criticism, and justification. But questions of justification

13 Along with the personal and moral reactive attitudes Strawson also distinguishes a third group, the self-reactive attitudes. These are "associated with demands on oneself for others" (FR, 29/71).

> are internal to the structure or relate to modifications internal to it. The existence of the general framework of attitudes itself is something we are given with the fact of human society. As a whole, it neither calls for, nor permits, an external 'rational' justification. Pessimist and optimist alike show themselves, in different ways, unable to accept this. (FR, 35/78–9)

With this point in place Strawson goes on to conclude that "if we sufficiently, that is radically, modify the view of the optimist, his view is the right one" (FR, 36/80). What needs to be done is to fill the "lacuna" in the optimist's account—the neglect of our human attitudes and feelings—without falling into the "obscure and panicky metaphysics of libertarianism" (FR, 36/80). We must ask the pessimist to "surrender his metaphysics" and the optimist to concede that there is "something vital" missing in his account. If both parties can agree to these terms then "reconciliation" will be achieved.

Strawson's naturalistic turn in "Freedom and Resentment" has two important features and each draws on (related) historical sources. The first is the role of *emotion* in Strawson's account of the nature and conditions of moral responsibility. Strawson makes clear that his interest in reactive attitudes constitutes a return to the moral sense tradition in ethical theory.[14] Although Strawson does not cite any specific influences, several key claims and insights in his paper can be found (in a more fully worked out form) in the writings of predecessors such as David Hume (Hume, 1739–40) and Adam Smith (Smith, 1759). The most fundamental of these claims is that our moral reactions to other human beings is a given of our human nature and involve feelings and attitudes produced by our beliefs about conduct and character. It is, nevertheless, a significant fact that the moral sense tradition is itself divided on the free will issue.[15]

Another important dimension of Strawson's "naturalistic turn" is his way of discrediting the skeptical challenge by reference to the inescapable, psychological mechanisms that guide human thought and action. This is a theme that Strawson directly addresses in his more recent work *Skepticism and Naturalism* (hereafter abbreviated as SN), where he explicitly acknowledges both his historical sources and the wider application of this form of naturalism to philosophical skepticism. In this context, Strawson observes that Thomas Reid drew "an explicit parallel between our natural commitments to belief in external things and our natural proneness to moral or quasi-moral response" (SN, 33). In this respect, says Strawson: "we see Reid aligning himself with Hume the naturalist against Hume the skeptic" (SN, 33). He elaborates on this theme further below:

14 "It is a pity that talk of the moral sentiments has fallen out of favour …" (FR, 35/79).

15 Hume, for example, is clearly an optimist/compatibilist. In contrast with this, Butler defends a pessimist/libertarian position [*Analogy of Religion*, Pt. I, Ch. 4.]. Smith is simply silent on the problem of free will. On the striking resemblance between Hume's and Strawson's strategy as it concerns moral responsibility and the free will problem see, Russell (1995), esp. Ch. 5. For a more general discussion of Strawson's views as they relate to the compatibilist tradition see McKenna (2004).

> [Reid's point …] is that argument, reasonings, either for or against the skeptical position, are, in practice, equally inefficacious and idle … Where Nature thus determines us, we have 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 rationality or irrationality, justification or lack of justification, of this or that particular judgment or belief can come up. I then played roughly the same game, as one might put it, with the moral life. We are naturally social beings; and given with our natural commitment to social existence is a natural commitment to that whole web or structure of human personal and moral reactive attitudes, feelings, and judgments of which I spoke. Our natural disposition to such attitudes and judgments is naturally secured against arguments suggesting that they are in principle unwarranted or unjustified just as our natural disposition to belief in the existence of body is naturally secured against arguments suggesting that it is in principle uncertain. (SN, 39)

Strawson's "naturalist way" with the moral skeptic not only parallels the naturalist way with the skeptic about the existence of the external world, it also parallels (Hume's) naturalist way of dealing with skepticism about induction.[16] Clearly, then, as Strawson indicates, his own naturalistic way of refuting pessimism and moral skepticism, as it arises out of the free will dispute, has parallels with the way that Hume, Reid and others have dealt with structurally similar skeptical challenges.[17]

III. Critical Themes Concerning "Freedom and Resentment"

At the time "Freedom and Resentment" first appeared Anglo-American philosophy was tightly in the grip of "ordinary language" or "conceptual" analysis. As we have noted, around the middle of the twentieth century it was widely held that the right way to solve the free will problem was to provide the correct analysis of the logical relations among the terms "freedom", "causation", "responsibility" and "could have done otherwise". Lying behind this approach is the assumption that there is some identifiable, determinate meaning to each of the relevant terms and that this procedure of "conceptual analysis" could put an end to the whole controversy. Related to this assumption is the view that language embodies all the logical distinctions that are needed for everyday moral conversation and practice. Confusion in this sphere, therefore, is a peculiar product philosophical reflection rather than a real difficulty that presents itself in ordinary moral life.

It is evident that Strawson's naturalistic approach to the free will problem involves a fundamental shift away from this kind of philosophical methodology. However, although Strawson's naturalistic turn recommends that we begin with the

---

16 SN, 18–19; and FR, 35n7/79n7. In the latter passage Strawson suggests that any attempt to justify the general framework of our reactive attidudes is misguided in much the same way that any attempt to justify induction is mistaken. For Strawson's views on induction see Strawson (1952), Ch. 9.

17 According to Strawson, Hume's naturalist way of dealing with skepticism finds a "powerful later-day exponent of a closely related position" in Wittgenstein (SN, 10,14f). This indicates that, in some respects, Strawson's own strategy in "Freedom and Resentment" takes a Wittgensteinian approach.

facts of human moral psychology, as opposed to a pure conceptual analysis, there are features of his methodology that follow the ordinary language approach in a number of ways. It may be argued, for example, that Strawson's naturalistic commitments in "Freedom and Resentment" manifest his own preference for "descriptive" over "revisionary" metaphysics. This distinction is introduced by Strawson in *Individuals*, which was published three years before "Freedom and Resentment".

> Descriptive metaphysics is content to describe the actual structure of our thought about the world, revisionary metaphysics is concerned to produce a better structure … The idea of descriptive metaphysics is liable to be met with skepticism. How should it differ from what is called philosophical, or logical, or conceptual analysis? It does not differ in kind of intention, but only in scope or generality. Aiming to lay bare the most general features of our conceptual structure, it can take far less for granted than a more limited and partial conceptual inquiry. [Strawson (1959), p. 9]

Strawson goes on to say that while metaphysics may be an instrument of conceptual change, there is, nevertheless, "a massive central core of human thinking which has no history … there are categories and concepts which, in their most fundamental character, change not at all" [Strawson (1959), p. 10]. Strawson also makes clear that his work in *Individuals* is intended as a contribution to descriptive as opposed to revisionary metaphysics. The same "descriptive" spirit is evident in "Freedom and Resentment". Strawson wants to account for the *natural* foundation of our "concepts and practices associated with moral responsibility" in the "web" of human reactive attitudes and feelings. Although he does not deny that there are "local and temporary" variations in the way our moral sentiments and reactive attitudes may be manifest and directed. (FR, 36/80) he insists, nevertheless, that "in the absence of *any* forms of these attitudes it is doubtful whether we should have anything that *we* could find intelligible as a system of human relationships, as a human society". (FR, 36/80, Strawson's emphasis). This aspect of his approach in "Freedom and Resentment" is entirely consistent with his more general descriptive commitments and orientation.

These features of Strawson's naturalism and descriptivist metaphysics may be challenged from two related points of view. In the first place, it may be argued that our human commitment to the framework of reactive attitudes or moral sentiments is not so essential to human nature as Strawson suggests.[18] The reactive attitudes are not permanent, inescapable features of human nature, like other more basic emotions such as fear or love. Rather they should be viewed as cultural artifacts produced in particular historical and social conditions. Reactive attitudes, the critic maintains, involve a set of socially acquired dispositions and expectations and we can well imagine more radical shifts and transformations in human relationships and attitudes than Strawson allows for. Related to this point, it may also be argued that Strawson's form of descriptivist metaphysics and naturalism commits him to a "static" and "conservative" account of our basic concepts and practices as they relate to moral responsibility. Although Strawson allows that there is "endless room for modification, redirection, criticism and justification", he pointedly blocks-off

---

18 See, for example, Watson (1987; reprinted in this volume, pp. 117–43); or Wallace (1994; excerpt reprinted in this volume, pp. 159–87).

any radical revisionist account of responsibility that would take us in the direction of a purely utilitarian or consequentialist theory (FR, 34–5/78–9). The critic may argue that our attitudes and practices in this sphere are more flexible and plastic than Strawson's picture suggests and that our theoretical reflections may push us in the direction of more fundamental changes of a "revisionist" kind.[19]

A central theme of "Freedom and Resentment" is that both optimists and pessimists seek to "over-intellectualize the facts" as they relate to justifying responsibility. More specifically, it is a mistake, Strawson argues, to aim to provide any "external 'rational' justification" for responsibility. Beyond this, however, Strawson is clear that reason has an important role to play "inside the general structure or web of human attitudes and feelings" that he is concerned with. Be this as it may, it remains a matter of some debate about the exact role that reason can play in justifying our reactive attitudes. Here too, we encounter two related difficulties. The first concerns Strawson's understanding of the cognitive aspects of emotion. Should moral sentiments or reactive attitudes be interpreted as mere feelings, without propositional content, or should they be regarded as necessarily involving beliefs of some relevant kind?[20] Similarly, what role does reason play in justifying the (moral) demands and standards associated with the reactive attitudes and feelings?[21] Are these demands and standards part of the "given framework" of human nature or do they vary with local conventions and customs? The deeper issue here concerns the open or noncommittal nature of Strawson's theoretical commitments as they relate to moral reasons and justification *within* the framework of reactive attitudes.[22]

Strawson's discussion in "Freedom and Resentment" turns on several sharp dichotomies that may also be questioned. Among these is the fundamental opposition between the "objective" and "participant" stances. A number of problems present themselves in relation to this dichotomy. It may be suggested, for example, that we may altogether avoid reactive attitudes toward an individual (e.g. on the ground that they are severely incapacitated) but still engage in other forms of personal, emotional response or engagement (e.g. parental love). Another general difficulty relating to this dichotomy is how it is possible to "dispel" or set-aside our reactive attitudes when we take up the objective stance in the "normal" case without thereby discrediting them? In other words, it may be argued that the constraints imposed by embracing the objective stance as an acceptable option for inter-personal relations

19 For a strong argument for this point see Pereboom (2001). For an interesting way to *support* Strawson along revisionist lines, see Vargas (2004; reprinted in this volume, pp. 301–20).

20 See, for example, Bennett (1980; see also "Accountability (II)" at pp. 47–69 of this volume); Wallace (1994; excerpt reprinted in this volume, pp. 159–87). The background worry here is that Strawson's account depends on a feeling theory of emotion that would make it impossible to assess (tokens of) emotions or reactive attitudes and feelings as themselves reasonable or unreasonable. This difficulty has deeper roots in the moral sense tradition dating back to Hume's views on this subject. On this issue see Russell (1995), Ch. 6.

21 On this point see Ayer (1980; reprinted in this volume, pp. 37–46) and Wolf (1981; reprinted in this volume, pp. 71–85).

22 For differing positions on this issue, see Magill (1997; and see the excerpt reprinted in this volume, pp. 189–202) and Kelly (2002; reprinted in this volume, 275–98).

are such that it is not possible to continue to endorse the reactive attitudes as in any way appropriate in these circumstances. If these observations are correct then the objective/participant dichotomy is more problematic for Strawson's project than he has acknowledged.[23]

Another dichotomy that that has attracted critical comment concerns the gap between holding and being responsible. Critics, such as John Martin Fischer and Mark Ravizza, have argued that Strawson is insufficiently sensitive to this gap and that this vitiates much of his strategy in "Freedom and Resentment".

> … Strawson's theory may reasonably be said to give an account of what it is for agents to be held responsible, but there seems to be a difference between being *held* responsible and actually *being* responsible. Surely it is possible that one can be held responsible even though one in fact is not responsible, and conversely that one can be responsible even though one is actually not treated as a responsible agent. By understanding responsibility primarily in terms of our actual practices of adopting or not adopting certain attitudes towards agents Strawson's theory risks blurring the difference between these two issues. [Fischer and Ravizza (1993), p.18—their emphasis]

It is important to be careful about interpreting the exact nature of this criticism. It is clear that Strawson wants to allow that in *particular* cases our reactive attitudes may fail to properly track moral responsibility (e.g. we feel an inappropriate reactive attitude because we have incorrect beliefs about the agent's intentions or state of mind). Similarly, sometimes we may fail to feel any sentiment of approval or disapproval but would do so if we were fully and properly informed about the agent's conduct and character. These are not, however, the sort of difficulties that concern Fischer and Ravizza.

The problem that they are concerned with is the possibility that there could be a *systematic* lack of correlation between our reactive attitudes and their appropriate and legitimate objects. It could be, for example, that an entire community has its reactive attitudes switched on or off in the wrong way and at the wrong times. The very possibility of this suggests that there is more to being responsible than what is generally targeted by our reactive attitudes and feelings. A deeper difficulty, related to this, is that what does or does not make an individual an appropriate target of reactive attitudes must depend on some relevant interpretation of *moral capacity*—the capacities that establish a person as a "normal adult" who is capable of full participation in the moral community.[24] This is a crucial issue, since the pessimist/incompatibilist will argue that among the relevant capacities we must consider is the ability to act otherwise or (libertarian) free will. Unless Strawson can provide some

---

23 On these matters, see G. Strawson (1986; see also excerpt reprinted herein, pp. 87–115); Smilansky (2000; reprinted in this volume, pp. 237–56), (2001); and Zimmerman (2001; reprinted herein, pp. 257–74).

24 Several commentators have fixed upon this point. See, for example, Nagel (1986), Ch. 7; Watson (1987; reprinted in this volume, pp. 117–43); Russell (1992; reprinted, this volume, pp. 145–58), (1995); McKenna (1998; reprinted, this volume, pp. 203–20); and Scanlon (1998).

plausible account of moral capacity then it would appear there is a significant gap or "lacuna" in his own version of compatibilism.

Another problem for Strawson's theory, critics may argue, involves his ambiguous remarks about the relationship between resentment and retribution. Although Strawson maintains that "savage or civilized, we have some belief in the utility of [retributive] practices" (FR, 34/78) this leaves it open as to what role utilitarian considerations play in justifying punishment in general, as well as particular cases of punishment.[25] It is not clear, for example, whether the institution of punishment requires some "external 'rational' justification"—in contrast with the whole framework of reactive attitudes that serve to ground and structure retributive practices and the institutions associated with them.[26] Nor is it clear if our retributive practices, as distinct from our reactive attitudes which motivate them, are a "given" of human nature and something that it is practically inconceivable that we could altogether suspend or abandon.[27] Obviously it is one thing to claim that our emotional nature is fixed and incapable of fundamental alteration and quite another to say that the institutions and practices associated with them are also incapable of being eliminated or removed from human life. Certainly this is an issue that requires further interpretation and analysis. What is needed, therefore, is a more detailed account of the implications of Strawson's naturalistic account of responsibility for a theory of retributive justice.[28]

Perhaps the most fundamental problem with Strawson's theory in the eyes of his critics is that he has too little to say about the extent to which our reflections concerning the (historical) origins or sources of character and conduct can inhibit—if not altogether undermine—the (normal) operation of our reactive attitudes. The difficulty here can be brought to light by means of "implantation" cases. Suppose that our basic dispositions, which shape and direct our attitudes and intentions towards others, were somehow "implanted" by means of an artificial technique of some kind (e.g. neuro-surgery or genetic engineering). According to Strawson' theory, our reactive attitudes would still continue to operate in cases of this kind. That is to say, allowing that the agent is a "normal" (i.e. rational) adult who is capable of manifesting good or ill will towards others "theoretical" worries about implantation will not and cannot dislodge or discredit our reactive attitudes. Contrary to this view, however, critics will argue that implantation evidently eliminates the agent's moral responsibility and so there is something wrong with Strawson's theory.[29]

---

25 For a discussion of the general utilitarian justification for the reactive emotions see Magill (1997; excerpt reprinted in this volume, pp. 189–202) and Zimmerman (2001; reprinted, this volume, pp. 257–74).

26 On this see Russell (1995), Ch. 10.

27 See, for example, Mackie (1982). Mackie shows how naturalistic principles of a Strawsonian kind may be used to defend strong forms of retributivist punishment.

28 It is beyond the scope of this collection to take up the relevance of Strawson's views in "Freedom and Resentment" for the related topic of *distributive* justice. See, however, the valuable discussion in Scheffler (1992).

29 Similar difficulties arise in the theological context regarding worries about moral responsibility in circumstances where God is the ultimate source of the agent's character and conduct. (See Hume's influential discussion of this issue in his *Enquiry Concerning Human Understanding*, Sect. VIII.)

This line of criticism, based on worries about the history, origins or source of character and conduct, cannot be evaded by insisting that implantation is "abnormal" and that this is what excuses the agent. We may, for example, alter the factors at work in conditioning the agent's basic character and disposition and make them natural and/or ordinary social processes of various kinds.[30] In these circumstances the agent still has no control over the kind of moral character she has acquired and it is *this*—not implantation *per se*—that makes it impossible to sustain our commitment to reactive attitudes. Strawson's theory, the critic continues, simply presupposes that (historical) considerations of this kind are irrelevant to the legitimacy and functioning of our reactive attitudes. Most philosophers maintain that this view of things is neither psychologically credible nor philosophically defensible.[31] The fact that Strawson has so little to say about this general problem suggests, to some readers, that Strawson's naturalistic approach never adequately confronts the real problem lying at the heart of the free will dispute and thus fails to solve it.

The themes and criticisms described above are all raised in one form or another in the essays that follow. For the purpose of our introduction there is no need to summarize or paraphrase each individual essay. It is better for the contributors to speak for themselves in their own words.

IV. The Aim of This Volume

All the essays in this collection make a significant contribution to the interpretation and criticism of the strategy and arguments that are advanced in "Freedom and Resentment". Taken as a group, they constitute an important set of developments in the contemporary free will debate. The primary purpose of bringing this material together is to give our readers easy access to these contributions, which in some cases are not easily found. There are also several essays in this collection that present new material that has not appeared in print before or else substantially revise earlier work. This collection provides an opportunity to contrast and compare the various responses to "Freedom and Resentment" in a unified format that allows the reader to consider these responses and criticisms in a more systematic and coherent fashion. This process, we hope, will stimulate new avenues of criticism and encourage further development of the themes arising from Strawson's work. Finally, as with any project of this kind, due to limited space and publication costs we have not been able to include every relevant and worthwhile contribution that has already appeared in print. Several of these works are cited in the "suggestions for further reading" that appears at the end of this volume. The suggestions that we have made are not intended to cover all material relevant to "Freedom and Resentment" published over the past four decades. Our aim is simply to guide the reader to a few more notable contributions that are directly relevant to the particular discussions and arguments taken up in this volume.

---

30 This is one crucial point in Watson (1987; this volume, pp. 117–43).

31 For one noteworthy source of dissent, see Frankfurt (1975). In Strawson's defence, McKenna (1998) resists Watson's historical objections to Strawsonian compatibilism.

References

Ayer, A.J., “Freedom and Necessity”, in *Philosophical Essays* (New York: St Martin’s Press, 1954), pp. 3–20.

—— “Free Will and Rationality”, in van Straaten (ed.), *Philosophical Subjects* (1980).

Bennett, Jonathan, “Accountability”, in van Straaten (ed.), *Philosophical Subjects* (1980).

Berofsky, Bernard (ed.), *Free Will and Determinism* (New York: Harper & Row, 1966).

Butler, Joseph , “On Resentment” [Sermon 8], in *The Works of Joseph Butler*, Vol. 2.

—— *The Analogy of Religion*, in *The Works of Joseph Butler*, Vol. 1.

—— *The Works of Joseph Butler*, 2 vols, Samuel Halifax (ed.), (Oxford: Oxford University Press, 1849).

Campbell, C.A., “Is ‘Freewill’ a Pseudo-Problem?”, *Mind*, 60 (1951), 446–65. Reprinted in Bernard Berofsky, (ed.), *Free Will and Determinism*.

Chisholm, Roderick, “Human Freedom and the Self”, *The Lindley Lectures* (Department of Philosophy, University of Kansas, 1964).

—— “He Could Have Done Otherwise”, *Journal of Philosophy*, 64 (1967), 409–18.

Edwards, Paul, “Hard and Soft Determinism”, in Hook (ed.), *Determinism and Freedom* (1958).

Fischer, John M., “Responsibility and Control”, *Journal of Philosophy*, 89 (1982), 24–40.

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

Frankfurt, Harry, “Alternate Possibilities and Moral Responsibility”, *Journal of Philosophy*, 68 (1969), 829–39.

—— “Freedom of the Will and the Concept of a Person”, *Journal of Philosophy*, 68 (1971), 5–20.

—— “Three Concepts of Free Action”, *Proceedings of the Aristotelian Society*, supp. vol. 2 (1975), 113–25.

Ginet, Carl, “Might We Have No Choice?” in Lehrer (ed.), *Freedom and Determinism* (1966).

—— *On Action* (Cambridge: Cambridge University Press, 1990).

Hobart, R.E., “Free Will as Involving Determination and Inconceivable Without It”, *Mind*, 43 (1934), 1–27.

Hook, Sidney (ed.), *Determinism and Freedom in the Age of Modern Science* (New York: MacMillan, 1958).

Hume, David, *A Treatise of Human Nature* [1739–40], David Fate Norton and Mary J. Norton (eds), (Oxford: Oxford University Press, 2000).

—— *An Enquiry Concerning Human Understanding* [1748], Tom L. Beauchamp (ed.), (Oxford: Oxford University Press, 2000).

Kane, Robert, *The Significance of Free Will* (Oxford: Oxford University Press, 1996).

Kelly, Erin, “Doing Without Desert”, *Pacific Philosophical Quarterly*, 83 (2002), 180–205.

Lehrer, Keith (ed.), *Freedom and Determinism* (New York: Random House, 1966).

Mackie, John, “Morality and the Retributive Emotions” [1988], reprinted in Mackie, *Persons and Values: Selected Papers Volume II* (Oxford: Oxford University Press, 1985), pp. 206–19.

Magill, Kevin, *Experience and Freedom: Self-Determination Without Illusions* (London: MacMillan, 1997).

McKenna, Michael, “The Limits of Evil and the Role of Moral Address: A Defense of Strawsonian Compatibilism”, *Journal of Ethics*, 2 (1998), 123–42.

—— “Compatibilism”, *The Stanford Encyclopedia of Philosophy* (Summer 2004 Edition), Edward N. Zalta (ed.), http://plato.stanford.edu/archives/sum2004/entries/compatibilism.

—— “Where Strawson and Frankfurt Meet”, *Midwest Studies in Philosophy*, 29 (2005), 163–80.

Moore, G.E., *Ethics* (Oxford: Oxford University Press, 1912).

Nagel, Thomas, *The View From Nowhere* (Oxford: Oxford University Press, 1986).

Nowell-Smith, H.P., “Free Will and Moral Responsibility”, *Mind*, 57 (1948), 45–61.

—— “Ifs and Cans”, *Theoria*, 26 (1960), 85–101.

Pereboom, Derk, “Alternate Possibilities and Causal Histories”, *Philosophical Perspectives*, 14 (2000), 119–38.

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

Russell, Paul, “Strawson’s Way of Naturalizing Responsibility”, *Ethics*, 102 (1992), 287–302.

—— *Freedom and Moral Sentiment: Hume’s Way of Naturalizing Responsibility* (Oxford: Oxford University Press, 1995).

Scanlon, T.M., “The Significance of Choice”, in Sterling M. McMurrin (ed.), *The Tanner Lectures on Human Values* (Salt Lake City: University of Utah Press, 1998).

Scheffler, Samuel, “Responsibility, Reactive Attitudes, and Liberalism”, *Philosophy and Public Affairs*, 21 (1992), 299–323.

Schlick, Moritz, “When Is a Man Responsible?” reprinted in Berofsky (ed.), *Free Will and Determinism* (1966).

Schoeman, Ferdinand (ed.), *Responsibility, Character, and the Emotions: New Essays in Moral Psychology* (Cambridge: Cambridge University Press, 1987).

Smart, J.J.C., “Free Will, Praise, and Blame”, *Mind*, 70 (1963), 291–306.

Smilansky, Saul, *Free Will and Illusion* (Oxford: Oxford University Press, 2000).

—— “Free Will: From Nature to Illusion”, *Proceedings of the Aristotelian Society*, 101 (2001), 71–95.

Smith, Adam, *The Theory Of Moral Sentiments* [1759], Knud Haakonssen (ed.), (Cambridge: Cambridge University Press, 2002).

Strawson, Galen, *Freedom and Belief* (Oxford: Oxford University Press, 1986).

Strawson, P.F., *Introduction to Logical Theory* (London: Methuen, 1952).

—— *Individuals: An Essay in Descriptive Metaphysics* (London: Methuen, 1959).

—— "Freedom and Resentment", *Proceedings of the British Academy*, 48 (1962), 187–211. Reprinted in Strawson (1974) and in Watson (1982).
—— *Freedom and Resentment and Other Essays* (London: Methuen, 1974).
—— *Skepticism and Naturalism* (New York: Columbia University Press, 1985).
Taylor, Richard, "Determinism and the Theory of Agency" in Hook (ed.), *Determinism and Freedom in the Age of Modern Science* (1958).
—— *Metaphysics* (Englewood Cliffs, NJ: Prentice Hall, 1974).
van Straaten, Zak (ed.), *Philosophical Subjects: Essays Presented to P. F. Strawson* (Oxford: Oxford University Press, 1980).
Vargas, Manuel, "Responsibility and the Aims of Theory: Strawson and Revisionism", *Pacific Philosophical Quarterly*, 85 (2004), 218–41.
Wallace, R. Jay, *Responsibility and the Moral Sentiments* (Cambridge MA: Harvard University Press, 1994).
Watson, Gary (ed.), *Free Will* (Oxford: Oxford University Press, 1982).
—— "Responsibility and the Limits of Evil: Variations on a Strawsonian Theme", in Schoeman (ed.), *Responsibility, Character, and the Emotions* (1987).
Wolf, Susan, "The Importance of Free Will", *Mind*, 90 (1981), 386–405.
Zimmerman, David, "Thinking With Your Hypothalamus: Reflections on a Cognitive Role for the Reactive Emotions", *Philosophy and Phenomenological Research*, 63 (2001), 521–41.