Taking the Straight Path:

P. F. Strawson's Later Work on

Freedom and Responsibility

Benjamin De Mesel

Institute of Philosophy, KU Leuven

P. F. Strawson's 'Freedom and Resentment' (2008 [1962]) plays a central role in responsibility theory, but it is barely known that Strawson wrote about freedom and responsibility in other works and that he explicitly rejected a central argument put forward in 'Freedom and Resentment'. Strawson's reply to Jonathan Bennett (1980; Strawson 1980) and his book *Scepticism and Naturalism* (1985a) are relatively well known, but replies to Simon Blackburn (1998; Strawson 1998a), David Pears (1998; Strawson 1998b), Ernest Sosa (1998; Strawson 1998c) and, especially, Rajendra Prasad (1995; Strawson 1995) have been virtually ignored. The same can be said about his book chapters 'Liberty and Necessity' (Strawson 2011c [1983]) and 'Freedom and Necessity' (Strawson 1992b).[1]

My primary aims are (1) to present some of Strawson's later work on freedom and responsibility, (2) to show how it sheds light on 'Freedom and Resentment' (despite appearing to contain claims that are incompatible with it), and (3) to identify problems with interpretations of and objections to 'Freedom and Resentment' in the light of Strawson's later remarks.

I will start, in section 1, with Strawson's reply to Prasad (1995). Prasad argues that one of Strawson's arguments in 'Freedom and Resentment' against the relevance of determinism to moral responsibility is flawed. In his reply, Strawson straightforwardly accepts Prasad's criticism and repudiates his earlier argument. In section 2, I will show how Strawson's remarkable acceptance of Prasad's criticism, closely related to a criticism later developed by John Martin Fischer (2014), motivates him to advance a direct argument for the claim that determinism is compatible with freedom and responsibility. The later Strawson is clearly (and more unambiguously than in 'Freedom and

---

1. These book chapters are almost identical. 'Liberty and Necessity' was first published in a 1983 collection on Spinoza and republished in *Philosophical Writings* (2011). A slightly altered version, 'Freedom and Necessity', became part of a series of lectures delivered in the Collège de France in the spring of 1985 and published later that year in French under the title *Analyse et métaphysique*. The lectures were published in English in 1992 under the title *Analysis and Metaphysics*. For more on the history of the lectures, see Strawson (1992a: viii). I would like to thank an anonymous reviewer for pressing me to clarify this.

Resentment') a classical compatibilist. He claims that the ability to do otherwise is a necessary condition of responsibility and provides a list of additional conditions, including a knowledge condition. Although this definitely involves a shift of emphasis when compared to 'Freedom and Resentment', I will argue that there is no need to see a major change in Strawson's thought.[2] Strawson's later remarks are compatible with the thoughts expressed in 'Freedom and Resentment', and they do not detract from its significance.

Strawson's later emphasis on the conditions of responsibility, including a knowledge condition and the ability to do otherwise, raises questions about the place of the reactive attitudes in his account, which I will discuss in section 3. Drawing upon Strawson's reply to Pears (1998), I will highlight the importance of quality of will and introduce a key distinction between two questions: 'How does the concept of responsibility operate?' and 'Why do we have the concept of responsibility?' I will argue that the reactive attitudes, lying at the *origin* of our distinction between responsible and non-responsible agents, figure essentially in Strawson's answer to the second rather than the first question, and I will indicate how Strawson's position relates to response-dependent and response-independent accounts of responsibility.

1. Prasad and Strawson on Determinism and Responsibility

1.1 Prasad's Criticism

In 1995, the Indian Council of Philosophical Research published a collection of essays on P.F. Strawson's work. *The Philosophy of P.F. Strawson* (1995), edited by Pranab Kumar Sen and Roop Rekha Verma, has been

---

2. My distinction between Strawson's earlier and later work is based solely on publication dates; it is not meant to suggest that Strawson's views on freedom and responsibility underwent major changes. 'Freedom and Resentment', published in 1962, is earlier work; the other works I discuss, written in the 1980s or later, are later works. Campbell (2017: 30) also thinks that Strawson's later work 'seeks to clarify and expand on the view given in his previous work', but he does not discuss Strawson's later emphasis on the conditions of responsibility, including the freedom to do otherwise, which is central to my paper.

virtually overlooked in the literature on Strawson, in contrast to a later volume with the same title edited by Lewis Hahn (1998).[3] The collection edited by Sen and Verma contains an introduction by Strawson ('My Philosophy'), a bibliography of his work, and sixteen essays with replies. The essays cover the whole of Strawson's work. Among the authors are Michael Dummett, Hilary Putnam, Quassim Cassam, Michael Luntley, Akeel Bilgrami and Mark Platts.

Prasad's 'Reactive Attitudes, Rationality and Determinism' (1995) is a rich commentary on 'Freedom and Resentment' to which I cannot do full justice here. I will focus on one of its criticisms. Prasad notes that Strawson lists three types of conditions which make it inappropriate to have a reactive attitude. 'The first group [known in the literature as the group of excuses] I shall abbreviate as the indeliberateness of the action, the first sub-group of the second [known as the first sub-group of exemptions] as the unusual impairment of the agent, and its second sub-group [the second sub-group of exemptions] as the agent's immunity from moral appraisal' (Prasad 1995: 353). According to Strawson in 'Freedom and Resentment', when we suspend our reactive attitudes,

> it is *never* the consequence of the belief that the piece of behaviour in question was determined in a sense such that all behaviour *might be*, and, if determinism is true, all behaviour *is*, determined in that sense. For it is not a consequence of any general thesis of determinism which might be true that nobody knows what he is doing or that everybody's behaviour is unintelligible in terms of conscious purposes or that everybody lives in a world of delusion, or that nobody has a moral sense, i.e., is susceptible of self-reactive attitudes, etc. (Strawson 2008 [1962]: 19; cited in Prasad 1995: 357–358)

---

3. The idea for the collection was conceived on the occasion of Strawson's visit to India in 1987 to 1988. At the time, Strawson's work had been studied and taught for nearly twenty years in India, where he was 'the best known and also the closest to the philosophic community' 'of all the leading among the living philosophers of the West' (Sen and Verma 1995: vii).

Prasad remarks that 'the above is an argument to the effect that determinism does not lead to, or imply, suspension of a reactive attitude because it does not lead to, or imply, any one of the three types of inhibitors [indeliberateness of action, unusual impairment of agent, agent's immunity from moral appraisal]' (Prasad 1995: 358). He reformulates Strawson's argument as follows:

> If any one of the three types of inhibitors exists, then it is inappropriate to feel or have any reactive attitude.
>
> It is not that if determinism is true, at least one of the three inhibitors exists.
>
> Therefore, it is not that if determinism is true, it is inappropriate to feel or have any reactive attitude. (Prasad 1995: 358)

Prasad then formalizes the argument, symbolizing 'Determinism is true' as D, 'It is inappropriate to feel or have any reactive attitude' as ~R, and 'At least any one of the three inhibitors exists' as $(I_1 \lor I_2 \lor I_3)$:

(1) $(I_1 \lor I_2 \lor I_3) \supset \sim R$
$\sim (D \supset (I_1 \lor I_2 \lor I_3))$
$\therefore \sim (D \supset \sim R)$

Prasad observes that (1) is invalid, because determinism might make the reactive attitudes inappropriate by itself, without implying that any one of the three inhibitors exists. He asks what needs to be done to make the argument valid:

> It will become a valid argument if we add another premise to the effect that if it is inappropriate to have a reactive attitude, then at least one of the three inhibitors exists: ~R $\supset (I_1 \lor I_2 \lor I_3)$.
>
> Then the argument assumes the form:

(1a) $(I_1 \lor I_2 \lor I_3) \supset \sim R$
$\sim R \supset (I_1 \lor I_2 \lor I_3)$
$\sim (D \supset (I_1 \lor I_2 \lor I_3))$
$\therefore \sim (D \supset \sim R)$

> (1a) is valid as it is not possible for its conclusion to be false without at least one of its premises being false.
>
> I do not know if Strawson would like to strengthen (1) as (1a). Sometimes I feel his argument is (1), and sometimes that it is (1a). (Prasad 1995: 358–359)

But even if Strawson's argument is (1a) and thus valid, he is in trouble, because

> the first two premises, taken together, state that a reactive attitude is made inappropriate *if and only if* the conditions referred to by $I_1 \lor I_2 \lor I_3$ exist. Therefore, to use them as premises is to assume that nothing except $I_1 \lor I_2 \lor I_3$ can make a reactive attitude inappropriate, or inhibit it, and thereby to *assume* that determinism cannot do this. To assume the latter is to assume that, if true, determinism is irrelevant to inhibiting our reactive attitudes. But the argument has been given to *prove* the latter and therefore it cannot assume it without being vicious. (Prasad 1995: 359)

Strawson cannot *assume* that only I_1, I_2 and I_3 can make reactive attitudes inappropriate; he has to *argue* for it. Prasad is happy to grant that determinism may not involve or imply the existence of any one of Strawson's inhibitors, but why could determinism not be an inhibitor in its own right? If determinism is true, there is no ability to do otherwise, and it seems inappropriate to adopt reactive attitudes towards people if they could not have acted otherwise than they did (Prasad 1995: 356). Prasad puts the point as follows: 'That determinism is not an inhibitor of reactive attitudes has to be proved as an independent thesis, and cannot be taken as having been proved by showing that

none of a certain set of inhibitors is a consequence of, or implied by, it. If a bullet can itself kill the game, it does not have to emit another missile to kill it' (Prasad 1995: 360). To prove 'as an independent thesis' that determinism is not an inhibitor requires taking what Prasad calls 'the straight path',

> whose destination is said to be either determinism's falsity, as libertarians claim to have arrived at, or its compatibility with freedom of the will, as the optimists among determinists do, or the impossibility of moral appraisal the pessimists among them find inescapable. We are thus back to the classical battle between the libertarians and determinists. (Prasad 1995: 360–361)

1.2 Strawson's Reply
Strawson's two-page reply to Prasad contains a remarkably candid reaction to Prasad's criticism:

> he [Prasad] shows conclusively that one of my arguments against the relevance of the thesis of determinism to the question of moral assessment and responsibility is either invalid or question-begging. I argue that the truth of determinism would not entail the presence in any particular case of one or more of the normal 'inhibitors' of moral reactions. But do I contend that such presence is merely sufficient or also a necessary condition of the inhibition? If I contend only the former, the argument is inconclusive; if the latter, it begs the question. (Strawson 1995: 430)

This is, as far as I can tell, the *only* place in Strawson's *oeuvre* where he repudiates an argument or claim made in 'Freedom and Resentment'. Not only does he accept Prasad's diagnosis of his argument, but he also accepts the advice to take the straight path:

> Let me concede that we rightly regard an action as an appropriate subject for moral judgement only if we regard the agent as one who could have acted, and chosen to act, otherwise than he did; only if he acted freely. The question is: what does this mean? It certainly means that he was not subjected to overwhelming external or internal compulsion to act as he did. It means more: that he did not lack the material and mental resources (he was rich or strong or intelligent enough) to do otherwise; and that he had the opportunity to do otherwise (there were no evidently insuperable obstacles in the way of doing so). Besides these somewhat negative considerations, it means also, ideally, that he knew what he was doing, was aware of other possibilities and chose to do what he did in the light of his beliefs about the facts and his attitudes (including moral attitudes) and preferences. All this seems to be about as full a statement as could be required of what it is to act freely, and hence, in the relevant sense, though not perhaps in every sense, responsibly. But I find in none of this an explicit or implicit denial of a thesis of determinism, stated, for example, in such simple and familiar terms as 'Every event has a cause'.
>
> If, in saying the above, I have dwindled into a mere compatibilist — in the company, say, of Hume — I am content with that. (Strawson 1995: 431)

Strawson accepts Prasad's challenge and takes, however sketchily, the straight path. He concedes that responsibility requires the ability to do otherwise and specifies what he means by the ability to do otherwise in a way that is meant to be compatible with the truth of determinism. He provides a list of conditions for responsibility and acknowledges that this puts him in the company of 'mere' compatibilists. I will discuss this 'mere' or 'straight' compatibilism in the next section. In section 3, I will explain how the reactive attitudes, abundantly present in

'Freedom and Resentment' but conspicuously absent in the quotation above, could be fitted into the straight compatibilist picture.

2. 'So Do I Emerge As a Straight Compatibilist?'

2.1 Compatibilism and the Freedom to Do Otherwise

In order to shed more light on the power of Prasad's criticism, I will relate it to an objection levelled against Strawson by Fischer (2014). Fischer's suggested answer to his objection will then be compared to Strawson's 'straight compatibilist' reply to Prasad (2.1). I will show that straight compatibilism is central to Strawson's later thought on freedom and responsibility, which is marked by explicit specifications of their conditions (2.2). I will indicate what this tells us about common interpretations of 'Freedom and Resentment' (2.3).

Prasad's criticism is similar to, but subtly different from, a criticism of 'Freedom and Resentment' developed by Fischer (2014). Fischer calls attention to Strawson's first category of inhibitors, Prasad's 'indeliberateness of action' category. This category includes all considerations 'which might give occasion for the use of the phrase "He couldn't help it", when this is supported by such phrases as "He was pushed", "He had to do it", "It was the only way", "They left him no alternative", etc.' (Strawson 2008 [1962]: 7–8). According to Fischer, this suggests 'that when an agent is not free to do otherwise — could not have done otherwise — he is not legitimately deemed morally responsible for the behavior in question […] But this lands one squarely in the traditional metaphysical debates about the relationship between causal determinism and "could have done otherwise" and also the relationship between "could have done otherwise" and moral responsibility' (Fischer 2014: 99). Fischer does not deny that 'He couldn't help it', etc. can be given compatibilist construals, but he suggests that Strawson cannot simply *assume* that the compatibilist construal is the relevant one. He cannot just bypass questions about the relevance of determinism to his inhibitors, and doing so makes his argument 'at best, incomplete' (Fischer 2014: 103).

Fischer's diagnosis of Strawson's argument is similar to Prasad's, but there is a difference. Fischer questions whether determinism and the lack of freedom to do otherwise that one might take it to entail would make the inhibiting conditions mentioned by Strawson ('He couldn't help it') obtain universally. That is, he targets the second premise of argument (1), the third of (1a). Prasad is prepared to accept these premises, but he argues that, even if they are true, Strawson has not established the irrelevance of determinism to responsibility. He targets the second premise of (1a). Taken together, Fischer's and Prasad's criticisms yield a compelling conclusion that holds regardless of one's views about the relation between determinism, Strawson's inhibiting conditions and responsibility: Strawson's argument cannot be made to work unless he takes the straight path.

And Strawson's path is very straight indeed, straighter than Fischer's. According to Fischer, 'He couldn't help it', 'He had to do it' and 'It was the only way' suggest that the agent lacked the freedom to do otherwise. If determinism is true, we have no such freedom or, in Fischer's terminology, no 'regulative control'. But the expressions can also be understood in another way: instead of suggesting that the agent lacked the freedom to do otherwise (regulative control), the expressions could be taken to suggest that the agent did not act freely or, in Fischer's terminology, had no 'guidance control'. Freedom to do otherwise (regulative control) is threatened by determinism but unnecessary for responsibility; acting freely (guidance control) is necessary for responsibility but not threatened by determinism. Guidance control requires ownership of the mechanism that issues in the agent's behavior (which excludes manipulation cases) and the reasons-responsiveness of that mechanism. These requirements can be met if determinism is true. Thus, Fischer shows that there is a reading of 'He couldn't help it', etc. so that determinism would not make the inhibiting conditions referred to by these expressions obtain universally (Fischer 2014: 105–115).

In contrast to Fischer, Strawson does *not* disambiguate the relevant expressions ('only if we regard the agent as one who could have acted,

and chosen to act, otherwise than he did; only if he acted freely'). He sees no need for compatibilists to give up on the freedom to do otherwise. He does *not* concede to the incompatibilist, as Fischer does, that there is a kind of freedom, though not the kind relevant to moral responsibility, that we lack if determinism is true. The libertarian's conception of free will, associated by Strawson with formulae such as 'contra-causal freedom' (Strawson 2008 [1962]: 25) and 'uncaused acts of will' (Strawson 1980: 260), is not just to be rejected as irrelevant to moral responsibility; it cannot even coherently be stated. This point was made (though scarcely defended) in 'Freedom and Resentment' (Strawson 2008 [1962]: 25) and repeated later (Strawson 1998a: 170). But if we grant that the libertarian's conception of free will is incoherent, are we then not compelled to say the same about the freedom to do otherwise, with which the libertarian's conception of free will is frequently associated? Strawson answers this question in his later work in a clear and unambiguous way. The freedom to do otherwise is not threatened by determinism, as concessive compatibilists à la Fischer think. The straight compatibilist can have it all.

2.2 Conditions of Freedom and Responsibility

One might think that Strawson's reply to Prasad was once only and not representative of his thought. But that would be mistaken. Strawson makes similar points at several places in his later work. He writes the following about the relation between determinism and the ability to do otherwise in 'Freedom and Necessity':

> Now it certainly is generally held — it is a thesis, one might say, of the common moral consciousness — that the appropriateness of these attitudes and feelings [reactive attitudes], the applicability of these notions, requires, in respect of any occasion on which these attitudes and notions are in question, that the agent could have acted otherwise than he did on that occasion. But — so the argument runs — if the thesis of determinism is true, then

it is not true of any agent on any occasion that that agent could have acted otherwise than he did on that occasion. Hence, if the thesis of determinism is true, the attitudes and notions in question are never appropriate.

> [...] It is certainly true that often, in the context of a moral judgment (especially if disapprobative) one may utter the words, 'He could have acted otherwise', or other words to the same effect. But are such words, as then uttered, really equivalent to 'There was no sufficient natural impediment or bar, *of any kind whatsoever, however complex*, to his acting otherwise'? I find it difficult, as others have found it difficult, to accept this equivalence. The common judgment of this form amounts rather to the denial of any sufficient natural impediment *of certain specific kinds or ranges of kinds*. For example, 'He could (easily) have helped them (instead of withholding help)' may amount to the denial of any lack on his part of adequate muscular power or financial means. Will the response, 'It simply wasn't in his nature to do so' lead to a withdrawal of moral judgement in such a case? I hardly think so; rather to its reinforcement. (Strawson 1992b: 137; see also Strawson 2011c [1983]: 150)

This passage shows that Strawson found his thoughts about the freedom to do otherwise important enough to include in *Analysis and Metaphysics*. My point is not that Strawson came to think about the freedom to do otherwise in response to Prasad.[4] Rather, it is that the

4. As explained in note 1, 'Freedom and Necessity' is almost identical to 'Liberty and Necessity', published in 1983. 'Liberty and Necessity' and 'Freedom and Necessity' were both published *before* Strawson's reply to Prasad. Why, then, was Strawson much more explicit about the freedom to do otherwise in his later works than he had been in 'Freedom and Resentment'? I entirely agree, on this point, with the suggestion of an anonymous reviewer: at least part of the explanation for this change of emphasis is that, in 1962, it could be more or less taken for granted that the freedom to do otherwise was relevant to moral responsibility. Works published after 'Freedom and Resentment' by

freedom to do otherwise is prominent in his later thought. He knew that his account in 'Freedom and Resentment' was vulnerable to the kind of criticism formulated by Prasad and Fischer.

I cannot evaluate the accuracy of Strawson's rough analysis of 'could have done otherwise' here. I would only like to draw attention to the way in which Strawson sets up the problem about determinism and the freedom to do otherwise in the first part of the quotation, which is reminiscent of the way in which Paul Russell (2017a [1992]) has formulated an influential objection to Strawson. Russell distinguishes between type-pessimism, which claims that we need a justification for the fact that we are susceptible to the reactive attitudes and do not have one if determinism is true, and token-pessimism, which claims that we are never justified in entertaining any tokens of the reactive attitudes if determinism is true. Russell argues that Strawson's type-naturalism, which claims that liability to the reactive attitudes is natural to humans and requires no rational justification, constitutes a good answer to the type-pessimist. But Strawson does *not* have a good answer to the token-pessimist. He acknowledges, in 'Freedom and Resentment', that our reactive attitudes may, in particular instances, be inappropriate. Why exclude, then, that the truth of determinism could make it the case that they are *never* appropriate, that the inappropriateness generalizes if determinism is true? Russell thinks (and this sounds familiar by now) that Strawson's account is 'at best, incomplete' (Russell 2017a [1992]: 45; Russell 2017b [2004]: 75). Strawson cannot simply assume that determinism does not threaten moral responsibility; he owes us a compatibilist construal of the conditions of responsibility. A compatibilist construal of one such condition, the freedom to do otherwise, has been given in the quotation above and in the reply to Prasad.

_____

Chisholm (1966), Frankfurt (1969), and Lehrer (1976), among others, made it impossible to take this for granted. Thus, the relevance of the freedom to do otherwise had to be made explicit in Strawson's works from the 1980s and 1990s.

But perhaps *other* conditions of responsibility are in tension with determinism? In his reply to Prasad, Strawson writes that 'we rightly regard an action as an appropriate subject for moral judgment only if we regard the agent as one who could have acted, and chosen to act, otherwise than he did; only if he acted freely […] and hence, in the relevant sense, though not perhaps in every sense, responsibly' (Strawson 1995: 431). Strawson suggests that 'could have acted, and chosen to act, otherwise' is *equivalent* to 'acted freely' and (in the relevant sense) 'acted responsibly'. He lists six conditions, three negative and three positive ones. The negative conditions are:

(1) The agent was not subjected to overwhelming external or internal compulsion.

(2) The agent did not lack the material or mental resources to do otherwise.

(3) The agent had the opportunity to do otherwise (there were no insuperable obstacles).

The positive conditions are:

(4) The agent knew what he was doing.

(5) The agent was aware of other possibilities.

(6) The agent chose to do what he did in the light of their beliefs about the facts and their attitudes (including moral attitudes) and preferences.

Strawson adds that this is about 'as full a statement as could be required' of what it is to act responsibly and claims (with a confidence that will undoubtedly seem unwarranted to the incompatibilist) to find 'in none of this an explicit or implicit denial of a thesis of determinism'.

How does Strawson's analysis of 'could have done otherwise' in the reply to Prasad relate to his analysis in 'Freedom and Necessity' ('the denial of any sufficient natural impediment *of certain specific kinds or ranges of kinds*')? I suggest the following: An agent could have acted

otherwise if the negative conditions (1)–(3) are satisfied. They could have acted, *and chosen to act*, otherwise if the positive conditions (4)–(6) are also satisfied. If both negative and positive conditions are satisfied, the agent acted freely and responsibly.

Apart from the reply to Prasad, there are other passages in which Strawson mentions requirements or conditions of freedom and responsibility. Here is a passage from his reply to Pears (1998):

> There [in 'Freedom and Resentment'] I spoke of our natural responses, personally or vicariously felt, to certain states of mind we discern as manifested by agents in their actions — notably their good or ill will or indifference towards ourselves or others. Circumstances which exclude manifestation of such states, so manifested, do not evoke the responses in question (e.g., gratitude or approbation, resentment or indignation) — and such circumstances are precisely the generally recognized defeating conditions of responsibility.
>
> Now, let us turn briefly from defeating conditions to what we would normally regard as the generally necessary conditions of someone's acting, as we say, freely or voluntarily, and hence being responsible for what he does. They seem to include the following: the agent knows what he is doing, is aware of what at least seemed to him other possibilities of action and chooses to do what he does in the light of his belief about the facts and of his attitudes, principles, and preferences. (Note, parenthetically, that the notion of choice makes no sense here unless it at least seems to the agent as if other possibilities of action are open to him.) The satisfaction of what we normally regard as these positive conditions of responsibility carries with it at least the possibility, though by no means always the actuality, of the agent's behavior manifesting one or another of those states of mind which typically

> invoke in an observer (or patient) of his action the reactive attitudes I spoke of. (Strawson 1998b: 261)

This is a complex passage. Note, first, the formulation 'acting […] freely or voluntarily, and hence being responsible'. Voluntariness is not seen as a *condition* of responsibility (at the same level as, for example, a knowledge condition). Rather, to act voluntarily *is* to act responsibly. Second, Strawson distinguishes between 'defeating' and 'generally necessary' conditions or, in the terms I used before, between negative and positive conditions. The positive conditions are the same as those listed in the reply to Pears, with a small change in (5), where 'other possibilities' has become 'what at least seemed to him other possibilities'. (This change is not insignificant. One might think that, if determinism is true, there are no other possibilities, but it may still seem to the agent that there are.)

The negative conditions of responsibility are specified in a way that is different from what we saw in the reply to Prasad. Strawson now identifies them with the conditions/circumstances which exclude manifestations of good or ill will or indifference, circumstances which do not evoke the reactive attitudes. I will say more about the reactive attitudes in section 3. For now, let us have a look at the relation between the negative conditions (1)–(3) and the conditions which exclude manifestations of good or ill will or, in other words, the groups of 'inhibitors' (I_1, I_2, I_3). It will be readily apparent that they cannot be neatly mapped onto one another: negative condition (1) does not correspond to I_1, etc. I_1 contains the conditions referred to by 'He didn't know', 'It was the only way' and 'They left him no alternative'. These excusing conditions correspond, albeit roughly, to the absence of Strawson's *positive* conditions of responsibility (4)–(6). I_2 contains the conditions referred to by 'He wasn't himself', 'He has been under very great strain recently' and 'He was acting under post-hypnotic suggestion'. Strawson characterizes these conditions as ones in which the agent is normal but the circumstances are not. I_3 contains the conditions referred to by 'He's only a child', 'He's a hopeless schizophrenic',

'His mind has been systematically perverted' and 'That's purely compulsive behaviour on his part'. Here the circumstances may be normal, but the agent is not (Strawson 2008 [1962]: 7–8). The category of exempting conditions (including I_2 and I_3) corresponds, again roughly, to Strawson's category of *negative* or defeating conditions. Conditions (1) and (3) (and perhaps also the part about material resources in condition (2)) seem to correspond to the first sub-group of exempting conditions. Condition (2), and especially the part about mental resources, corresponds more closely to the second sub-group of exempting conditions.[5] If there are no exempting/defeating conditions of responsibility (that is, if the negative conditions are satisfied) *and* no excusing conditions (that is, if the positive conditions are satisfied), we get the 'possibility, though by no means always the actuality, of the agent's behavior manifesting one or another of those states of mind [good will, ill will, indifference] which typically invoke in an observer (or patient) of his action the reactive attitudes'.

I have not evaluated Strawson's analysis of the freedom to do otherwise, and neither will I evaluate his analysis of the conditions of responsibility. Rather, my aim has been to show how central the straight compatibilism, the link between responsibility and freedom and the idea that a list of conditions of responsibility can be provided are to his later thought. The following passage, from his reply to Blackburn (1998), confirms this:

> It is quite true, indeed tautological, that an agent can be justly blamed for an action only when he can be held morally responsible for it; and it is true that he is responsible for so acting only if he acted freely. But 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. Human beings, or

5. Note that Strawson avoids, in his later work, the language of (ab)normality which some of his commentators found problematic in 'Freedom and Resentment' (see, for instance, Russell 2017a [1992]).

the more judicious among them, are really quite skilled in determining whether, and to what degree, the conditions of moral responsibility are satisfied. It would be tedious to rehearse them, since this has been done again and again. So we can relax: the whole issue between determinists and libertarians is an irrelevance; and the fact that it has been so long and earnestly debated is but one more illustration of the tendency of philosophers to raise a dust and then complain they cannot see. So do I emerge as a straight compatibilist? If so, *ainsi soit-il*. (Strawson 1998a: 170)

2.3 Interpretations of 'Freedom and Resentment'

What are the consequences of Strawson's compatibilism, and of his emphasis on freedom in discussions of responsibility, for contemporary readings of 'Freedom and Resentment'? First, he seems to have provided direct answers to at least some charges of incompleteness levelled against his account in 'Freedom and Resentment' by Prasad, Fischer and Russell. Though sketchy, his answer is of the right, straight kind demanded by his critics.

Second, the freedom to do otherwise plays a crucial role in Strawson's answers. This is remarkable, because it has often been claimed that 'Freedom and Resentment' was groundbreaking precisely because it *removed* the focus on questions about free will and determinism from debates about responsibility. Fischer (2014: 113) claims to 'have agreed with Strawson that questions about the relationship between causal determinism and freedom to do otherwise can be separated from issues about moral responsibility'. As David Shoemaker (2020: 212) puts it, 'Strawson aimed to free us from the quest for freedom'. According to Russell (2017b [2004]: 67), whereas traditional or classical debates about responsibility focused on the problem of freedom, Strawson directed his attention elsewhere, namely to the role of the reactive attitudes. I believe that, in light of the above, it would be more accurate to say that Strawson aimed to *reorient* (rather than free us from) the quest

for freedom. His attention to quality of will and the reactive attitudes was not meant to direct us *away from* the problem of freedom. Rather, it was meant to direct us *to* a place where the kind of freedom relevant to moral responsibility could be found.

Third, the conditions of responsibility listed by Strawson do not radically diverge from conditions that have traditionally been put forward, and they are not *meant* to diverge from them ('It would be tedious to rehearse them [the conditions of moral responsibility], since this has been done again and again'). They seem to include a knowledge condition (condition 4), a voluntariness condition (condition 6; note, though, that Strawson's own use of 'voluntary' as equivalent to 'freely' is broader than this) and a control condition (condition 1). It becomes difficult, then, to understand Strawson as offering a radical *alternative* to traditional analyses of responsibility in terms of knowledge, voluntariness and control, as Shoemaker (2017) does.

One could try to resist the idea that Strawson's later work on freedom and responsibility has such direct consequences for interpretations of 'Freedom and Resentment'. Perhaps there has simply been a change in his thought (as 'dwindled into a mere compatibilist' might seem to suggest). First, however, it is noteworthy that many of the quoted passages occur in replies to commentators such as Prasad and Pears whose pieces are explicitly framed as discussions of 'Freedom and Resentment'. One expects Strawson, in his replies to such criticisms, to *clarify* his position in the paper under discussion (or, if a problem has been identified, to address it in a way that is compatible with his earlier statements) and to flag changes in his thought, as he does in no uncertain terms when he repudiates the argument criticized by Prasad. Second, the move from (1) a radical alternative to traditional analyses of responsibility in terms of freedom, knowledge, voluntariness, etc. (supposedly provided in 'Freedom and Resentment') to (2) something that seems to come fairly close to such a traditional analysis (provided by the 'later' Strawson) would be very significant. It is unlikely that Strawson would not have noticed it. Third, there is no need to postulate a change, because, as I will show in the next section, Strawson's

later remarks are compatible with the thoughts expressed in 'Freedom and Resentment', and they do not detract from its significance.[6]

3. Responsibility, Quality of Will and the Reactive Attitudes

3.1 Quality of Will Links Responsibility to the Reactive Attitudes
It is undeniable that the moral sentiments or reactive attitudes are essential to Strawson's account of responsibility in 'Freedom and Resentment', but they may not seem essential to his later thought, as exemplified in the passages quoted above. His list of conditions in the reply to Prasad makes no reference to the reactive attitudes. His list of conditions in the reply to Pears *does* refer to them, but since the list seems roughly equivalent to the list in the reply to Prasad, one may wonder whether the reactive attitudes are essential to his later thought; after all, it appears to be possible to specify the conditions of responsibility *without* referring to the reactive attitudes. And if they are inessential, then something distinctive about Strawson's account, the aspect of our moral lives that he directed our attention to (Russell 2017b [2004]: 67), the very thing that makes his account significant and different from traditional accounts, seems lost.

6. A reviewer asks how the following two claims fit together. (1) Strawson gave up on one of his main points in 'Freedom and Resentment', and (2) the later work (which rejects that point) does not represent a change of views. Strawson indeed gave up on the point criticized by Prasad (see section 1), but he did not seem to regard this (and, as I hope to show in this paper, did not need to regard it) as a major problem for the view presented in 'Freedom and Resentment'. The point he gave up on was basically that there is no need to explain how determinism is compatible with freedom of the will (and, more specifically, with the freedom to do otherwise). In his later work, Strawson accepts that there *is* such a need, and he takes what I call the 'straight path'. Thus, in his later work, Strawson sees the need to explain something that he did not explain (and did not think he had to explain) in his earlier work, but the way in which he explains it in his later work is perfectly compatible with everything he says in 'Freedom and Resentment' (except, of course, for the idea that there is no need to explain how determinism is compatible with the freedom to do otherwise). One might say that this amounts to a change of view, but I do not think (and Strawson did not seem to think) that it is a major one.

In this section, I will indicate how the reactive attitudes fit into the picture of Strawson's straight compatibilism, as introduced in the previous sections. I will emphasize the importance of quality of will in Strawson's account (3.1), elucidate a key distinction between the questions 'How does our concept of responsibility operate?' and 'Why do we have the concept of responsibility?' (3.2) and situate Strawson's account in relation to the distinction between response-dependence and response-independence (3.3). I take the following passage from the reply to Pears, already partly quoted above, as my starting point:

> I agree that we should consider seriously the origin — the fount and origin — of the distinction we draw in practice between cases where responsibility is ascribed and cases where it is not. But I think Pears gives insufficient weight […] to just those features of the fount and origin which I was at pains to emphasize in 'Freedom and Resentment'. There I spoke of our natural responses, personally or vicariously felt, to certain states of mind we discern as manifested by agents in their actions — notably their good or ill will or indifference towards ourselves or others. Circumstances which exclude manifestation of such states, so manifested, do not evoke the responses in question (e.g., gratitude or approbation, resentment or indignation) — and such circumstances are precisely the generally recognized defeating conditions of responsibility.
>
> […] The satisfaction of what we normally regard as the positive conditions of responsibility carries with it at least the possibility, though by no means always the actuality, of the agent's behavior manifesting one or another of those states of mind which typically evoke in an observer (or patient) of his action the reactive attitudes I spoke of.
>
> […] we *could* say that no one is *ultimately* responsible for his or her actions, thereby doing a sort of justice to one

apparently ineliminable natural-philosophical conviction. But *ultimate* responsibility in this sense is neither here nor there. We have the empirically founded distinctions with which in practice we worked; and once we are fully clear about the real nature of those foundations, foundations in phenomena which we are quite good at discriminating, we see equally clearly that no such general doctrine as determinism, in any of its forms, would possibly disturb them.

> So, it will be seen I agree with the conclusion reached by Professor Pears; but I hope to have made its foundations more secure by referring again to the phenomenology of the moral life. (Strawson 1998b: 260–262)

Strawson repeats the point about ultimate responsibility made in his reply to Blackburn. Two further aspects of this passage merit attention. The first concerns the relation between the reactive attitudes and 'good or ill will or indifference', or what is known as 'quality of will'. The reactive attitudes are described as natural responses to the quality of will we discern as manifested by agents in their actions. *The satisfaction of the conditions of responsibility makes it possible for an agent's behavior to manifest a certain quality of will, and behavior manifesting a certain quality of will typically evokes reactive attitudes.* This way of putting things seems to leave little room for readings of Strawson in which the reactive attitudes play an essential role but quality of will does not (see also Strawson's characterization of the reactive attitudes as 'essentially reactions to the quality of others' wills towards us' in 'Freedom and Resentment'; Strawson 2008 [1962]: 15). It also suggests that, according to Strawson, quality of will should not be thought of as a condition of responsibility in the sense in which knowledge, control, etc. are conditions of responsibility. Quality of will links responsibility to the reactive attitudes, and responsibility is a condition of quality of will rather than the other way round.

I propose to have a look, in this regard, at Shoemaker's (2017) response-dependent account of responsibility.[7] Although Shoemaker develops a theme from 'Freedom and Resentment' rather than provide an interpretation of it, it may be instructive to bring out the tension between his account and Strawson's. Shoemaker starts from the idea that traditional accounts of responsibility are response-*in*dependent: they attempt to specify the necessary and sufficient conditions of responsibility without making essential reference to our responsibility responses, that is, without referring to the reactive attitudes. Conditions that have been proposed include knowledge, voluntariness, control, quality of will and a history condition. All such analyses, however, regardless of how the conditions are combined or whatever additional conditions are invoked, are vulnerable to false negatives and/or false positives. Thus, we need a different kind of account, a response-*dependent* one, which specifies the necessary and sufficient conditions of responsibility by making essential reference to the reactive attitudes. Shoemaker's own version of response-dependence can, for my purposes, be summarized as follows: an agent is morally responsible if and only if, and in virtue of the fact that, it would be fitting to adopt a reactive attitude towards them.

How does Shoemaker's account relate to Strawson's view? First, in his reply to Prasad, Strawson specifies the conditions of responsibility (including a knowledge, voluntariness and control condition) without reference to the reactive attitudes. This is a feature of response-independent accounts. Second, Shoemaker regards quality of will as a (candidate) condition of responsibility, but it does not play that role in Strawson's thought. Third, Shoemaker downplays the importance of quality of will and gives pride of place to the reactive attitudes, while quality of will and the reactive attitudes are inextricably linked in Strawson's account. Fourth, it has been argued that Shoemaker has difficulties to account for the fact that we use 'She didn't know', etc. to explain why we do not hold an agent responsible, that is, for the

7. Shoemaker uses 'blameworthiness' instead of 'responsibility', but the difference is not important for my purposes.

explanatory relations between knowledge, voluntariness and control, on the one hand, and responsibility on the other (Heyndels and De Mesel 2018: 450). Strawson can straightforwardly account for these explanatory relations.

3.2 *The Natural Foundations of Our Concept of Responsibility*

I do not think that Shoemaker situates quality of will and the reactive attitudes where Strawson situates them. Strawson's view can be elucidated by focusing on a second important aspect of the quotation from his reply to Pears: the use of terms such as 'fount', 'origin' and 'foundations' in descriptions of the role of the reactive attitudes. Strawson thinks that the reactive attitudes are 'features of the fount and origin' 'of the distinction we draw in practice between cases where responsibility is ascribed and cases where it is not'. His references to 'the phenomenology of the moral life' (that is, to the role of the reactive attitudes) are meant to clarify (make 'fully clear') 'the real nature' of the 'foundations' of the 'distinctions [between responsible and irresponsible agents] with which in practice we worked' (Strawson 1998b: 260–262).

Strawson's use of 'foundations' in descriptions of the role of the reactive attitudes can be fruitfully related to his use of the same term in some of his methodological work (Strawson 1963, 2011a [1956], 2011b [1967]). There he distinguishes between several philosophical tasks, two of which are of particular importance here. The first is, very roughly, the task of explaining how our concepts operate. The second is the task of explaining

> why it is that we have such concepts and types of discourse as we do […]. This is not an historical enquiry. It attempts to show the natural foundations of our logical, conceptual apparatus in the way things happen in the world, and in our own natures. […] It might reasonably be maintained, or ruled, that full understanding of a concept is not achieved until this kind of enquiry is added

to the activities […] which I mentioned first. (Strawson 1963: 515–516)

Full understanding of the concept of responsibility is not achieved until we have answered at least two questions. The first question is, 'How does our concept of responsibility operate?' An answer to that question can be provided by investigating the conditions of application of the concept, the conditions which something must fulfill to satisfy it. The second question is, 'Why do we have the concept of responsibility?' An answer to that question shows the foundations of the concept of responsibility in 'the way things happen in the world, and in our own natures'.

Strawson's references to the 'fount', 'origin' and 'foundations' of the distinction we draw between those who are and those who are not responsible in 'phenomena which we are quite good at discriminating' (compare 'the way things happen in the world') and 'our natural responses' (compare 'our own natures') suggest that the reactive attitudes figure essentially in Strawson's answer to the *second* question, the question why we have the concept of responsibility. His lists of conditions of responsibility in the replies to Prasad and Pears contain his answer to the *first* question, the question about the operation of the concept of responsibility, specifying its conditions of application. There is no tension or conflict between what Strawson says about quality of will and the reactive attitudes, on the one hand, and what he says about the conditions of responsibility, on the other, because they figure in two different but related kinds of enquiry which can be 'added to' one another. An explanation of responsibility in terms of reactive attitudes is not an *alternative* to an explanation in terms of knowledge, control, etc.

How are the two kinds of enquiry related? Strawson insists, in 'Freedom and Resentment', upon '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'

(Strawson 2008 [1962]: 5). He emphasizes 'how much we actually mind, how much it matters to us, whether the actions of other people — and particularly of *some* other people — reflect attitudes towards us of goodwill, affection, or esteem on the one hand or contempt, indifference, or malevolence on the other' (Strawson 2008 [1962]: 5–6). We are naturally concerned about the quality of will manifested in the actions and attitudes of others, and our reactive attitudes express that concern. We have the concept of responsibility because we are so concerned; our natural concern about quality of will and our need to express that concern are the natural foundations of the concept of responsibility.[8]

I have suggested that, according to Strawson, responsibility is a condition of quality of will rather than the other way round. It is impossible to manifest quality of will without acting responsibly. The possibility of acting responsibly (satisfying all conditions of responsibility) without manifesting any quality of will (good or ill will or indifference) is nothing more than the possibility of acting responsibly while having a 'neutral' will (neither good nor bad nor indifferent).

8. Watson (2014) distinguishes between a basic concern (about quality of will) and a basic demand (to be treated with good will). He claims that the concern underlies the demand, which is 'an intelligible expression of the basic concern. But it is only with the basic demand that a distinctive stance of holding responsible emerges' (Watson 2014: 19). I wish to remain neutral here about the question whether it is the demand or the concern that is the natural foundation of our concept of responsibility. A commentator who thinks that it is the concern rather than the demand that underlies responsibility is Beglin (2018, 2020). According to Beglin's concern-based strategy, basic human concerns 'ultimately explain why we have responsibility practices' (Beglin 2018: 618). His account 'goes beyond considering *how* we engage with morally responsible agents and focuses on *why* we engage with them in that way' (Beglin 2020: 2361). Thus, Beglin also seems to distinguish between a 'how' and a 'why' question, although the questions are slightly different from mine. I cannot discuss Beglin's account in detail here, so I will just mention two differences between his and Strawson's accounts as presented in this paper. First, Beglin seems to situate the reactive attitudes at the level of the 'how' question, and the concern at the 'why' level, whereas I suggest that, according to Strawson, the concern for quality of will and the reactive attitudes belong together at the 'why' level. Second, Beglin has doubts about Strawson's identification of the basic concern(s) with a concern about quality of will (Beglin 2020: 2356).

– 13 –

Strawson refers to the natural foundations of the concept of responsibility in our concern for quality of will in order to show that *conditions are relevant to moral responsibility because they are relevant to quality of will*. Thus, the link between responsibility and its natural foundation in a human concern serves to explain why some conditions are relevant to responsibility and others are not. Whether an agent was subjected to overwhelming compulsion, for instance, is relevant to assessing their responsibility because it is relevant to assessing their quality of will. The conditions of responsibility can be recovered by attending to the conditions of quality of will, and because the reactive attitudes are essentially reactions to quality of will, the conditions under which the latter are appropriate will be the conditions of responsibility. This is the gist of Strawson's famous remark that 'only by attending to this range of [reactive] attitudes can we recover from the facts as we know them as sense of what we mean, i.e. of *all* we mean, when […] we speak of […] responsibility' (Strawson 2008 [1962]: 24).

If the question is whether determinism is relevant to responsibility, we have to ask whether it is relevant to quality of will. Would the truth of determinism make it the case that actions or attitudes never manifest good or ill will or indifference? Would it invalidate the distinction between someone treading on my hand accidentally (manifesting no quality of will) and their doing so 'in contemptuous disregard of my existence or with a malevolent wish to injure me' (Strawson 2008 [1962]: 6) (manifesting ill will)? Because of the foundation of responsibility in our concern about quality of will, those who argue that determinism is relevant to responsibility are committed to the claim that it is relevant to quality of will.

To the question 'What are the conditions of responsibility?', a subquestion of the more general question 'How does our concept of responsibility operate?', Strawson answers with a traditional list of conditions. He takes the straight path when he explains that these conditions can be understood in a way that is compatible with the truth of determinism. But explaining *what* the conditions of responsibility are and how they can be understood in a compatibilist way is not

equivalent to explaining *why* these are the conditions of responsibility and why the compatibilist construal is the relevant one. That is where the second kind of enquiry comes in, centering on the question 'Why do we have the concept of responsibility?' We have it because it serves our need to express a natural concern for the quality of will manifested in people's attitudes and actions. If we see that responsibility is so rooted, we are in a position to see *how the conditions of responsibility are to be understood*. They are to be understood in a way ('all we mean') that is relevant to quality of will, in a way in which they could reasonably be understood *as conditions of quality of will*. Determinism is irrelevant to responsibility because it is irrelevant to quality of will.

This argument can, of course, be challenged. Critics could bite the bullet and argue that the truth of determinism would make it the case that nobody ever manifests good or ill will. They could try to break the link between quality of will and (some senses of) responsibility or argue that 'quality of will' is ambiguous or unclear. But there is no room, in light of the above, to downplay the weight that quality of will bears in Strawson's account, which is arguably what Fischer, Prasad and Russell have done.[9] They ask why Strawson *assumes* that the conditions of responsibility, or the conditions under which the reactive attitudes are appropriate, have to be construed in a compatibilist way. But Strawson does not assume this. He argues that the conditions of responsibility are basically the conditions under which it is possible for an agent's actions to manifest quality of will. He believes that, if this is true, it makes things much more difficult for the incompatibilist,

---

9. By contrast, Scanlon (1988), Beglin (2018, 2020), Hieronymi (2020) and McKenna (2012) give a lot of weight to quality of will in their interpretations of 'Freedom and Resentment'. McKenna's account comes close to the one I attribute to Strawson (see McKenna 2012: 61, note 3). He explains 'quality of will' in a helpful way to which Strawson would have been sympathetic — namely, in terms of the *value* or *worth* (quality) of an agent's *regard* or *concern* for others (will) (McKenna 2012: 58–60). It is beyond the scope of this paper to discuss the exact relation between McKenna's account and the one I attribute to Strawson. One important difference seems to be that McKenna, like Shoemaker, regards quality of will as a condition of responsibility rather than the other way round.

because it seems much more implausible to claim that determinism threatens quality of will than that it threatens responsibility. If he assumes anything, it is that determinism would not make it the case that actions or attitudes never manifest any quality of will. His critics are challenged to show that it would.

3.3 Response-Dependence or Response-Independence?

Is Strawson's account of responsibility response-(in)dependent? An in-depth discussion is outside the scope of this paper, but some preliminary remarks may help to see where it could go. First, we have seen that, in his reply to Prasad, Strawson specifies the conditions of responsibility without referring to the reactive attitudes. If an account of responsibility is response-dependent in virtue of the fact that it makes essential reference to the reactive attitudes in its specification of the conditions of responsibility, Strawson's account is response-*independent*. Or, to put it otherwise, if we stay at the level of the first question, 'What are the conditions of responsibility?', no reference to the reactive attitudes is necessary. But Strawson insists that full understanding of a concept cannot be achieved until we answer the second question, 'Why do we have the concept of responsibility?' If *that* is the question (or at least part of the question) that accounts of responsibility must seek to answer, then Strawson's answer *does* make essential reference to the reactive attitudes. If an account of responsibility is response-dependent in virtue of the fact that it makes essential reference to the reactive attitudes in a full explanation of the concept, Strawson's account is response-*dependent*. We can only understand what responsibility is, and how its conditions are to be understood, if we understand that the concept of responsibility is rooted in our natural concern for quality of will (expressed in the reactive attitudes). Perhaps, then, the best and unsurprising answer to the question whether Strawson's account is response-dependent or response-independent is 'It depends on what you understand by response-(in)dependent'.

4. Conclusion

My aim in this paper has been to present Strawson's later, neglected work on freedom and responsibility. I have highlighted three aspects. First, Strawson explicitly rejects an argument put forward in 'Freedom and Resentment' in response to a criticism by Prasad. Second, Strawson takes up the challenge presented by Prasad, Fischer and Russell to take the 'straight path', that is, to be straightforward about the relation between determinism, freedom, the ability to do otherwise and the conditions of responsibility. His view is a classical compatibilist one.[10] Third, Strawson clarifies the relation between responsibility, quality of will and the reactive attitudes. The latter do not figure essentially in his answer to the question 'What are the conditions of responsibility?', but they do play an essential role in his answer to the question 'Why do we have the concept of responsibility?' We only have it, Strawson suggests, because of our natural concern about the quality of will with which people act, a concern expressed in our reactive attitudes.

I have indicated that, although Strawson's later work definitely involves a shift of emphasis when compared to 'Freedom and Resentment', his overall account of freedom and responsibility is coherent.

10. This means, as a reviewer has rightly noted, that Strawson's view, as presented in this paper, is vulnerable to many of the objections faced by classical compatibilism (see, for instance, the Consequence Argument in van Inwagen 1983), which may well make his view more problematic for many contemporary commentators than they thought. The reviewer suggests that it is a significant advantage for an approach such as Fischer's that it can avoid these challenges. A detailed critical comparison between Strawson and Fischer falls outside the scope of this paper, but I would like to emphasize that, in my view, the significance of Strawson's contribution does not lie primarily in his characterization of the conditions of moral responsibility. Strawson acknowledges that 'it would be tedious to rehearse them [the conditions of moral responsibility], since this has been done again and again' (Strawson 1998a: 170). I distinguish in this paper between two questions: (1) 'How does our concept of responsibility operate?', and (2) 'Why do we have the concept of responsibility?' While Strawson's answer to the first question is not very original, as he himself seems to acknowledge, I believe that the significance of his contribution to the debate about responsibility lies primarily in his answer to the second question. So even if his answer to the first question has some disadvantages in comparison to other answers, Strawson's contribution remains, in virtue of his answer to the second question, highly original and significant.

The later work helps to better understand this overall account and its significance and to identify problems with interpretations of and objections to 'Freedom and Resentment'.

There is more to Strawson's later work on freedom and responsibility than I have been able to discuss here. He emphasizes, perhaps in even stronger terms than in 'Freedom and Resentment', the depth of the reactive attitudes' 'entrenchment in our whole conception of what it is to be human' (Strawson 1995: 420). They 'form an essential part of our conception of the human' (Strawson 1995: 420) and 'govern all our relations with one another and our reactions to human behavior in general' (Strawson 1998b: 259). He links our sense of freedom to the experience of deliberation, our sense of self and the experience of agency (Strawson 1992b: 135). He agrees with Prasad 'that unless the determinist thesis is spelled out (precisified?) in much fuller physical and psychological detail, we remain unclear as to what either its affirmation or its denial effectively amounts to' (Strawson 1995: 432), suggesting that, in 'Freedom and Resentment', he had overestimated the possibility of answering questions about determinism and responsibility without knowing what exactly the thesis of determinism is (Strawson 2008 [1962]: 11). And there is an interesting discussion in 'Freedom and Necessity' about the dependence of the mental and behavioral on the physical and the scope of physical explanations of human behavior (Strawson 1992b: 139–142). All this material deserves further discussion and critical engagement. I hope that the present paper gives some indication of the rewards such engagement can bring.

References

Beglin, David. (2018) 'Responsibility, Libertarians, and the "Facts as We Know Them". A Concern-Based Construal of Strawson's Reversal'. *Ethics* 128, 612–625.

Beglin, David. (2020) 'Two Strawsonian Strategies for Accounting for Morally Responsible Agency'. *Philosophical Studies* 177, 2341–2364.

Bennett, Jonathan. (1980) 'Accountability'. In Zak van Straaten (ed.), *Philosophical Subjects*. Oxford, Clarendon, 14–47.

Blackburn, Simon. (1998) 'Relativization and Truth'. In Lewis Edwin Hahn (ed.), *The Philosophy of P. F. Strawson*. Chicago, Open Court, 151–167.

Campbell, Joe. (2017) 'P. F. Strawson's Free Will Naturalism'. *International Journal for the Study of Skepticism* 7, 26–52.

Chisholm, Roderick. (1966) 'Freedom and Action'. In Keith Lehrer (ed.), *Freedom and Determinism*. New York, Random House, 11–40.

Fischer, John Martin. (2014) 'Peter Strawson and the Facts of Agency'. In David Shoemaker and Neal A. Tognazzini (eds.), *Oxford Studies in Agency and Responsibility*. *Volume 2*. Oxford, Oxford University Press, 93–116.

Frankfurt, Harry. (1969) 'Alternate Possibilities and Moral Responsibility'. *Journal of Philosophy* 66, 829–839.

Hahn, Lewis Edwin (ed.). (1998) *The Philosophy of P. F. Strawson*. Chicago, Open Court.

Heyndels, Sybren and Benjamin De Mesel. (2018) 'On Shoemaker's Response-Dependent Theory of Responsibility'. *Dialectica* 72, 445–451.

Hieronymi, Pamela. (2020) *Freedom, Resentment, and the Metaphysics of Morals*. Princeton, Princeton University Press.

Lehrer, Keith. (1976) 'Can in Theory and Practice. A Possible Worlds Analysis'. In Myles Brand and Douglas Walton (eds.), *Action Theory*. Dordrecht, Reidel Publishing, 242–271.

McKenna, Michael. (2012) *Conversation and Responsibility*. Oxford, Oxford University Press.

Pears, David. (1998) 'Strawson on Freedom and Resentment'. In Lewis Edwin Hahn (ed.), *The Philosophy of P. F. Strawson*. Chicago, Open Court, 245–258.

Prasad, Rajendra. (1995) 'Reactive Attitudes, Rationality and Determinism'. In Pranab Kumar Sen and Roop Rekha Verma (eds.), *The Philosophy of P. F. Strawson*. New Delhi, Indian Council of Philosophical Research, 346–376.

Russell, Paul. (2017a [1992]) 'Strawson's Way of Naturalizing Responsibility'. In Paul Russell, *The Limits of Free Will. Selected Essays*. Oxford, Oxford University Press, 29–45.

Russell, Paul. (2017b [2004]) 'Moral Sense and the Foundations of Responsibility'. In Paul Russell, *The Limits of Free Will. Selected Essays*. Oxford, Oxford University Press, 67–96.

Sen, Pranab Kumar and Roop Rekha Verma (eds.). (1995) *The Philosophy of P. F. Strawson*. New Delhi, Indian Council of Philosophical Research.

Scanlon, T. M. (1988) 'The Significance of Choice'. In Sterling M. McMurrin (eds.), *The Tanner Lectures on Human Values*. Salt Lake City, University of Utah Press, 151–177.

Shoemaker, David. (2017) 'Response-Dependent Responsibility; or, A Funny Thing Happened on the Way to Blame'. *The Philosophical Review* 126, 481–527.

Shoemaker, David. (2020) 'Responsibility. The State of the Question'. *The Southern Journal of Philosophy* 58, 205–237.

Sosa, Ernest. (1998) 'P. F. Strawson's Epistemological Naturalism'. In Lewis Edwin Hahn (ed.), *The Philosophy of P. F. Strawson.* Chicago, Open Court, 361–369.

Strawson, P. F. (1963) 'Carnap's Views on Constructed Systems versus Natural Languages in Analytic Philosophy'. In Paul Arthur Schilpp (ed.), *The Philosophy of Rudolf Carnap.* LaSalle, Open Court, 503–518.

Strawson, P. F. (1980) 'P. F. Strawson Replies'. In Zak van Straaten (ed.), *Philosophical Subjects*. Oxford, Clarendon, 260–296.

Strawson, P. F. (1983) 'Liberty and Necessity'. In Nathan Rotenstreich and Norma Schneider (eds.), *Spinoza. His Thought and Work.* Jerusalem, Israel Academy of Sciences and Humanities, 120–129.

Strawson, P. F. (1985a) *Scepticism and Naturalism. Some Varieties.* London, Methuen.

Strawson, P. F. (1985b) *Analyse et métaphysique.* Paris, Vrin.

Strawson, P. F. (1992a) *Analysis and Metaphysics.* Oxford, Oxford University Press.

Strawson, P. F. (1992b) 'Freedom and Necessity'. In P. F. Strawson, *Analysis and Metaphysics*. Oxford, Oxford University Press, 133–142.

Strawson, P. F. (1995) 'Replies'. In Pranab Kumar Sen and Roop Rekha Verma (eds.), *The Philosophy of P. F. Strawson.* New Delhi, Indian Council of Philosophical Research, 398–433.

Strawson, P. F. (1998a) 'Reply to Simon Blackburn'. In Lewis Edwin Hahn (ed.), *The Philosophy of P. F. Strawson.* Chicago, Open Court, 168–173.

Strawson, P. F. (1998b) 'Reply to David Pears'. In Lewis Edwin Hahn (ed.), *The Philosophy of P. F. Strawson.* Chicago, Open Court, 259–262.

Strawson, P. F. (1998c) 'Reply to Ernest Sosa'. In Lewis Edwin Hahn (ed.), *The Philosophy of P. F. Strawson.* Chicago, Open Court, 370–372.

Strawson, P. F. (2008 [1962]) 'Freedom and Resentment'. In P. F. Strawson, *Freedom and Resentment and Other Essays*. London, Routledge, 1–28.

Strawson, P. F. (2011a [1956]) 'Construction and Analysis'. In P. F. Strawson, *Philosophical Writings*. Oxford, Oxford University Press, 30–38.

Strawson, P. F. (2011b [1967]) 'Analysis, Science and Metaphysics'. In P. F. Strawson, *Philosophical Writings*. Oxford, Oxford University Press, 78–90.

Strawson, P. F. (2011c [1983]) 'Liberty and Necessity'. In P. F. Strawson, *Philosophical Writings*. Oxford, Oxford University Press, 146–156.

van Inwagen, Peter. (1983) *An Essay on Free Will*. Oxford, Clarendon.

Watson, Gary. (2014) 'Peter Strawson on Responsibility and Sociality'. In David Shoemaker and Neal A. Tognazzini (eds.), *Oxford Studies in Agency and Responsibility. Volume 2*. Oxford, Oxford University Press, 15–32.