**Being and Holding Responsible**

**Reconciling the Disputants Through a Meaning-Based Strawsonian Account**

This version of the article has been accepted for publication, after peer review (when applicable) but is not the Version of Record and does not reflect post-acceptance improvements, or any corrections. The Version of Record is available online at: https://link.springer.com/article/10.1007/s11098-021-01737-7. Use of this Accepted Version is subject to the publisher’s Accepted Manuscript terms of use https://www.springernature.com/gp/open-research/policies/acceptedmanuscript-terms.

*Abstract*

A fundamental question in responsibility theory concerns the relation between being responsible and our practices of holding responsible. ‘Strawsonians’ often claim that being responsible is somehow a function of our practices of holding responsible, while others think that holding responsible depends on being responsible, and still others think of being and holding responsible as interdependent. Based on a Wittgensteinian reading of Strawson, I develop an account of the relation between being and holding responsible which respects major concerns of all parties in this debate. I characterize the way in which being responsible depends on holding responsible as *genealogical*, and the way in which holding responsible depends on being responsible as *justificatory*. I show how my account cuts across received ways of carving up the debate, and how it allows for all the kinds of fallibility about moral responsibility that are worth wanting.

What is the relation between being morally responsible and our practices of holding people morally responsible? This is a fundamental question in contemporary responsibility theory (for a recent overview of foundational questions in the field, see Shoemaker 2020). ‘Strawsonians’ about responsibility often claim that being responsible is somehow a function of our practices of holding responsible (Shoemaker 2017, Wallace 1994, Watson 1987 and 2014), while others think that holding responsible depends on being responsible (Brink and Nelkin 2013, Fischer and Ravizza 1993, Smith 2007, Zimmerman 1988 and 2016), and still others think of being and holding responsible as interdependent (De Mesel 2018a, Manata 2017, McKenna 2012). This paper has two aims. First, I hope to make a historical and exegetical contribution by presenting what I take to be P.F. Strawson’s view on the relation between being and holding responsible, taking the whole of his *oeuvre* into account and paying particular attention to his account of meaning and Wittgenstein’s influence on it.[[1]](#footnote-1) I will call my reconstruction of Strawson’s view Strawson’s view. Second, I hope to make a substantive philosophical contribution by showing how Strawson’s view, as opposed to existing ‘Strawsonian’ views, respects major concerns of all parties in the debate about being and holding responsible.

I begin by taking seriously Strawson’s (2008a: 24) aim in ‘Freedom and Resentment’ to recover ‘from the facts as we know them a sense of what we mean, i.e. of *all* we mean, when, speaking the language of morals, we speak of desert, responsibility, guilt, condemnation, and justice’. How is it possible to recover from the facts as we know them a sense of what we mean by something? In section one, I outline Strawson’s Wittgensteinian answer to this question. The core idea is that the meaning of an expression is constituted by the rules for its use, and these rules originate in empirical regularities which have acquired a normative status. In section two, I discuss the consequences of Strawson’s Wittgensteinian account of meaning for the relation between being and holding responsible. On Strawson’s view, being responsible depends on holding responsible *and* the other way around, but the dependence relations are different. I characterize the way in which being responsible depends on our practices of holding responsible as *genealogical*, and the way in which holding responsible depends on being responsible as *justificatory*. I explain how Strawson’s view challenges a prominent view about Strawson’s methodology, according to which Strawson turns away from conceptual issues in ‘Freedom and Resentment’ in favor of a description of actual moral psychology.

In section three, I compare Strawson’s view to Shoemaker’s (2017) related, response-dependent view of responsibility. I argue that the former has at least two important advantages over the latter, and that it cuts across Shoemaker’s way of carving up the debate about being and holding responsible, because it does not unequivocally fall under any of the categories specified by Shoemaker. I will show that, depending on one’s understanding of the distinction between response-dependence and response-independence, Strawson’s view can be labeled as either a sophisticated version of response-dependence, or as neither response-independent nor response-dependent.

In section four, I discuss a possible objection, namely that Strawson’s view does not allow for the possibility of collective fallibility about responsibility. I distinguish several ways of understanding the idea of collective fallibility, and I argue that Strawson allows for all the kinds of collective fallibility worth wanting. I conclude by indicating how Strawson’s view respects major concerns of all parties in the debate about being and holding responsible.

1. The Background: Aspects of Strawson’s Wittgensteinian Account of Meaning

In this section, I will present some aspects of Strawson’s Wittgensteinian account of meaning that will become relevant in what follows. My aim is neither (1) to defend this account, nor (2) to provide a full overview, nor (3) to defend the claim that Strawson’s account of meaning is broadly Wittgensteinian. The latter is rather uncontroversial and recognized by Strawson himself (Strawson 1992: Ch. 8, Strawson 2008b, Strawson 2008c: Ch. 4), although its plausibility depends on one’s reading of Wittgenstein.[[2]](#footnote-2)

How can we recover from the facts a sense of what we mean by something? Wittgenstein (2009a: §43) famously held that, for a large class of cases, the meaning of an expression is its use. This can be understood in two ways. First, it may be thought that explaining what an expression means is describing how *it is actually used,* describing *regularities*. But this idea cannot account for the crucial distinction between actual and correct use (Strawson 2008c: 63). The notion of correctness is related to the notion of a standard of correctness or *rule.* According to the second and more plausible reading of ‘meaning is use’, to explain what an expression means is to explain how *it is to be used,* to specify a (set of) rule(s) for its use (Strawson 2011a: 211)*.* Expressions can be used correctly or incorrectly, in agreement or in conflict with the rules.

Although there is a categorial distinction between the descriptive level of regularities and the normative level of rules, the two are normally related. The rules for the use of an expression are regularities which have ‘hardened’ into norms (Wittgenstein 1978: 324). They are ‘descriptive-turned-normative items, originating in regularities already existing in linguistic exchanges’ (Bangu 2021: 452). Rules are not facts, but rules can be *recovered from* the facts because they originate in factual regularities. The existence of regularities in use, of agreement in use, is a precondition for the existence of semantic rules, but these rules do not normally make reference to agreement. If people did not use ‘circle’ to refer to a kind of plane figure, we would not have the rule that a circle is a plane figure. Although this rule originates in agreement, it does not state *that* we agree in our use of ‘circle’ or that we use ‘circle’ in this or that way. Rather, it states how ‘circle’ *is to be used.* In Wittgenstein’s (1978: 365) words, agreement is a framework condition or ‘precondition of our language-game, it is not affirmed in it’.

To explain the meaning of an expression is to provide (a) rule(s) for its use. These rules are fluid, flexible and unsystematic (Wittgenstein 2009a: §142, Strawson 2011a: 226, 231). They are present in our practices. Whether something is a rule is not determined by its form, but by its *role*. Semantic rules are cited and referred to in explanations of meaning, in justifications of use, in answers to questions like ‘How do you know that this is an *x*?’, in teaching and correcting others. Thus, ‘the bishop moves diagonally’ may be used to *describe* what is regularly happening or to state a *rule* of chess, by reference to which we explain the use of the bishop in chess, justify (give reasons for) our use of ‘bishop’, answer how we know that this is the bishop, teach others to play or correct them.

Semantic rules for the use of expressions are in some ways similar to the syntactic rules of school grammar: they originate in how we speak, but they do not state *that* we speak thus-and-so. They are hardened regularities endowed with normative force. Wittgenstein (2009: §122) and Strawson (1992: 7) both use ‘grammar’ to refer to the set of semantic rules for the use of an expression, and they call these rules ‘grammatical’. Grammatical rules are rooted in our practices (‘We frame rules in the light of our study of our practice’, Strawson 2011a: 231; ‘There cannot be law where there is no custom, or rules where there are not practices’, Grice and Strawson 1956: 153), they grow out of how we speak and act (Strawson 1992: 7).

Grammatical rules for the use of ‘*x*’ determine what *counts as x*. Only plane figures count as circles. Wittgenstein’s (2009a: §497) suggestion that grammatical rules are arbitrary is directed against the idea that they can be justified as correct by reference to reality. It is tempting to think that ‘a circle is a plane figure’ is a description of how things are in reality. But that a circle is a plane figure was never discovered by observation or experience. Instead, ‘a circle is a plane figure’ is a rule for the use of ‘circle’ which lays down what is to be called ‘circle’. We could have had different rules, originating in different regularities. It is a contingent fact that we have the rules that we have (that ‘circle’ refers to a plane figure), but that does not make the connections spelled out in the rules any less necessary (a concept not connected to the concept of a plane figure is not the concept of a circle).

It is perhaps misleading to connect grammar with arbitrariness. We associate arbitrariness with choice or whim, with what can be easily dispensed with or altered, but grammar is not arbitrary in *these* senses. A grammatical rule is not a whimsical stipulation, because it is firmly rooted in our practices. These practices could have been different in the sense that reality does not dictate what they have to be, but the world we live in and our human nature make some practices and conceptual possibilities more attractive, one could say more natural, than others. Our concepts are dependent on empirical facts for their usability and usefulness. These ‘very general facts of nature’ (Wittgenstein 2009b: §§365-366, Strawson 2008b: 170) are ‘not curiosities …, but facts that no one has doubted and which have escaped notice only because they are always before our eyes’ (Wittgenstein 2009a: §415). An example of a general fact of nature is the fact that the colors of things do not change every few seconds. If they did, our color concepts would be unusable. Grammar is non-arbitrary in a sense because, given how the world is and who we are, some grammatical structures serve us better than others.

That some rules and concepts serve us better than others explains, at least partly, why some regularities harden into rules and others do not.[[3]](#footnote-3) (I am aware of the fact that ‘serve us better’ can be understood in many different ways, and I am using it in a deliberately open-ended way.) We distinguish between green and blue in the way we do, because having a unitary concept ‘grue’ instead of having both ‘green’ and ‘blue’ would not allow us to draw some of the distinctions that we want to draw or make it more cumbersome to draw these distinctions. That we find it important to distinguish between green and blue has to do with our nature (without our natural ability to distinguish between green and blue, the distinction would not play the role that it plays for us) and empirical facts about the world (if nothing were either green or blue, the distinction would lose its importance).

It is the philosopher’s task to provide an overview of the rules for the use of expressions, to elucidate our *concepts* (Wittgenstein 2009: §122, Strawson 1975: 13)*.[[4]](#footnote-4)* Concepts are abstractions from the uses of expressions: if expression ‘*x*’ in language L is used according to the same rules as expression ‘*y*’ in language M, ‘*x*’ and ‘*y*’ express the same concept. The philosopher elucidates *our* concepts, the ones that we have, not the ones that we could or should have had. There could have been different regularities, different rules and different concepts. If it were a rule, originating in regularities of use, that a circle is a piece of furniture, the concept of a circle determined by *that* rule would not be *our* concept of a circle.

According to Strawson (and here he differs from Wittgenstein), the philosopher’s task is not only to elucidate our concepts, but also to explain

[…] 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. A form which propositions exemplifying this strand in philosophy may often take, is the following: if things (or we) were different in such-and-such ways, then we might lack such-and-such concepts or types of discourse. […] It might reasonably be maintained […] that full understanding of a concept is not achieved until this kind of enquiry is added to the activities of comparing, contrasting and distinguishing which I mentioned first. (Strawson 1963: 515-516)

I will come back to this difference between Strawson and Wittgenstein in the next section.

I have been writing about the *meaning* of *expressions*. Instead of asking about the meaning of ‘circle’, one might ask what a circle essentially *is.* Wittgenstein (2009a: §371, §373) famously wrote that ‘Essence is expressed in grammar’ and that ‘Grammar tells what kind of object anything is’. Strawson (1997: 4) agrees. They are not *only* talking about expressions: ‘the question of what imagination essentially is, is as much about the word “imagination” as my question [what “imagination” means]’. The question ‘What is imagination?’ *also* ‘asks for the clarification of a word; but it makes us expect a wrong kind of answer’ (Wittgenstein 2009a: §370). According to Wittgenstein and Strawson, there is no philosophically important difference between an explanation of the meaning of ‘imagination’ and an explanation of what imagination essentially is. In both cases, we need an overview of the rules for the use of ‘imagination’.

1. Being and Holding Responsible: Strawson’s View

*2.1 Responsibility and the Facts as We Know Them*

The aspects of Strawson’s account of meaning outlined above help to understand Strawson’s (2008a: 24) aim in ‘Freedom and Resentment’ to ‘recover from the facts as we know them a sense of what we mean […] when, speaking the language of morals, we speak of desert, responsibility, guilt, condemnation, and justice’. This is about *meaning* and *language*, but because essence is expressed in grammar, we can say that Strawson wants to investigate what moral responsibility *is.* In a methodology paper published shortly after ‘Freedom and Resentment’, he writes that ways of using ordinary language

[…] may give rise to philosophical problems; that the concepts employed in these [ordinary] activities [such as pleading in law courts] may […] call for philosophical clarification. How do we conceive of *responsibility*? (Strawson 1963: 505; Strawson’s italics)

The concept of responsibility calls for clarification, which can only be provided by looking closely at our ways of using language: ‘the actual use of linguistic expressions remains his [the philosopher’s] sole and essential point of contact with the reality which he wishes to understand, conceptual reality’ (Strawson 2011b: 90).

Strawson explains what we mean by ‘responsibility’. To explain the meaning of an expression is to provide an overview of (at least some of) the rules for its use. These rules originate in our practices of holding responsible, which exhibit regularities. One such regularity is that we tend not to adopt reactive attitudes towards people if we believe that they did not know what they were doing. This regularity has hardened into the rule that people are not to be held responsible if they did not know what they were doing. Because it is *pro tanto* appropriate to hold agents responsible iff they are responsible, this means that people are not responsible if they did not know what they were doing.[[5]](#footnote-5) That people are not responsible if they did not know what they were doing is a rule for the use of ‘responsible’.[[6]](#footnote-6) We explain why people are not responsible by reference to their ignorance, justify our use of ‘responsible’ by reference to the fact that people knew what they were doing, teach others what responsibility is by reference to the knowledge condition, correct them if they hold people responsible who did not know what they were doing, etc.

Strawson’s aim is not just to articulate a sense of what we mean by responsibility, it is to recover such a sense *from the facts as we know them.* This is accomplished in two ways. First, he shows how the rules for being responsible are rooted in our practices of holding responsible, in particular in what we do and say (‘He did not know’, etc.) when we refrain from holding people responsible. It is a contingent fact that we have these rules and not others (that ‘responsible’ is not to be used if agents did not know what they were doing). But that does not make the connections spelled out in the rules any less necessary, because the rules determine what responsibility *means.* If there were a rule that ‘responsible’ is not to be used if agents have red hair, this rule would specify *another* concept.

Second, our practices of holding responsible do not come from nowhere. They could have been different, but some general facts about us and the world we live in make it understandable that we have them. Strawson (2008a: 5) insists on what he calls a ‘central commonplace’, namely ‘the very great importance that we attach to the attitudes and intentions towards us of other human beings’. We are concerned about whether the actions of others reflect attitudes of good or ill will or indifference, and that basic human concern seems to lie at the root of our practices of holding responsible. The point of mentioning this is to explain why we have the practices and concepts that we have, to show that there is an important sense in which our grammar, although it cannot be justified by reference to reality, is not arbitrary. The central commonplace makes our practices intelligible, but it does not necessitate them. We could have translated our concern into a different set of practices, into different regularities, rules and concepts.

The first way of recovering a sense of what we mean by responsibility from the facts as we know them might be called ‘conceptual’: Strawson shows how our *concept* of responsibility is rooted in our *practices* of holding responsible. The second way might be called ‘naturalistic’: Strawson shows how our *practices* of holding responsible are rooted in *natural facts* about us and the world we live in. The conceptual (more Wittgensteinian) and naturalistic (more Humean) strands are complementary, and they are both indispensable to an accurate understanding of Strawson’s view. My focus in this paper is on the conceptual strand, because my main aim is to shed light on the relation between being responsible and our practices of holding responsible. By contrast, many commentators have emphasized the naturalistic, or ‘social naturalistic’ strand (see, for instance, Beglin 2018, Hieronymi 2020, Watson 2014), while neglecting the conceptual one. My point is not that these commentators are mistaken; on the contrary, they highlight extremely important aspects of Strawson’s view. Rather, the problem is that naturalistic views are incomplete, and that, consequently, they cannot explain some of the things that Strawson’s view can explain. First, their focus is on the relation between our practices of holding responsible and our natural concerns, and they do not have much to say on the relation between being responsible and our practices of holding responsible. Second, as I will show in section 2.4, they leave Strawson’s emphasis on meaning, and the frequent occurrence of ‘concept(ual)’ in ‘Freedom and Resentment’, unexplained.[[7]](#footnote-7) Third, as I will show in section 4, Strawson’s view makes it possible to improve on naturalistic responses (such as Beglin 2018) to a worry formulated by Todd (2016): can Strawsonians avoid the consequence that, if we were to start blaming children, they would be blameworthy?

*2.2 Justificatory and Genealogical Dependence*

My remarks about Strawson’s project in ‘Freedom and Resentment’ have been sketchy and incomplete, highlighting only aspects of it that are relevant to my argument, but they allow me to give a preliminary account of the relation between being responsible and holding responsible. First, those who think that holding responsible depends on being responsible emphasize that it is only appropriate to hold agents responsible if, and because (or, as it is sometimes put, in virtue of the fact that) they *are* responsible (Brink and Nelkin 2013, Fischer and Ravizza 1993, Smith 2007, Zimmerman 1988 and 2016). I take this to be a platitude that any plausible account of the relation between being and holding responsible must respect, but those who claim that being responsible depends on holding responsible struggle to accommodate it. Nothing in what I have said excludes this platitude, however, and I explicitly endorse it. This simply *is* a sense in which holding responsible depends on being responsible. Let us call this relation between holding and being responsible a relation of *justificatory dependence*: we can justify why we *hold* an agent responsible by reference to the fact that they *are* responsible.

Second, those who think that being responsible depends on our practices of holding responsible *also* have a point. The dependence relation is not justificatory: we do not justify ascriptions of responsibility by reference to practices of holding responsible, but by reference to the rules for being responsible. We do not say that an agent is responsible because we are disposed to hold them responsible or because they merit being held responsible; rather, we justify ascriptions of responsibility by saying that the agent knew what they did, had alternatives, and so on. But, as we have seen, these rules do not come from nowhere, they are present in our practices of holding responsible and have grown out of regularities. Bangu (2021) has recently offered a convincing account of how, according to Wittgenstein, grammatical rules emergefrom our practices, of how custom makes for grammatical law, and calls this view *genealogical.* The term is meant to capture the idea that our practices are the place where our rules come from. Bangu (2021: 453) explains that

[…] although […] ‘genealogical’ typically connotes ‘historical’, it is important to realize that the kind of ‘origins’ I have in mind […] need not be of a literally historical nature, in the sense of being found to have happened in the past. As understood here, the act of hardening-into-a-rule happens (or happened) not only ‘in the past’, but is reaffirmed in the present.

Following Bangu’s terminology, I will characterize the relation between being responsible and our practices of holding responsible as a relation of *genealogical dependence.*

Justification and genealogy are forms of explanation, justificatory dependence and genealogical dependence are forms of explanatory dependence. It has often been remarked that genesis and justification should not be conflated. To treat what belongs to the *context of formation* of a concept as if it belonged to its *context of justification* is to commit the so-called ‘genetic fallacy’. Queloz (2020: 2022) argues that genealogies can ‘give us reasons to cultivate certain concepts […], but they cannot give us reasons for *particular* actions or beliefs – they cannot tell one, say, whether to be truthful to someone on a particular occasion’. A genealogical account can yield a ‘pragmatic vindication’ for having a concept, ‘by showing that, given certain needs and purposes and given certain facts about us and our environment, certain ways of going on are rational because *apt* responses’ (Queloz 2020: 2014; see also Queloz 2021a and Queloz 2021b). Strawson seems to do precisely this: by showing how our concept of responsibility is rooted in practices of holding responsible, practices that are in turn rooted in a basic human concern, he gives us (a) reason(s) for having that concept. It is a concept worth having because having it constitutes an apt response to a basic human concern. In Strawson’s words (see the quote in section one), showing ‘the natural foundations of our conceptual apparatus in the way things happen in the world, and in our own natures’ contributes to the understanding of concepts in the sense that it may explain ‘why it is that we have such concepts […] as we do’.

Importantly, however, although genealogies can justify our having certain concepts, they cannot justify applications of these concepts in particular cases. The ways in which we justify having the concept of responsibility are not the ways in which we justify the application of ‘responsible’ in particular cases, and to explain why we have the concept of responsibility is not to explain what ‘responsibility’ means. The ability to explain what ‘circle’ means does not imply the ability to explain why we have the concept of a circle. Genealogical considerations do not normally enter into explanations of meaning (unless the application of a concept is justified by reference to a formation process).

*2.3 Epistemic or Metaphysical Priority?*

Some might be tempted to conclude that, if Strawson’s view is correct, holding responsible depends on being responsible. This conclusion would be compatible with the idea that our practices of holding responsible are *epistemically prior* to being responsible. As McKenna (2012: 39-41) explains, looking at our practices of holding responsible may help us to understand what responsibility is, ‘much as we might look to activity in a cloud chamber to understand the nature of an electron’. The explanatory priority of holding responsible in the order of knowledge, however, does not entitle us to draw symmetrical conclusions about the order of being: ‘it is the electron that is more fundamental or basic in the order of being in comparison with the goings on in the cloud chamber’. If Strawson’s point is merely epistemic, as ‘recover’ suggests, it is compatible with the fact that being responsible is prior to holding responsible in the order of being. Those who think that holding responsible is prior to being responsible need more than epistemic priority (McKenna 2012: 40, Shoemaker 2017: 486).

The kind of genealogical priority at issue here is priority in the order of being. Our practices of holding responsible are more than a suitable context of discovery for understanding the nature of being responsible. It is not just the case that, given our epistemic capacities, our practices of holding responsible happen to be the best place to look for the rules of being responsible, which would be like saying that our practices of actual language use happen to be the best place to look if we want to know what our grammatical rules (in the school grammar sense) are. Our practices are not just the best context of discovery, they are also the *context of* *formation* of our concepts. They constitute the framework within which our concepts are born and have their lives. Our practices determine the rules for being responsible, what being responsible *is,* but cloud chambers do not determine what electrons are. This is a way in which holding responsible is prior to being responsible in the order of being.

It may be asked whether holding or being responsible is *metaphysically* prior. The answer depends on one’s conception of metaphysics. Because Wittgenstein tends to equate explanations of what ‘*x*’is to explanations of what *x* means, and because genealogical considerations do not normally enter into explanations of meaning, the way in which being responsible depends on our practices of holding responsible is not metaphysical in his sense of the term. Things are different for Strawson, who suggests (see previous section) that a metaphysical explanation of *x* includes not only (1) an explanation of its meaning (by reference to by the rules for its use), but also (2) an explanation of why we have the concept of *x* (how it is rooted in our practices, how these practices are in turn rooted in facts about us and the world we live in). Genealogical explanations are explanations of the second kind.

We can now catch a glimpse of how disputants in the debate about being and holding responsible could be reconciled (or, at least, how they could be brought closer to each other). Is being responsible dependent on holding responsible or *vice versa*? Strawson’s view is: both, but the metaphysical dependence relations are different. Being responsible is genealogically dependent on our practices of holding responsible, holding responsible is justificatorily dependent on being responsible. This is, in a sense, what has been called an *interdependence thesis* (McKenna 2012), but ‘interdependence’ is somewhat misleading. It suggests that being responsible is dependent on holding responsible in the way in which holding responsible is dependent on being responsible. It carries in its wake the idea of symmetrical dependence and the charge of circularity. McKenna (2012: 54) sees this. Some have tried to show that, although there may be a circle, it is virtuous rather than vicious (De Mesel 2018a, Manata 2017). In Strawson’s view, however, there *is* no circle, because the way in which being responsible depends on holding responsible differs from the way in which holding responsible depends on being responsible.

*2.4 A Note on Strawson’s Methodology and Two Possible Objections*

Although my description of Strawson’s project has been sketchy, it makes clear why a prominent view of his methodology is inaccurate. According to McKenna and Russell (2016: 5),

Strawson’s strategy in ‘Freedom and Resentment’ involves turning away from conceptual issues about the analysis of ‘freedom’ and ‘responsibility’ and taking a closer look at what actually goes on when we *hold* a person responsible. That is to say, his methodology depends less on conceptual analysis and more on a descriptive account of actual moral psychology.

It is true that Strawson provides a descriptive account of actual moral psychology, but this does not involve turning away from conceptual issues. On the contrary, Strawson attends to our nature and the practices that express it *in the service of* his conceptual analysis of responsibility, as the frequent occurrence of ‘concept(ual)’ in his paper suggests. Conceptual analysis is a matter of looking closely at our ways of language use, embedded in practices which, though contingent, have roots in our nature, and full understanding of a concept is not achieved until we understand not only what it means, but also why we have it. Strawson (2008a: 24) criticizes the intellectualist inclination of conceptual analysts to go beyond the facts as we know them, not the enterprise of conceptual analysis as such.

Before going on, I will briefly respond to two possible objections. First, my Strawsonian account of the relation between being and holding responsible is based on a *general* account of meaning. It may be thought to underplay the differences between the concept of responsibility and other (kinds of) concepts. I do not want to deny these differences, but I believe that we do not need to appeal to them in order to bring the disputants in the debate about being and holding responsible closer to each other.

Second, one may not accept a Wittgensteinian/Strawsonian account of meaning. Such a normative account is difficult to square, for example, with an externalist semantics in the tradition of Putnam and Kripke. Defending it in detail is beyond the scope of this paper, but it is important to note that normative accounts of meaning are among the standard options in contemporary philosophy of language (for an overview, see Glüer and Wikforss 2018). Contemporary accounts close to the one I have offered have been defended by Robert Brandom and Paul Horwich. Even if one does not accept them, my proposal is philosophically significant. I claim that Strawson’s conception of the relation between being and holding responsible is a corollary of his account of meaning. Thus, the reason why some find Strawson’s story unconvincing may have to do with his account of meaning, or with his view on the relation between meaning and essence, rather than with his specific remarks about being and holding responsible.[[8]](#footnote-8)

1. Response-Dependence or Response-Independence?

Shoemaker (2017) has recently developed a response-dependent view of responsibility. In this section, I will argue that Strawson’s view has at least two important advantages over Shoemaker’s, and that it cuts across Shoemaker’s way of carving up the debate about being and holding responsible. I will show that, depending on one’s understanding of the distinction between response-dependence and response-independence, Strawson’s view can be labeled as either a sophisticated version of response-dependence, or as neither response-independent nor response-dependent.

Shoemaker distinguishes three possible ways of accounting for the relation between being and holding responsible. The first is dispositional response-dependence: an agent is responsible for something if and only if people are typically disposed to hold them responsible for it.[[9]](#footnote-9) The second is response-*in*dependence: holding an agent responsible is appropriate if and only if, and in virtue of the fact that, they are antecedently responsible. What makes them responsible is independent of our practices of holding responsible. Shoemaker argues, rightly in my view, that dispositional response-dependence and response-independence are implausible. The third account, which Shoemaker defends and takes to be a development of Strawson’s position, is fitting response-dependence: someone is responsible for something if and only if, and in virtue of the fact that, they merit being held responsible for it (Shoemaker 2017: 508).

There are some problems with Shoemaker’s account, as I have pointed out in work with Sybren Heyndels (Heyndels and De Mesel 2018). I will summarize two of these problems and indicate how Strawson’ view can deal with them. The first problem concerns one of Shoemaker’s reasons for rejecting dispositional response-dependence. He observes that it would be strange to say, ‘I deem you responsible because most people would in fact (be disposed to) hold you responsible’. But would it not be equally strange to say, ‘I deem you responsible because you merit being held responsible’? Shoemaker’s reason for rejecting dispositional response-dependence seems to count against his own view as well.

I can see no way out of this problem for Shoemaker (Heyndels and I discuss a possible response on behalf of Shoemaker, but we do not find it convincing; see Heyndels and De Mesel 2018: 446-447). But what about Strawson’s View? Would it not be strange to say, ‘I deem you responsible because, both now and in the past, our moral community has regularly treated some relevant fact as a responsibility-grounding fact to the point that this regularity is now a constitutive rule’?[[10]](#footnote-10) This would indeed be a very strange thing to say, but Strawson’s view, in virtue of the central distinction between genealogical and justificatory explanation, has the resources to explain why it would be strange. Recall that, although genealogical explanations can justify our having certain concepts, they cannot justify applications of these concepts in particular cases (section 2.2). If asked why we have the concept of responsibility, an answer referring to the practices of our moral community (‘because, now and in the past, …’) would be apt. But if we are asked why someone is responsible, if we are asked to *justify* the application of ‘responsible’ in a particular case, a different kind of explanation is called for. In justifying why we apply ‘responsible’ in a particular case, we have to refer to the rules that constitute the meaning of ‘responsible’. We then say things like, ‘I deem you responsible because you knew what you were doing’ (referring to the rule that people are not responsible if they did not know what they were doing, see section 2.1), ‘I deem you responsible because you could have acted otherwise’, and so on. Thus, the distinction between genealogical and justificatory explanation helps to explain why ‘I deem you responsible because, both now and in the past, …’ would be strange.

The second problem with Shoemaker’s account has to do with the normative-descriptive distinction. Shoemaker rejects dispositional response-dependence for its lack of a normative element: people are not responsible in virtue of the fact that they are *typically* held responsible, but in virtue of the fact that they *merit* being held responsible. Without this normative element, there is no possibility of typical responses being mistaken, a possibility that any plausible account of responsibility has to allow for. However, the move from typical responses to merited responses is problematic for Shoemaker. First, it invites the further question of what *makes* certain responses merited, and this question leads back to the idea that what makes it appropriate to hold an agent responsible is that the agent *is* responsible, an idea that Shoemaker associates with response-*in*dependent views. Second, the move from typical responses to merited responses protects Shoemaker’s account from the charge that a change in typical responses would result in a change in responsibility status: if we typically held the intellectually disabled responsible, they would not, in virtue of that fact, *be* responsible. Our responses may change or could have been different, but what is fitting remains constant. But if radical changes in our responses do not affect responsibility, then how is responsibility response-dependent?

Strawson’s view does not have these problems. The idea that what makes it appropriate to hold an agent responsible is that the agent *is* responsible is accommodated by the justificatory dependence of holding responsible on being responsible. The genealogical dependence of being responsible on holding responsible leaves room for a sense in which a radical change in typical responses couldaffect responsibility. It would not change the *facts* (such as the fact that the intellectually disabled are not responsible in our sense of the term), but it could change the *rules,* and thereby the *meaning,* of being responsible. I will explain this in the next section.

Strawson’s view cuts across received ways of carving up the debate about being and holding responsible, because it does not unequivocally fall under any of the categories specified by Shoemaker (dispositional response-dependence, response-independence, fitting response-dependence). It respects Shoemaker’s best reasons for rejecting dispositional response-dependence (we do not justify ascriptions of responsibility by reference to our practices) and response-independence (which does not recognize that our practices make knowledge, etc. the conditions for being responsible), while avoiding problems with Shoemaker’s fitting response-dependence view.

It may be asked whether, on Strawson’s view, responsibility is response-dependent or response-independent.[[11]](#footnote-11) The answer depends on how ‘response-dependent’ and ‘response-independent’ are understood. The first option is to understand the distinction as exhaustive. A property must be either response-dependent or response-independent, and the possibility of its being neither is excluded *a priori.* A property is response-dependent if it somehow (genealogically, justificatorily, or otherwise) metaphysically depends on our responses, and response-independent if it does not metaphysically depend on our responses in any way. If we understand the distinction between response-dependence and response-independence in this way, then Strawson’s view is response-dependent. It will then be a sophisticated response-dependent view, which is able to accommodate a major concern (see the platitude mentioned in section 2.2) of response-independence theorists.

The second option is to understand the distinction as non-exhaustive. The idea here is that responsibility is response-independent if being responsible does not metaphysically depend on holding responsible (the same idea of response-independence as on the exhaustive view), and response-dependent if being responsible metaphysically depends on holding responsible, while holding responsible does not metaphysically depend on being responsible. If the distinction is understood in this way, there is room for views that are neither response-independent nor response-dependent. A view, such as Strawson’s view, according to which being responsible somehow metaphysically depends on holding responsible *and* holding responsible somehow metaphysically depends on being responsible, would then be neither response-independent (because being depends on holding in some way) nor response-dependent (because holding depends on being in some way).

I wish to remain neutral between these two ways of understanding the distinction between response-dependence and response-independence. If I have been able to show that Strawson’s view is distinct from other views in the debate and has some important advantages over these views, it does not seem to matter much whether it is labeled as a sophisticated version of response-dependence, or as neither response-dependent nor response-independent. The key point is that Strawson’s view respects major concerns of all parties in the debate (see also the conclusion of this paper).

1. The Possibility of Collective Mistakes

In this section, I will discuss a possible objection to Strawson’s view, namely that it does not seem to allow for the possibility of collective fallibility about responsibility. If we accept that our practices of holding responsible determine the rules for being responsible, do we then not exclude the possibility that our practices might be radically mistaken? If there were a practice of holding the profoundly intellectually disabled responsible, would they then *be* responsible? (Shoemaker 2017: 497) Animal trials have been popular, which suggests that we once thought it appropriate to hold animals responsible (McGeer 2019: 308), but we do not want to say that they *were* responsible. How does Strawson’s view account for the fact that our ways of treating people with dyslexia (‘You should just try harder!’) were not just different but wrong? (McKenna 2012: 50)

Strawson’s view allows for several kinds of collective fallibility about responsibility. First, although our practices determine the rules for being responsible, we may be mistaken about what these rules are. There is a difference between the rules we observe and the rules we *think* we observe, between practice and the theory of our practice. We may think that we observe *this* rule, while in fact we observe *that* one. This kind of mistake is often the result of being insufficiently attentive to our practices, the kind that Strawson’s opponents in ‘Freedom and Resentment’, prone to overintellectualize the facts, tend to make. However, this kind of fallibility does not help much with the examples of dyslexia, animal trials and the intellectually disabled, because these are examples in which our *practices* go wrong, rather than examples in which our theories are mistaken.

Second, we can be mistaken about whether our rules apply in particular (sets of) cases. People might in general observe the rule that knowledge of what one is doing is necessary for being responsible, but be mistaken about the capacities for knowledge of people who are in fact intellectually disabled. Or they might in general observe the rule that one is not responsible for what one cannot help, but be mistaken about the fact that some people cannot help reading poorly. These are mistakes about the facts, about whether the rules apply, and such mistakes are by no means excluded by Strawson’s view.

Semantic rules are flexible, fluid, unsystematic and subject to change. Perhaps we can conceive of people who hold animals or the intellectually disabled responsible while knowing all the facts about them, so that their mistake is not a factual mistake, but the result of having different rules. Perhaps they have a rule stating that whoever or whatever causes something bad is responsible for it. It is important to recall, at this point, that rules are constitutive of meaning. If the rules are different, then what is determined by the rules will also be different. If their rules are verydifferent from ours, what these rules determine will not be the same thing as what our rules determine. Radically different rules will not determine what *moral responsibility* is, they will determine *another concept.*[[12]](#footnote-12)If people were to hold animals or the profoundly intellectually disabled responsible, and if they were to do it systematically while knowing all the relevant facts, the rules governing their practices would be radically different from our rules, and they would have radically different concepts. Of course, the animals could be *called* ‘morally responsible’, but ‘morally responsible’ would in that case express a concept that is different from ours. Because it is only appropriate to hold agents responsible if they *are* responsible, and because agents would not *be* responsible in this scenario, it would not be *appropriate* to hold animals and the intellectually disabled responsible. This is the gist of a recent Strawsonian answer to the problem of collective fallibility which I developed in collaboration with Heyndels (De Mesel and Heyndels 2019; for a similar account, see Bengtson 2019).

A version of this problem has been forcefully put forward by Todd (2016): how could Strawsonians avoid the consequence that, if we were to *hold* children responsible, they would *be* responsible? Beglin’s (2018) answer is, basically, that holding children responsible would not be an expression of our basic concerns. Given our actual basic concerns (‘this is an empirical matter’, Beglin 2018: 622), it seems highly unlikely or even practically inconceivable that, as we are now, we would start holding children responsible. That is correct, I think, and it is probably the best answer that is available to naturalistic Strawsonians. But Todd might not be satisfied. He could ask, ‘What if we were different? What if we had different basic concerns, reflected in the practice of holding children responsible? Would children then *be* responsible?’ Beglin (2018: 622) writes that ‘it is not obvious that rational agents who relate to each other in fundamentally different ways […], would or should operate with the same conceptual equipment as we do’, but he does not explain this point. I believe that that is not a coincidence: as explained in section 2.1, naturalistic views such as Beglin’s focus on the relation between concerns and practices, and they do not have much to say about the relation between practices and concepts. Strawson’s view, by contrast, can explain Beglin’s point: having radically different practices means that there will be radically different regularities and radically different rules. We might then *call* children ‘morally responsible’, but ‘responsible’ would express a concept that is different from ours, because it would be constituted by radically different rules originating in radically different practices.

So one response to the problematic practices of animal trials and holding the intellectually disabled responsible is to say that these practices would not change the fact that animals and the intellectually disabled *are* not responsible, because the concept of being responsible is constituted by *our* rules, not by whatever the rules in very different communities happen to be. More radically, I believe that even the conceivability of holding animals and the intellectually disabled responsible, which is taken for granted in the examples, has to be questioned. Strawson (2008a: 10, 15) emphasizes that holding responsible is *essentially* a reaction to perceived quality of will. If there is no mistake about the fact that animals and the profoundly intellectually disabled cannot display quality of will, then they *cannot* be held responsible. Our reactions to them will not *count as* ways of holding responsible (in our sense of holding responsible).

This response might seem merely verbal, for surely we once treated animals in ways that are strikingly similar to some of the ways in which we treat those whom we hold responsible. We punished them, for instance, for doing things they could not help doing. McGeer takes this to suggest that we thought of animals as responsible beings, but that is historically incorrect. The practice of punishing animals was neither grounded in factual mistakes about their capacities (people did not think that they had a capacity to know what they were doing, for instance), nor in a radically different concept(ion) of responsibility (people did not think that knowledge, etc. were irrelevant to responsibility). Rather, it was the corollary of a strikingly different way of thinking of *punishment,* as Friedland (2012) has masterfully demonstrated. *Our* concept of punishment is linked to the idea of responsibility (although the link is probably less straightforward than one might think, see Kneer and Machery 2019), but those who punished animals thought of punishment as a kind of ritual in which communities could overcome what had happened. The point of animal trials was to determine to what extent the accused animals were causally involved in what happened, not to determine whether they were morally responsible for it.

According to Strawson’s view, the meaning of ‘responsibility’ is determined by *our* rules, but why would ourrules be privileged or correct? Two senses of ‘correct’ have to be distinguished. First, according to Wittgenstein and Strawson, semantic rules cannot be mistaken in the sense of not corresponding with how things really are (although we can be mistaken about what our rules are). Grammar is arbitrary. We cannot discover that a circle is not a plane figure, because the rule that a circle is a plane figure stipulates what a circle *is,*what is to be called a ‘circle’. This is the sense in which semantic rules cannot be the wrong ones. Second, some rules and concepts may have their origin in, or keep in place or foster practices that are prudentially or morallyor otherwisewrong. These practices are wrong, and we could say that there is something wrong with the rules and concepts that structure them, but not in the sense of their not corresponding to reality. We could say, for example, that there is something wrong with the rule that whoever causes something bad is responsible, because having that rule sustains practices of punishing the intellectually disabled, which is morally wrong. We might want to correct these wrongs by changing our practices through the construction of rules and concepts that are less likely to foster them, but this would be a project of conceptual engineering rather than a Wittgensteinian/Strawsonian project of conceptual analysis. The latter focuses on analyzing our concepts as they are rather than on making them more precise, more likely to further our goals, etc.[[13]](#footnote-13)

Let me summarize how Strawson’s view allows for the possibility of collective mistakes about moral responsibility. We cannot be collectively mistaken in the sense that our rules cannot fail to correspond with an independent reality. But (1) we can be mistaken about what our rules are. (2) We can also make factual mistakes, a result of which being that we can be mistaken about whether our rules apply in particular cases. We did not know that people now diagnosed with dyslexia could not help reading poorly, and so we did not apply our rule that those who cannot help doing what they do are not responsible for it. (3) Our semantic rules sometimes have their origin in, or keep in place or foster practices that are prudentially or morally or otherwise wrong. In these cases, we might say that our rules are the wrong ones, but not in the sense that they do not correspond to reality. I believe that these kinds of collective fallibility are all the kinds worth wanting for an account of the relation between being and holding responsible.

1. Conclusion

Like ‘Freedom and Resentment’, this paper is ‘intended as a move towards reconciliation; so is likely to seem wrongheaded to everyone’ (Strawson 2008a: 2). It challenges widespread views about Strawson’s methodology and the character of his account of responsibility. I believe that Strawson’s view, based on a Wittgensteinian/Strawsonian account of meaning, respects fundamental concerns (though not necessarily all of them; see footnote 8) of all parties in the debate about being and holding responsible. The platitude that it is only appropriate to hold agents responsible if and because they are responsible, recognition of which is of fundamental importance for those who think that holding responsible depends on being responsible, is captured by the idea that holding responsible is justificatorily dependent on being responsible. The insight that being responsible somehow depends on our practices of holding responsible, which is of fundamental importance for many Strawsonians, is captured by the idea that being responsible is genealogically dependent on holding responsible. The idea that being responsible depends on holding responsible and *vice versa,* which is of fundamental importance for those who hold an interdependence thesis, has been elucidated in a way that avoids charges of circularity. Moreover, Strawson’s view has important advantages over Shoemaker’s related view and allows for the possibility of collective mistakes about moral responsibility.

References

Baker, G.P. and P.M.S. Hacker. (2005) *Wittgenstein. Understanding and Meaning.* Oxford, Blackwell.

Baker, G.P. and P.M.S. Hacker. (2009) *Wittgenstein. Rules, Grammar and Necessity.* Oxford, Wiley-Blackwell.

Bangu, Sorin. (2021) ‘Hard and Blind. On Wittgenstein’s Genealogical View of Logical Necessity’. *Philosophy and Phenomenological Research* 102, 439-458.

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

Bengtson, Audun. (2019) ‘Responsibility, Reactive Attitudes and Very General Facts of Human Nature’. *Philosophical Investigations* 42, 281-304.

Boghossian, Paul. (2006) *Fear of Knowledge. Against Relativism and Constructivism.* Oxford, Clarendon.

Brink, David and Dana Nelkin. (2013) ‘Fairness and the Architecture of Responsibility’. In David Shoemaker (ed.), *Oxford Studies in Agency and Responsibility. Volume 1.* Oxford, Oxford University Press, 284-313.

De Mesel, Benjamin. (2018a) ‘Moral Responsibility and the Moral Community. Another Reply to Zimmerman’. *The Journal of Ethics* 22, 71-92.

De Mesel, Benjamin. (2018b) ‘Are Our Moral Responsibility Practices Justified? Wittgenstein, Strawson and Justification in ‘Freedom and Resentment’’. *British Journal for the History of Philosophy* 26, 603-614.

De Mesel, Benjamin. (2021) ‘P.F. Strawson Was Neither an Externalist Nor an Internalist About Moral Responsibility’. *European Journal of Philosophy* 29, 199-214.

De Mesel, Benjamin and Sybren Heyndels. (2019) ‘The Facts and Practices of Moral Responsibility’. *Pacific Philosophical Quarterly* 100, 790-811.

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

Friedland, Paul. (2012) *Seeing Justice Done. The Age of Spectacular Capital Punishment in France.* Oxford, Oxford University Press.

Glock, Hans-Johann. (2009) ‘Meaning, Rules, and Conventions’. In David K. Levy and Edoardo Zamuner (eds.), *Wittgenstein’s Enduring Arguments.* Abingdon, Routledge, 156-178.

Glüer, Katrin and Åsa Wikforss. (2018) ‘The Normativity of Meaning and Content’. In Edward Zalta (ed.), *The Stanford Encyclopedia of Philosophy,* Spring 2018 Edition.

Grice, H.P. and P.F. Strawson. (1956) ‘In Defense of a Dogma’. *The Philosophical Review* 65, 141-158.

Hacker, P.M.S. (2019) *Wittgenstein. Meaning and Mind.* Oxford, Wiley-Blackwell.

Heyndels, Sybren. (2019) ‘Strawson’s Method in ‘Freedom and Resentment’’. *The Journal of Ethics* 23, 407-423.

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.

Kneer, Markus and Edouard Marchery. (2019) ‘No Luck for Moral Luck’. *Cognition* 182, 331-348.

Manata, Paul. (2017) ‘The Conversational Theory of Moral Responsibility and McKenna’s Interdependence Thesis’. *The Philosophical Quarterly* 67, 61-83.

McGeer, Victoria. (2019) ‘Scaffolding Agency. A Proleptic Account of the Reactive Attitudes’. *European Journal of Philosophy* 27, 301-323.

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

McKenna, Michael and Paul Russell (eds.). (2016) *Free Will and Reactive Attitudes.* Abingdon, Routledge.

Queloz, Matthieu. (2020) ‘How Genealogies Can Affect the Space of Reasons’. *Synthese* 197, 2005-2027.

Queloz, Matthieu. (2021a) ‘The Self-Effacing Functionality of Blame’. *Philosophical Studies* 178, 1361-1379.

Queloz, Matthieu. (2021b) *The Practical Origins of Ideas.* Oxford, Oxford University Press.

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. Fault Lines in the Foundations’. *Southern Journal of Philosophy* 58, 205-237.

Sider, Theodore. (2011) *Writing the Book of the World.* Oxford, Clarendon.

Smith, Angela M. (2007) ‘On Being Responsible and Holding Responsible’. *The Journal of Ethics* 11, 465-484.

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 (IL), Open Court, 503-518.

Strawson, P.F. (1975) ‘Semantics, Logic and Ontology’. *Neue Hefte für Philosophie* 8, 1-13.

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

Strawson, P.F. (1997) *Entity and Identity and Other Essays.* Oxford, Clarendon.

Strawson, P.F. (2008a [1962]) ‘Freedom and Resentment’. In his *Freedom and Resentment and Other Essays.* London, Routledge, 1-28.

Strawson, P.F. (2008b [1954]) ‘Wittgenstein’s *Philosophical Investigations’.* In his *Freedom and Resentment and Other Essays.* London, Routledge, 147-185.

Strawson, P.F. (2008c [1985]). *Scepticism and Naturalism. Some Varieties.* London, Routledge.

Strawson, P.F. (2011a [1952]) *Introduction to Logical Theory.* New York, Routledge.

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

Thomasson, Amie L. (2015) *Ontology Made Easy.* Oxford, Oxford University Press.

Todd, Patrick. (2016) ‘Strawson, Moral Responsibility, and the “Order of Explanation”. An Intervention.’ *Ethics* 127: 208-240.

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

Watson, Gary. (1987) ‘Responsibility and the Limits of Evil. Variations on a Strawsonian Theme’. In Ferdinand Schoeman (ed.), *Responsibility, Character and the Emotions. New Essays in Moral Psychology.* Cambridge, Cambridge University Press, 256-286.

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.

Wittgenstein, Ludwig. (1978) *Remarks on the Foundations of Mathematics.* Oxford, Blackwell.

Wittgenstein, Ludwig. (2009a) *Philosophical Investigations.* Oxford, Wiley-Blackwell.

Wittgenstein, Ludwig. (2009b) ‘Philosophy of Psychology. A Fragment’. In his *Philosophical Investigations.* Oxford, Wiley-Blackwell, 182-243.

Zimmerman, Michael J. (1988) *An Essay on Moral Responsibility.* Totowa (NJ), Rowman and Littlefield.

Zimmerman, Michael J. (2016) ‘Moral Responsibility and the Moral Community. Is Moral Responsibility Essentially Interpersonal?’ *The Journal of Ethics* 20, 247-263.

1. I focus in this paper on the influence of Wittgenstein’s ideas on meaning. I have argued in other work that Strawson’s account of responsibility is influenced by Wittgensteinian ideas in epistemology (De Mesel 2018b) and the philosophy of mind (De Mesel 2021). [↑](#footnote-ref-1)
2. I characterize Strawson’s account as ‘*broadly* Wittgensteinian’, because Strawson (2008b) distances himself from Wittgenstein at some points. The disagreements between Strawson and Wittgenstein are unimportant for my argument. My reading of Wittgenstein is inspired by Baker and Hacker (2005, 2009) and Hacker (2019). For an excellent summary of Wittgenstein’s views on meaning, rules and conventions, and the place of these views within subsequent developments in the philosophy of language, see Glock (2009). [↑](#footnote-ref-2)
3. See Bangu (2021: 454; see also 456): ‘There is no hidden property of these rules that puts them on the list. […] It is precisely the rule’s previous *public,* successful “career” as a regularity that puts it on the list’. [↑](#footnote-ref-3)
4. For a convincing account of Strawson’s methods in ‘Freedom and Resentment’, see Heyndels (2019). [↑](#footnote-ref-4)
5. Strawson does not make this biconditional explicit, but it is clearly assumed. Smith (2007) has pointed out that there are cases in which agents are responsible, while holding them responsible would be inappropriate because others do not have the standing to blame, the action was insignificant, etc. This is why, following McKenna (2012: 36-38, 47-48), I have added a *pro tanto* clause. Watson (2014: 16) calls the biconditional an ‘age-old commonplace’. Although I hope to reconcile disputants in the debate about being and holding responsible, I can only hope to reconcile those who can accept the biconditional in some form (but this goes, as far as I can see, for most participants in the debate). [↑](#footnote-ref-5)
6. If Strawson wants an overview of the rules of ‘responsible’, then why does he not look directly at the use of ‘responsible’, instead of looking at our practices of holding responsible? For Strawson, the direct approach is bound to mislead, because it takes us immediately into well-worn, abstract philosophical discussions. Wittgenstein suggests that the question ‘What is meaning?’ should be dealt with by examining how we explain what something means, and that ‘What is length?’ can partly be answered by looking at how we measure length. According to Baker and Hacker (2005: 147), these moves ‘operationalize’ the original questions. They point to what we *do* and prevent the temptation to reify meanings. In Wittgenstein’s (2009a: §107) words, operationalizing brings us ‘back to the rough ground’. In Strawson’s (2008a: 25) words, it prevents us from overintellectualizing the facts. [↑](#footnote-ref-6)
7. Beglin (2018: 613) writes that his concern-based view ‘holds that what it means to be morally responsible is determined by the basic social concerns of which our practices are an expression’. Although he refers to meaning and being responsible, his actual argument is primarily about the relation between our practices of holding responsible and basic social concerns, ‘deep-seated features of human psychology and sociality’ (Beglin 2018: 620). As he puts it, ‘our moral responsibility practices and the reactive attitudes undergirding them are an expression of certain basic social concerns, and these concerns thus normatively structure our practices’ (Beglin 2018: 617-618). Watson (2014: 21) explains how our responsibility practices express our social sentimental nature: ‘The starting point of the argument [Strawson’s argument in ‘Freedom and Resentment’] is of course that our practices bottom out in sentiments and concerns, the susceptibility to which defines our sociality’. Hieronymi’s (2020) defense of Strawson’s position does not rely on conceptual or semantic points (terms such as ‘concept’ and ‘meaning’ do not occur in the index of the book), but it relies heavily on Strawson’s social naturalism, on the idea that ‘some or another system of demands and reactions will be given with the fact of society and that a system of *reciprocal* demands will be given with the fact of *human* society’ (Hieronymi 2020: 29; see also 97 and 105). I do not mean to suggest that there are no important differences between the views of Beglin, Watson, and Hieronymi, but it is beyond the scope of this paper to evaluate which of these views comes closest to Strawson. The crucial difference between Strawson’s view and these naturalistic views is the former’s attention to the conceptual strand in Strawson’s argument. [↑](#footnote-ref-7)
8. The idea that grammar is arbitrary, in the sense that it cannot be justified by reference to an independent reality, is controversial. Some think that concepts aim to carve nature at its joints, that there is a structure in reality which our concepts aim to reflect (Sider 2011). A reviewer has rightly pointed out that, if preserving *this* thought is an important concern of those who think that holding responsible depends on being responsible, then there is at least one important concern that Strawson’s view cannot respect. From a Wittgensteinian/Strawsonian perspective, our concepts do not aim to reflect the independent structure of reality, because there *is* no such concept-independent structure. (This claim may have to be nuanced with respect to natural kinds.) There is an independent *reality*, but it is not structured. Boghossian’s (2007: 37-38) image of ‘basic worldly dough’ and deflationary positions in contemporary metaphysics, such as Thomasson (2015), fit the Wittgensteinian/Strawsonian perspective much better than Sider’s structural realism. [↑](#footnote-ref-8)
9. Shoemaker uses ‘blameworthy’ instead of ‘responsible’, ‘anger’ instead of ‘holding responsible’, and he adds ‘under standard conditions’ to the formula. These changes, though significant, are not important for my purposes. [↑](#footnote-ref-9)
10. I would like to thank an anonymous reviewer for raising this point. [↑](#footnote-ref-10)
11. I would like to thank an anonymous reviewer for pressing this question. [↑](#footnote-ref-11)
12. While having radically different rules means not having a concept of moral responsibility at all, one might say that having less radically different rules only means having another concept (maybe ‘conception’ is a better term here) *of moral responsibility*. I take Strawson to suggest that only practices expressing a concern for the quality of will with which people act and a demandfor interpersonal regard count as practices of holding morally responsible. *Our* concept of moral responsibility, determined by rules about knowledge, control, etc., is rooted in our practices of holding morally responsible, and these practices represent one way (but not the only possible way) of expressing a concern for quality of will and a demand for interpersonal regard. [↑](#footnote-ref-12)
13. On Strawson’s relation to what is now called conceptual engineering, see his discussion of Carnap’s methodology (Strawson 1963). [↑](#footnote-ref-13)