Strawson's Method in 'Freedom and Resentment'

(This is the accepted manuscript version, the published version can be found under the following link: https://doi.org/10.1007/s10892-019-09302-5. Please use the published version for references.)

Introduction

While P.F. Strawson's essay 'Freedom and Resentment' has had many commentators, discussions of it can be roughly divided into two categories. A first group has dealt with the essay as something that stands by itself in order to analyse Strawson's main arguments and to expose its weaknesses (see for example Watson 1987; Russell 1992; Fischer 2014). A second group of commentators has looked beyond 'Freedom and Resentment' by emphasizing its Humean, Kantian or Wittgensteinian elements (some recent examples include Campbell 2017; Coates 2017; De Mesel 2018; Bengtson 2019). Although both approaches have their own merits, it is too often forgotten that Strawson was an original thinker with his own views on the nature of philosophical problems and how to appropriately deal with them. The aim of this article is to remedy this forgetfulness and to make sense of 'Freedom and Resentment' from a *Strawsonian* perspective, by looking at Strawson's own views on philosophical methodology.

It is a peculiar fact about the secondary literature on Strawson that there is such a sharp division between discussions of his work on Kant, metaphysics and the philosophy of language and the reception of his 'Freedom and Resentment', which has given rise to a specialized field in the literature all on its own. While this may be partially explained by the special status of 'Freedom and Resentment' itself, the present article is motivated by (and aims to justify) the idea that a better understanding of Strawson's 1962 essay can be achieved by analysing it in the

light of his broader philosophical project. I will especially focus on his views on philosophical methodology and argue that 'Freedom and Resentment' offers a clear example of how those central ideas are concretely put to work.

One feature of Strawson's philosophy that has caused a lot of confusion and critique is his adherence to both transcendental ('Kantian') and naturalist ('Humean') anti-sceptical strategies. While he sometimes claims that certain beliefs the sceptic calls into question cannot be intelligibly doubted, on other occasions he seems satisfied to argue that it is simply pointless (and not necessarily unintelligible) to question certain beliefs we are inescapably committed to. Whereas the former amounts to a transcendental strategy, the latter expresses a commitment to a naturalistic strategy. I will argue (section 1 and 2) that there is sufficient evidence to believe that Strawson endorses both transcendental and naturalistic strategies throughout his career and, more importantly, that these strategies are fully compatible. This result will then be used to offer an interpretation of Strawson's main argumentative strategy in 'Freedom and Resentment'. More specifically, I will critically discuss both a traditional naturalistic as well as a recent transcendental reading of 'Freedom and Resentment' and argue that these interpretations only *partially* grasp Strawson's intentions (section 3 and 4). In order to understand 'Freedom and Resentment' from a truly *Strawsonian* perspective, he should be understood as combining both naturalist and transcendental elements. I conclude that this opens up the prospect of what I shall call 'full-blooded Strawsonianism about moral responsibility', outline the two main challenges for this view and briefly discuss some options future Strawsonians might consider in order to answer these challenges (section 5).

1. Description, explanation and imagination

Strawson distinguishes between three tasks a philosopher should carry out: a descriptive, an explanatory and an imaginative (creative) task (Strawson 1963; Strawson 2011a; Strawson 2011b).[1] These three tasks have their own specific goals, which can be summarized as follows:

> (D) [The Descriptive Task] The philosopher *describes* our conceptual scheme in order to *delineate* the bounds of sense.
>
> (E) [The Explanatory Task] The philosopher *explains* how our concepts are rooted in certain natural facts in order to emphasize the *non-arbitrariness* of our actual conceptual scheme.
>
> (I) [The Imaginative Task] The philosopher *tests* the bounds of sense by imagining certain variations of our conceptual scheme.

The first task of the philosopher is to *describe* our conceptual scheme. Strawson characterizes such description as a kind of *analysis* and distinguishes between a reductive and a connective analysis (or elucidation). The aim of a reductive analysis is to break down a concept into more simple components. Its goal is to explicitly state the necessary and/or sufficient conditions for correct application of the concept(s) under investigation. A connective analysis, on the other hand, does not break down a concept into more simple terms, but highlights the interconnections between the concepts that make up our conceptual scheme. This leads to a less narrow model

[1] Two further remarks are necessary. First, Strawson (2011b) himself distinguishes *five* different strands. Three of them are descriptive: either the description is (1) therapeutic (in order to dissolve philosophical paradoxes and perplexities), (2) carried out for its own sake or (3) part of a project that aims at laying bare the most *general* features of our conceptual scheme (which he calls 'descriptive metaphysics' (Strawson 1959)). I include (1)-(3) all under the 'descriptive task'. Second, Strawson sometimes characterizes the 'explanatory' and the 'imaginative (creative)' task as both different exercises of our imaginative powers. In order to avoid confusion and to emphasize their differences, I label them as 'explanatory' and 'imaginative' respectively.

of philosophical analysis; one that Strawson declares to be 'more realistic and more fruitful' (Strawson 2011b: 19).[2]

The aim of making explicit these conceptual interconnections is to self-consciously delineate 'the bounds of sense'. Examples of such connective analysis can be found throughout Strawson's writings. In *Individuals* (1959), Strawson investigates the fundamental conceptual structure that is necessary for our conceptual scheme to include the general category of particulars. He argues that the abilities to identify particulars and to re-identify them in cases of non-continuous observation are necessary conditions for our conceptual scheme in the first place and that this, in turn, depends on 'the possibility of locating the particular things we speak of in a single unified spatio-temporal system' (Strawson 1959: 38). In *The Bounds of Sense*, he takes Kant's major achievement to be that he offers an analysis of the concept of experience by showing how it interconnects with concepts such as self-consciousness, objectivity, space, time and causation. The most straightforward example is to be found in Kant's transcendental aesthetics, where he argues that space and time are a priori forms of intuition. Strawson takes Kant here to offer an *analysis* of the minimal structure our concept of experience (which is an interplay of both intuitions and concepts) must have in order to count as a conception of *experience* in the first place: any possible object of empirical awareness must occur *somewhere* and *somewhen*.[3] The aim of doing so is to delineate the 'limits to what we can conceive of, or make intelligible to ourselves, as a possible general structure of experience' (Strawson 1966: 1).

[2] The same sentiment is shared by Donald Davidson: 'The lesson I take to heart is this: however feeble or faulty our attempts to relate these various basic concepts to each other, these attempts fare better, and teach us more, than our efforts to produce correct and revealing definitions of basic concepts in terms of clearer or even more fundamental concepts' (Davidson 1996: 264).

[3] 'We are confronted with the thought of this link being so vital that it cannot be broken without nullifying the whole conception of experience' (Strawson 1966: 50). For Strawson's discussion of Kant's argument, see Strawson (1966: 47-71).

The second task of the philosopher consists in telling an *explanatory* story as to *why* we, as human beings, have the concepts we have. And '[t]o ask this', Strawson continues, 'is to ask to be shown how the nature of our thinking is rooted in the nature of the world and in our natures' (Strawson 2011a: 36).[4] Although Strawson does not further clarify the nature of this explanatory strategy, he argues that such an explanation contributes to a fuller understanding of these concepts (Strawson 2011b: 86, 90). The aim of the explanatory task is to remind us that there are certain *limits* as regards the conceptual variations we can intelligibly conceive of.[5] In *The Bounds of Sense*, he writes:

> [t]he set of ideas, or schemes of thought, employed by human beings reflect, of course, their nature, their needs and their situation. They are not static schemes, but allow of that indefinite refinement, correction, and extension which accompany the advance of science and the development of social forms. At the stage of conceptual self-consciousness which is philosophical reflection, people may, among other things, conceive of variations in the character of their own situations and needs and discuss intelligibly the ways in which their schemes of thought might be adapted to such variations. But it is no matter for wonder if conceivable variations are intelligible only as variations within a certain fundamental general framework of ideas, if further developments are conceivable only as developments of, or from, a certain general basis (Strawson 1966: 44).[6]

[4] See Strawson (2011b: 86) as well: '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'. Compare Strawson (1963: 516).

[5] One might object that the use of 'limits' is too strong here. After all, in the quoted passage, Strawson seems to refer to the idea of a 'starting place' from which changes in our conceptual scheme can be imagined, rather than the limits of what can be (intelligibly) conceived of. While it is true that setting a starting place is not the same as setting a limit, it is clear from other phrases Strawson uses that he is also concerned with the limits or boundaries of what can be intelligibly expressed (the 'bounds of sense'). I would like to thank an anonymous reviewer for pressing this point.

[6] Compare: '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 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' (Strawson 2008: 23).

While Strawson does not give direct examples of this explanatory strategy, it is clearly plausible that certain contingent biological and physical facts play an important role in determining the contours of our actual conceptual scheme, which is (plausibly) an extension of, and embedded in (though not reducible to), our basic natural responses. As will become clear, the clearest example of such an explanatory strategy in Strawson's writings occurs in his 'Freedom and Resentment'. Here he investigates the concept of moral responsibility (and related concepts) and explains how this area of our conceptual scheme is rooted in our natural responses towards one another (and oneself), what Strawson calls our 'reactive attitudes'.[7]

The primary aim of the explanatory strategy is to answer the worry that a mere description of our conceptual scheme does not explain why we should 'choose' to adopt one scheme rather than another. The answer to this worry is that it is not just an arbitrary *choice* whether or not we adopt the conceptual scheme we actually have: we are, instead, naturally committed to it. The starting point of our philosophical investigations is *our* conceptual scheme, the scheme we are naturally attached to.

This is not to say that we cannot imagine certain variations *within* this natural framework. Whereas the explanatory task of the philosopher consists in emphasizing the embeddedness of our concepts in 'the nature of the world and our natures', the *imaginative* (and third) task 'consists in imagining ways in which, without things other than ourselves being different from what they are, we might view them through the medium of a different conceptual apparatus' (Strawson 2011b: 87). While this imaginative task certainly has a positive function, Strawson is keen to emphasize the dangers of such an exercise of the imagination. The positive function of the imaginative task is that it allows us, by conceiving of conceptual variations, to appreciate the complexity of the conceptual scheme we actually have. The imaginative task is,

[7] In a recent article, Bengtson (2019) has given an excellent account of what I call the 'explanatory strategy' in 'Freedom and Resentment' by drawing a parallel between Wittgenstein's remarks on the primitive reactions that lie at the origin of our language-games and Strawson's emphasis on the reactive attitudes.

in this sense, secondary to the philosopher's descriptive task: Strawson is a *descriptive* metaphysician, not a *revisionary* metaphysician. Descriptive metaphysics is continuous with the notion of a connective analysis and 'does not differ [from it] in kind of intention, but only in scope and generality' (Strawson 1966: 9; 2011b: 88). While some metaphysical projects can be (charitably) interpreted as proposing a *change* in our conceptual scheme (and are thus revisionary), they often amount to making 'merely rudimentary mistakes' (Strawson 2011b: 87) that result from a misrecognition of our actual use of linguistic expressions, which 'remains his [the philosopher's] sole and essential point of contact with the reality which he wishes to understand, conceptual reality' (Strawson 2011b: 90). The descriptive and explanatory tasks of the philosopher are exactly there to prevent us from the metaphysical illusions that arise when the faculty of imagination goes on holiday. [8] The philosopher thus does *three* things:

> (DEI) [The Threefold Task] The philosopher *describes* our concepts, *explains* how these concepts are rooted in 'the nature of the world and our natures' and *imagines* variations as to how our conceptual scheme might be different within the framework that reflects our human nature, needs and situation.

2. Naturalism and transcendental arguments in Strawson

Hilary Putnam (1998) notes a tension at the heart of Strawson's philosophical project. When 'wearing his Kantian hat' (*Individuals* and *The Bounds of Sense*), Strawson states that sceptical challenges are unintelligible (incoherent or meaningless). When 'wearing his Humean hat' (*Skepticism and Naturalism*), he claims that sceptical challenges are 'idle' or 'vain' and that we

[8] Such a view might be called 'methodologically conservative' in the sense that it aims to protect certain first-order beliefs against possible revision. For an interesting discussion, see Callanan (2019).

are simply naturally committed to certain beliefs (e.g. the existence of an external world). Whereas the Kantian Strawson claims that certain beliefs the sceptic calls into question cannot intelligibly be doubted, the Humean Strawson argues that it is simply pointless to doubt certain beliefs that we are inescapably committed to. But the claim that a sceptical challenge is unintelligible is different from the claim that such a challenge is pointless. While the former plausibly amounts to a semantic critique, the latter seems to amount either to a description of a psychological fact ('we cannot help believing it') or to an evaluation of the challenge in terms of its utility ('it is in vain' or 'it is pointless'). How can these Kantian and Humean sympathies be reconciled?

It would be misleading to attribute to Strawson a 'naturalistic turn' in his later writings.[9] Such a turn in Strawson's thinking would allegedly occur in his *Skepticism and Naturalism* (from 1985), where he indeed suggests a ('Humean') naturalistic strategy to deal with the sceptic. But Strawson's early writings contain such Humean elements as well. There is, first of all, the short debate between Strawson and Wesley Salmon on the problem of induction (1957-1958).[10] Salmon criticizes Strawson's claim that induction cannot be given a general justification by arguing that this commits Strawson to the further claim that inductive beliefs are merely *conventional* and, in the end, a matter of arbitrary choice. Referring to Hume, Strawson responds that the impossibility of giving a general justification for our basic canons of induction is compatible with the claim that we are naturally committed to them and therefore do *not* have a choice as to whether we 'adopt induction' or not.

Secondly, there are his two early methodological essays ('Construction and Analysis' from 1956 and 'Analysis, Science, and Metaphysics' from 1967) in which Strawson distinguishes between the descriptive, explanatory and imaginative tasks (as discussed in

[9] See Stern (2003) for such a view. For criticism of this reading, see Putnam (1998); Glock (2003); Callanan (2011).
[10] See Salmon (1957) and Strawson (1958). For a discussion, see Putnam (1998: 273-277).

section 1). The philosopher's explanatory task bears strong resemblances to the Humean naturalistic strategy he endorses in his response to Salmon. By explaining how our concepts are embedded in, and extensions of, our natural lives, the philosopher emphasizes the non-arbitrariness of our conceptual scheme. While some conceptual variations are certainly conceivable, they are so only against a 'general background' that reflects our human nature, needs and situation.

Thirdly, 'Freedom and Resentment' was written in 1960, published in 1962 and occurs exactly between his two 'Kantian' works (*Individuals* in 1959 and *The Bounds of Sense* in 1966). This essay, which was published more than twenty years before *Skepticism and Naturalism*, contains his most prominent pursuit of the naturalistic strategy by explaining how the concepts that have to do with responsibility, obligation and punishment are embedded in our emotional responses to one another (and oneself).

Lastly, Strawson himself, in his reply to Putnam (at the end of his career in 1998), asserts that '[t]he doctrine of the incoherence of sceptical doubt is not clearly incompatible with Humean naturalism' (Strawson 1998: 289). And this seems fair enough: why would it not be possible that our basic beliefs are *both* inescapable, in the sense that we are 'naturally committed' to them, and that their denial is inconceivable (unintelligible or meaningless)? Furthermore, the threefold distinction between description, explanation and imagination can shed further light on the compatibility between the 'Humean' and the 'Kantian' Strawson. Whereas the Humean strategy of emphasizing our natural commitment to certain beliefs can be seen as the philosopher's task to *explain* how our concepts are rooted in certain natural facts, the Kantian strategy of rendering certain doubts as meaningless can be seen as the philosopher's task to describe our conceptual scheme in order to delineate the bounds of sense. The descriptive and the explanatory tasks are compatible, and so are Strawson's Humean and Kantian sympathies.

It may still be argued that, while Strawson indeed endorses both naturalistic and transcendental arguments throughout his career, his belief in the strength of transcendental arguments faded. This is sometimes attributed to Stroud's influential critique (1968) of the kind of transcendental arguments Strawson defended. In his insightful discussion of Strawson's anti-scepticism, Callanan (2011) shows that this is based on a misunderstanding of the goal of Strawson's transcendental arguments. Stroud's critique only counts as an argument against those transcendental arguments that aim at a *direct* refutation of the sceptic, by arguing for the existence of a certain fact that the sceptic doubts. However, Strawson clearly emphasizes that he does not wish to make claims about the obtainment of facts in the world. Instead, he argues that there is something incoherent about the sceptic's challenge in the first place: Strawson does not *directly* (by making a claim about the world) but *indirectly* refute the sceptic, by undermining the intelligibility of her challenge. This aligns well with Strawson's numerous comments that he is concerned with delineating the 'bounds of sense', the realm within which something can intelligibly said or thought, rather than with a factual investigation.

Other philosophers have come to the same conclusion. Hacker (2001), for example, claims that Strawson did not take his transcendental arguments to prove the *existence* of the external world. He then continues that it 'would be absurd to argue from conceptual connections in thought to existential truths about the world' and that Strawson instead is concerned with the incoherence of his (sceptic) opponent (Hacker 2001: 363-364). Glock notes as well that Strawson does not investigate '*de re* essences but the conceptual framework of our thought and experience' (2012: 397). While his *indirect* refutation of the sceptic (showing the sceptic's doubt to be senseless) faces its own challenges (see section 5), it cannot be criticized for failing to directly establish a fact in the world for the very simple reason that that is not its aim in the first place. Strawson is indeed a self-proclaimed (descriptive) metaphysician, but he

takes the metaphysician to be involved in conceptual work and not as someone who discovers some deep and hidden truths about the world.[11]

Even if Strawson would have somewhat lost his belief in the strength of transcendental arguments at the end of his career, the fact remains that 'Freedom and Resentment' was written exactly in between his two major works that employ such arguments.[12] In fact, 'Freedom and Resentment' is an excellent example of how Strawson's threefold distinction is concretely put to work: Strawson *describes* our moral responsibility concepts (he gives an account of 'the facts as we know them'), *explains* how these concepts are rooted in our reactive attitudes towards one another (and oneself) and *imagines* variations as to how our conceptual scheme might be different *within* the human framework that reflects our nature, needs and situation. The result of his threefold strategy will be that one of the underlying assumptions of the position defended by 'the Pessimist', i.e. the possibility of a scenario in which we completely abandon our responsibility practices, is, 'for us as we are, practically inconceivable' (Strawson 2008: 11).

As I will argue, in order to interpret 'Freedom and Resentment' from a truly *Strawsonian* perspective, it is necessary to do justice to both his naturalistic and transcendental tendencies. In the following two sections, I consider both a transcendental as well as a naturalist reading. A discussion of what both readings get right and where they go wrong then allows for a reading of 'Freedom and Resentment' that does justice to both its transcendental and naturalist elements.

[11] An example of someone in the contemporary debate on metametaphysics who defends a similar approach is Amie Thomasson (see section 5 for a more elaborate discussion).

[12] Again, this remains a big 'if'. As Strawson makes clear in his 1998 response to Putnam: 'Putnam proposes to explore a tension he finds between my Kantian and my Humean sympathies, and shows his own sympathies to be very much more in tune with the former than with the latter. Since, to revert to my case, the former are in fact very much stronger than the latter, it is pertinent to ask: What is at issue?' (Strawson 1998a: 288)

3. Coates's transcendental reading: from justification to description and explanation

In a recent article, Coates (2017) develops a transcendental reading of Strawson's main argument in 'Freedom and Resentment'. According to him, Strawson offers a transcendental justification of our responsibility practices. De Mesel (2018: 604) has reconstructed Coates's argument as follows:

(P1) Our ordinary interpersonal relationships are justified.

(P2) Our moral responsibility practices are necessary conditions for the possibility of ordinary interpersonal relationships.

(C) Our moral responsibility practices are justified.

De Mesel argues, against Coates, that Strawson does not *justify* our responsibility practices.[13] Instead, Strawson criticizes both the Pessimist, who gives a metaphysical (libertarian) justification of our practices, and the utilitarian Optimist, who justifies our practices in utilitarian terms. Both of them fail to see that our responsibility practices as a *whole* cannot, and need not be, 'externally justified' (Strawson 2008: 23). Inspired by Wittgenstein's discussion of scepticism in his *On Certainty*, De Mesel reformulates Coates's argument as follows:

(P1*) Our ordinary interpersonal relationships are beyond being justified or unjustified.

(P2*) Our moral responsibility practices are necessary conditions for the possibility of ordinary interpersonal relationships.

(C*) Our moral responsibility practices are beyond being justified or unjustified.

[13] Compare Campbell's critique (2017) of the 'core assumption' in the secondary literature on Strawson's 'Freedom and Resentment'.

De Mesel's reformulation indeed fits better with Strawson's (and Wittgenstein's) clear rejection of the justificatory question as such. Nevertheless, he fails to give a *positive* account of what it means for something to be 'beyond being justified or unjustified'. Given that the latter remark is rather mysterious, it is hard to see how the reformulation of a clear argument into a less perspicuous one would constitute an improvement on the original.

A positive account is available when we reconsider Strawson's threefold distinction between description, explanation and imagination. For if we read 'Freedom and Resentment' in the light of this distinction, we see that Strawson is indeed not giving a *justification* of our responsibility practices but rather *describes* our conceptual scheme and *explains* how these concepts are rooted in our responsibility practices and the reactive attitudes towards one another (and oneself). The need for a justification already presupposes that there is something missing in our actual practices that needs to be filled. The utilitarian Optimist will argue that our responsibility practices are externally justified by their social utility. The Pessimist will develop his libertarian account of contra-causal freedom. But Strawson here wishes to make clear that asking for such a justification is in some sense already a step too far.

The reference to Wittgenstein's *On Certainty* (1975, henceforth 'OC') is helpful in this respect. [14] In a response to Moore, Wittgenstein claims that it is meaningless to say that one either *knows* to have hands (as Moore wishes to say) or *doesn't know* to have hands (as the sceptic would want to say).[15] 'I know', after all, is essentially linked to the possibility of justification (OC §40, §91, §175, §243, §484, §504) and it is exactly because of this link between knowledge and justification that Wittgenstein argues that our basic certainties

[14] Strawson himself refers to Wittgenstein's as an important source of inspiration: '… and here he [Putnam] and I both seem to find ourselves in good company, indeed the best – viz. with the Wittgenstein of *On Certainty*' (Strawson 1998a: 289); see as well Strawson (1985: 14-21).

[15] 'I know that a sick man is lying here? Nonsense! I am sitting at his bedside, I am looking attentively into his face. – So I don't know, then, that there is a sick man lying here? Neither question nor the assertion makes sense' (OC §10). See as well OC §58, §498, §521.

('hinges' or 'hinge propositions') are *beyond* the status of either being justified or unjustified (OC §359). By claiming to *know* one has hands and thus to *defend* the basic certainty that we have hands, we implicitly accept a challenge to this certainty as legitimate. But a basic certainty, Wittgenstein argues, is exactly that which cannot be legitimately challenged. Therefore, saying that he *knows* he has hands already concedes too much to the sceptic, for it implicitly endorses the legitimacy of the sceptic's challenge.

Similarly, to *justify* our responsibility practices would be to implicitly accept a challenge to our responsibility practices as legitimate. To give an 'external justification' for our responsibility practices would already be a step too far, for it presupposes that it *lacks* a justification and that, if there is no external justification to be given, our responsibility practices are unjustified. But our responsibility practices are outside the realm of being either justified or unjustified. To treat them as being within the 'game of giving and asking for justifications' would be to commit a category mistake. The reason for this, and here Strawson resembles Wittgenstein again, is that the denial of our responsibility practices is not something we can intelligibly conceive of in the first place. The game of giving or lacking justifications occurs *within* the bounds of sense. The denial of something intelligible is not something that is unjustified, but something that is *un*intelligible. To argue that Strawson *justifies* our responsibility practices is to misunderstand the importance of his claim that '[a]s a whole, it neither calls for, nor permits, an external "rational" justification' (Strawson 2008: 23).

4. The naturalist reading: from psychological impossibility to inconceivability

Paul Russell (1992: 289-292) distinguishes between a rationalistic strategy and a naturalistic strategy in 'Freedom and Resentment'. The rationalistic strategy aims to show that the truth of

determinism would not give us a reason to repudiate our responsibility practices. The naturalistic strategy argues that it is in an important sense *impossible* for us to repudiate our responsibility practices. Consider the following two claims of the Pessimist:

(P1) If determinism is true, then we ought to repudiate our responsibility practices.

(P2) It is possible for us to repudiate our responsibility practices.

The rationalistic strategy argues against (P1); the naturalistic strategy argues against (P2). As Russell observes, both strategies are related. If the naturalistic strategy is effective in arguing for the impossibility of repudiating our responsibility practices, the rationalistic strategy would be misguided. For we ought only to repudiate our responsibility practices if it is assumed that we *can* do so. As Russell argues, '[f]rom the point of view of the naturalistic strategy such an approach [the rationalistic strategy] is wholly mistaken' (Russell 1992: 291). This is Strawson's view as well. He claims that

> such a question [the question whether we (rationally) ought to abandon our responsibility practices] should seem real only to one who had utterly failed to grasp the purport of the preceding answer, the fact of our natural human commitment to ordinary inter-personal attitudes. This commitment is part of the general framework of human life, not something that can come up for review as particular cases can come up for review within this general framework. (Strawson 2008: 13)

But the opposite sentiment is possible as well. Gary Watson claims that the naturalistic strategy (which he refers to as the 'psychological argument') is the one that 'seems irrelevant to the issues of the essay' (Watson 2014: 21) and that the rationalistic strategy (which he refers to as the 'normative argument') is the one that matters. He writes:

> What puzzles me, however, is the prior question of how this claim is dialectically relevant. So what if it's true? What prompts "pessimism" or scepticism is the worry

> or conviction that there is or might well be a general "theoretical ground" for abandoning our sense of responsibility […] To add that even if that conviction is correct, we couldn't adjust our lives accordingly (because we could not accept or "absorb" its truth) leaves these interlocutors' basic position completely intact. What work, then, is the psychological inescapability argument supposed to be doing? (Watson 2014: 25-26)

The most crucial question concerns the kind of 'possibility' that features in (P2). It is common in the secondary literature on 'Freedom and Resentment' to understand Strawson as stating that it is 'psychologically impossible' to repudiate our responsibility practices.[16] Given that this assumption is so widespread, one would expect that Strawson talks about 'psychological impossibility' himself. But this is not the case: in fact, there is no mention of 'psychological impossibility' (or inescapability) in 'Freedom and Resentment' at all. Nevertheless, Strawson does talk about our natural commitment to our reactive attitudes and our incapacity to give up on them. To speak of 'psychological impossibility' in this regard does not seem wholly inappropriate: indeed, it is part of the philosopher's explanatory strategy. The question is however whether the emphasis on our natural commitment to our reactive attitudes should be seen as an independent argument, or only as one part of a broader strategy.

If Strawson's main argument were mainly stating a psychological fact about us, it would indeed be vulnerable to Watson's claim that we could be (that is, if determinism were true) inescapably committed to something we have no reason to adopt. Certainly, this would only make things worse: we would be inescapably committed to an illusion. But it is implausible

[16] One exception is Neal A. Tognazzini, who draws from Frankfurt's work on volitional necessity to argue that 'Strawson wants to say more than merely that it's psychologically impossible for us to shake our belief in moral responsibility' (Tognazzini 2014: 684). An elaborate discussion of Tognazzini's article lies outside the scope of the paper, but his talk about volitional necessity as a constraint on what considerations we count as (practical) reasons seems to fit well with some of the remarks in Wittgenstein's *On Certainty* (see section 3) as well as Strawson's aim to delineate the bounds of sense within which questions about truth, falsity and justification may arise.

that Strawson would endorse such an argument. In fact, there is very good evidence available that he does not take this to be his line of reasoning at all. In a response to a similar remark from Ernest Sosa, Strawson writes (in 1998 and thus, again, at the end of his career):

> It is not merely a matter of dismissing the demand for a justification of one's belief in a proposition on the ground that one can't help believing it. That would be weak indeed. The position is, rather, that he demand for justification is really senseless (Strawson 1998b: 370).

The 'psychological impossibility' reading does not quite capture Strawson's intentions. But how should Strawson then be understood? In a crucial passage, Strawson writes:

> It does not seem to be self-contradictory to suppose that this might happen. So I suppose we must say that it is not absolutely inconceivable that it should happen. But I am strongly inclined to think that it is, for us as we are, practically inconceivable. (Strawson 2008: 11)

In this passage, Strawson talks about 'practical inconceivability' and not about 'psychological impossibility'. He distinguishes between 'absolute inconceivability' (which he identifies with self-contradictoriness) and 'practical inconceivability'. What does this distinction amount to? The clearest occurrence of a similar distinction in Strawson's writings can be found in *Analysis and Metaphysics*, where he talks about the philosopher's aim to lay bare conceptual connections through a connective analysis. He argues that certain interconnections between our concepts are non-contingent and distinguishes between two kinds of such *non-contingency*. According to the first sense of non-contingency, something may be non-contingent if its non-existence would imply a contradiction. Our reactive attitudes would thus be non-contingent because their non-existence (which would be the result of a complete repudiation of our practices) would somehow imply a contradiction. But as Strawson explains, the question of non-contingency

'can also be understood in a much more interesting, though less definite, sense' (Strawson 1992: 25). This kind of non-contingency has to do with the strong conceptual interconnections the descriptive philosopher lays bare: can we conceive of a notion of experiencing something without that something occurring in a spatio-temporal framework? Can we intelligibly conceive of a lion that speaks? Can we conceive of something being a billiard ball but not having the causal powers we normally take billiards balls to have? Can we intelligibly conceive of human beings *not* being involved in ordinary inter-personal relationships? Certainly, these are perhaps 'logical' (non-contradictory) possibilities; but Strawson's notion of conceivability is more demanding: there is a sense in which a lion that speaks is not a lion anymore, in which a billiard ball that doesn't have the causal powers normally associated with billiard balls is not a billiard ball anymore, in which human beings are not human beings anymore if they are not involved in ordinary inter-personal relationships. While such a notion of '(in)conceivability' may not be, as Strawson admits, very 'definite', it seems clear that a reading of 'Freedom and Resentment' from a *Strawsonian* perspective must do justice to this ineliminable element of his main strategy.

5. Full-blooded Strawsonianism and its main challenges

Consider again the Pessimist's two main claims:

(P1) If determinism is true, then we ought to repudiate our responsibility practices

(P2) It is possible to repudiate our responsibility practices

Let a 'P1-strategy' be a strategy that aims at undermining the truth of (P1). Let a 'P2-strategy' be a strategy that aims at undermining the truth of (P2). The naturalistic reading of 'Freedom

and Resentment' makes the following two claims. First, Strawson's main strategy is a 'P2-strategy'. Second, he does so by arguing that it is 'psychologically impossible' to repudiate our responsibility practices. I agree with the first claim but disagree with the second. For Strawson, a 'P1-strategy' is indeed misguided, given that (P2) is a necessary condition for (P1). Russell is correct in maintaining that, from the point of view of the 'P2-strategy', the 'P1-strategy' is fundamentally misguided. Furthermore, Watson is right when he states that the claim that it is psychologically impossible to abandon our responsibility practices does not change anything about the justificatory status of our responsibility practices. To the contrary, it would make things even worse: we would be inescapably committed to an illusion (a practice that is fundamentally unjustified). The naturalistic reading goes wrong, however, because the claim that we are inescapably committed to our responsibility practices should not be seen as an independent argument. By making the claim, Strawson wishes to emphasize certain constraints on the *intelligibility* of a scenario in which we repudiate our responsibility practices. His claim that one of the Pessimist's underlying assumptions is *unintelligible* is the 'transcendental' element in 'Freedom and Resentment'.

Coates's recent reading of 'Freedom and Resentment' is correct in claiming that Strawson offers a transcendental argument against the sceptic, but incorrect in claiming that Strawson offers a transcendental *justification* of our responsibility practices. Such a reading does not take seriously enough, as Campbell (2017) and De Mesel (2018) have recently argued, Strawson's rejection of the justificatory question as such. Instead, Strawson *describes* our moral responsibility concepts, *explains* how they are rooted in our natural responses to one another (and oneself) in order to put constraints on the *intelligibility* of a scenario in which we imagine a total decay of our responsibility practices.

Therefore, I suggest not only that Strawson must be read as combining transcendental and naturalist arguments throughout his entire career, but that his 'Freedom and Resentment'

must be read as combining both transcendental and naturalist elements. Instead of either giving a justification or merely stating a psychological fact, he aims to undermine the intelligibility of the radical change that would lead to an abandonment of our responsibility practices. Giving up our responsibility practices would mean that we would have to give up our reactive attitudes; giving up our reactive attitudes would be to give up our ordinary inter-personal relationships; giving up our inter-personal relationships would be to give up our humanity. This, Strawson claims, is unintelligible.

Let us label 'full-blooded Strawsonianism' any position that argues along the lines outlined in this essay. It is clear that, in order to count as a legitimate contender in the debate on free will and moral responsibility, such a position faces many further challenges. In the remainder of this section, I will briefly discuss the two main challenges and consider possible responses a 'full-blooded Strawsonian' might consider.

First, something more has to be said about the notion of '(un)intelligibility'. While Strawson is rather unclear about how it should be construed, future Strawsonians must do better. One way to answer this challenge is by combining Strawsonianism about moral responsibility with contemporary deflationism in metametaphysics. Deflationists typically adhere to Carnap's distinction between internal and external questions. Amie Thomasson, one of the most recent defenders of the deflationary position, elucidates the internal/external distinction in terms of a *use*/*mention*-distinction. Internal questions are questions where the relevant term (for example 'proposition', 'number', 'electron' in questions such as 'Do propositions exist?', 'Do numbers exist?' or 'Do electrons exist'?) is being *used*, i.e. it is governed by the standard rules of use that make them part of the linguistic framework in which these terms are introduced in the first place. An external question is a question where the term is not governed by its standard rules of use and therefore only counts as being *mentioned*. If this is the case, however, an external question in which the term is *not* governed by its standard rules of use is either (1) meaningless,

for if the term is not governed by its actual rules of use it makes the term and the question in which it occurs meaningless, or (2) must be (more charitably) interpreted as a proposal to adopt a new set of rules. Given that the standard rules of use allow internal questions to be answered by simple conceptual or empirical means and given that the metaphysician does not take these existence questions to be settled in this straightforward manner, she must be asking a question ('Are there numbers?', 'Are there propositions?') in a different (external) sense and this renders the question either meaningless or simply different from the ordinary, internal question.

Similarly, the Strawsonian strategy can be seen as a way to describe the actual rules we follow when we *use* a responsibility-vocabulary in our actual practices. *Within* our responsibility practices, we adopt a certain set of rules that determine whether someone is morally responsible if certain conditions occur. Important criteria for being morally responsible may include an agent's quality of will, his knowledge of possible consequences, whether he did so intentionally etc. These conditions will be fulfilled in some cases and not in others; what is important is that there is a framework of rules according to which we can fairly reliably decide whether or not someone is morally responsible. In such circumstances, the term 'moral responsibility' is being *used*, that is it is governed by the standard rules of application and inferring that constitutes its meaning in the first place. However, if we depart too far from the actual set of rules that govern the responsibility-talk in our actual practices, we end up no longer *using* a term but only *mentioning* it. Such mentioning can then be charitably interpreted as a proposal to adopt a new set of rules; or, in Strawson's terms, as a revisionary metaphysical project.

Thomasson's deflationism already bears strong affinities to Strawson's philosophical project, which would make such a dialogue even more plausible. First, she observes that 'Carnap's interest lay in formal, technical languages, while I will be more concerned with existence questions that are (at least apparently) asked in ordinary English' (2015: 44).

Secondly, she argues that an emphasis on how our basic concepts have evolved throughout our natural history may help to counter a possible objection which states that the choice of *which* framework of rules we choose to adopt is arbitrary (2015: 43). Thirdly, she develops a positive account of metaphysics as being involved in conceptual work. It should be clear that these three points fit well with the account of Strawson's methodology outlined in this essay.[17]

The second challenge is that something more needs to be said about 'the ability to do otherwise'. It may be argued that the fact about whether or not an agent is able to do otherwise is a *relevant* fact within our responsibility practices and thus constitutes an important criterion to decide whether or not someone is morally responsible *within* our actual responsibility framework.[18] But the Strawsonian has a response here as well: what reason is there to believe that the expression 'the ability to do otherwise' in the mouth of the incompatibilist is the same one as our ordinary notion? In our actual practices, we take there to be certain rules for appropriately inferring from the obtainment of a set of conditions (physical fitness, the absence of manipulation, mental health etc.) that someone 'could have done otherwise'. But we never take the 'full list of physical laws' and the 'full description of the state of the universe at a certain time in the past' as relevant considerations from which we may appropriately infer that someone has the ability to do otherwise or not.[19] This seems to be Strawson's view as well:

> 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,*

[17] Strawson discusses Carnap's internal/external distinction in his 'Carnap's Views on Constructed Systems Versus Natural Languages in Analytic Philosophy'. While this would require a much more elaborate discussion, Strawson's talk about 'internal' and 'external' justifications of our responsibility practices might be influenced by Carnap's internal/external distinction as well. The idea of such a historical connection comes from Filip Buekens, who defends such an interpretation of 'Freedom and Resentment' in some of his lectures at KU Leuven.

[18] For a forceful critique along these lines, see Fischer (2014: 98-101).

[19] This view shows some affinities with contextualism in epistemology. Hawthorne has defended a similar contextualist account of free will and even refers to Strawson as one of his main influences (Hawthorne 2001: 72).

> *however complex*, to his acting otherwise"? I find it difficult, as others have found it difficult, to accept this equivalence. The common judgement of this form amounts rather to the denial of any sufficient natural impediment *of certain specific kinds or ranges of kinds*. (Strawson 1992: 136-137)

If some set of rules of use are different than the standard rules that govern our actual practices, then we may, following Thomasson's internal/external distinction, appropriately ask whether the incompatibilist here still *uses* a familiar term at all or has merely *mentioned* the term without adopting the standard rules that give life to it in the first place. The Pessimist's assumption that a scenario in which all our responsibility-talk is fundamentally mistaken is then either (1) *unintelligible* because it has departed too far from the actual rules according to which we, internal to our actual practices, treat each other as morally responsible, or (2) must be understood as a revisionary proposal that has not discovered anything potentially disturbing about moral responsibility but has merely stipulated a new meaning for it.

6. Conclusion

I argued, first, that Strawson combines both transcendental and naturalistic arguments throughout his career and that these strategies are compatible. A useful way to capture this compatibility is by taking seriously Strawson's threefold methodological distinction between the philosopher's descriptive, explanatory and imaginative tasks. This result was then applied to his 'Freedom and Resentment', which should be understood as combining transcendental and naturalistic elements as well. A merely transcendental or a merely naturalistic reading only

partially capture the full potential of Strawson's argumentative strategy. I concluded with a brief discussion of the challenges ahead for future Strawsonians.[20]

Bibliography

Allais, Lucy. 2014. Freedom and Forgiveness. In *Oxford Studies in Agency and Responsibility, Volume 2,* 33-63. Oxford: Oxford University Press, 2014.

Bengtson, Audun Benjamin. 2019. Responsibility, Reactive Attitudes and Very General Facts of Human Nature. *Philosophical Investigations* 42: 281-304.

Campbell, Joe. 2017. P.F. Strawson's Free Will Naturalism. *International Journal for the Study of Skepticism* 7(1): 26-52.

Callanan, John. 2011. Making Sense of Doubt: Strawson's Anti-Scepticism. *Theoria* 77: 261-278.

Callanan, John. 2019. Methodological Conservatism in Kant and Strawson. *British Journal for the History of Philosophy* 27: 422-442.

Coates, D. Justin. 2017. Strawson's modest transcendental argument. *British Journal for the History of Philosophy* 25: 799-822.

Davidson, Donald. 1996. The Folly of Trying to Define Truth. *The Journal of Philosophy* 93: 263-278.

[20] Special thanks to Stefan Rummens for invaluable comments on earlier drafts of this paper. I am grateful to Audun Bengtson, Arnold Burms, Stefaan Cuypers, Benjamin De Mesel, Bart Pattyn, Amie Thomasson and audiences at Dartmouth College and Trinity College Dublin for helpful feedback. I would also like to thank the editor and two anonymous reviewers of *The Journal of Ethics* for their insightful comments.

De Mesel, Benjamin. 2018. Are our moral responsibility practices justified? Wittgenstein, Strawson and justification in 'Freedom and Resentment'. *British Journal for the History of Philosophy* 26: 603-614.

Fischer, J.M. 2014. Peter Strawson and the Facts of Agency. In *Oxford Studies in Agency and Responsibility, Volume 2,* 93-116. Oxford: Oxford University Press.

Glock, Hans-Johann. 2003. Strawson and Analytic Kantianism. In *Strawson and Kant*, ed. H.J. Glock, Oxford: Oxford University Press.

Glock, Hans-Johann. 2012. 'Strawson's Descriptive Metaphysics'. In *Categories of Being: Essays on Metaphysics and Logic*, ed. L. Haaparanta and H.J. Koskinnen,, 391-419. New York: Oxford University Press.

Hacker, P.M.S. 2001. 'On Strawson's Rehabilitation of Metaphysics'. In *Wittgenstein. Connections and Controversies*, 345-370. Oxford: Oxford University Press.

Hawthorne, J. 2001. 'Freedom in Context'. *Philosophical Studies* 104: 63-79.

Putnam, Hilary. 1998. Strawson and Skepticism. In *The Philosophy of P.F. Strawson,* ed. Lewis Edwin Hahn, 273-287. Chicago: Open Court.

Russell, Paul. 1992. Strawson's Way of Naturalizing Responsibility. *Ethics* 102: 287-302.

Salmon, Wesley. 1957. Should We Attempt to Justify Induction?. *Philosophical Studies* 8: 33-48.

Stern, Robert. 2003. On Strawson's Naturalistic Turn. In *Strawson and Kant*, ed. H.J. Glock, Oxford: Oxford Unviersity Press.

Strawson, P.F. 1958. On Justifying Induction. *Philosophical Studies* 9: 20-21.

Strawson, P.F. 1959. *Individuals. An Essay in Descriptive Metaphysics.* London: Routledge.

Strawson, P.F. 1963. Carnap's Views on Constructed Systems versus Natural Languages in Analytic Philosophy. In *The Philosophy of Rudolf Carnap*, ed. P.A. Schilpp, 503-518. La Salle: Open court.

Strawson, P.F. 1966. *The Bounds of Sense*. *An Essay on Kant's Critique of Pure Reason.* London: Routledge.

Strawson, 1985. *Skepticism and Naturalism: Some Varieties.* New York: Columbia University Press.

Strawson, P.F. 1992. *Analysis and Metaphysics*. *An Introduction to Philosophy.* Oxford: Oxford University Press.

Strawson, P.F. 1998a. Reply to Hilary Putnam. In *The Philosophy of P.F. Strawson,* ed. Lewis Edwin Hahn, 288-292. Chicago: Open Court.

Strawson, P.F. 1998b. Reply to Ernest Sosa. In *The Philosophy of P.F. Strawson,* ed. Lewis Edwin Hahn, 370-372. Chicago: Open Court.

Strawson, P.F. Freedom and Resentment. Chapter 1 in *Freedom and Resentment and Other Essays*, 1-28. Abingdon: Routledge, 2008.

Strawson, P.F. 2011a. Construction and Analysis. In *Philosophical Writings*, 30-38. Oxford: Oxford University Press.

Strawson, P.F. 2011b. Analysis, Science and Metaphysics. In *Philosophical Writings*, 78-91. Oxford: Oxford University Press.

Stroud, Barry. 1968. Transcendental Arguments. *The Journal of Philosophy* 65: 241-256.

Thomasson, Amie. 2015. *Ontology Made Easy.* Oxford: Oxford University Press.

Tognazzini, Neal. 2014. Reactive Attitudes and Volitional Necessity. *Journal of Value Inquiry* 48: 677-689.

Watson, Gary. 1987. Responsibility and the Limits of Evil. In *Responsibility, Character, and the Emotions,* ed. Ferdinand Schoeman, 256-86. Cambridge: Cambridge University Press.

Watson, Gary. 2014. Peter Strawson on Responsibility and Sociality. In *Oxford Studies in Agency and Responsibility, Volume 2*, 15-32. Oxford: Oxford University Press.

Wittgenstein, Ludwig. 1975. *On Certainty*. ed. G.E.M. Anscombe and G.H. von Wright. trans. D. Paul and G.E.M. Anscombe. Oxford: Blackwell.