Are Our Moral Responsibility Practices Justified?

Wittgenstein, Strawson, and Justification in 'Freedom and Resentment'

Abstract

D. Justin Coates argues that, in 'Freedom and Resentment', P.F. Strawson develops a modest transcendental argument for the legitimacy of our moral responsibility practices. I disagree with Coates' claim that Strawson's argument provides a justification, in Wittgenstein's and/or Strawson's sense of that term, of our responsibility practices. I argue that my interpretation of Strawson solves some difficulties with Coates' argument, while retaining its advantages.

Keywords

P. F. Strawson; transcendental argument; moral responsibility; Wittgenstein; justification

1 Introduction

D. Justin Coates ('Strawson's Modest Transcendental Argument') argues that, in 'Freedom and Resentment', P.F. Strawson develops a modest transcendental argument for the legitimacy of our moral responsibility practices. There is much to recommend in Coates' paper, but I disagree with the claim that Strawson's argument provides a justification, in Wittgenstein's and/or Strawson's sense of that term, of our responsibility practices.

I will outline Coates' version of Strawson's argument and point to a problem with the premise that our ordinary reciprocal interpersonal relationships are justified (section two). In his discussion of this premise, Coates refers to Wittgenstein's *On Certainty*. I will argue that, in the light of *On Certainty*, our interpersonal relationships are beyond being justified or unjustified, and that this was also Strawson's view. If the premise of Strawson's argument is

modified according to this view, the conclusion of the argument is not that our responsibility practices are justified, but that they are beyond being justified or unjustified (section three). I will show how this interpretation of Strawson helps to solve some difficulties with Coates' reading of Strawson's distinction between internal and external justification (section four). Moreover, my modified version of Strawson's argument has all the advantages that Coates' argument has: it is anti-sceptical and avoids Coates' problems with naturalistic readings of Strawson (section five). I conclude that, if 'justification' is understood as Wittgenstein and Strawson understand it, Strawson does not provide a justification of our moral responsibility practices (section six).

2 The Transcendental Argument

According to Coates, Strawson offers 'a positive account of the justificatory status of our responsibility practices' (800). He offers it as a reply to the sceptic who doubts the legitimacy of these practices, and he does so by way of a modest transcendental argument. Our responsibility practices are justified because they are necessary for the possibility of ordinary interpersonal relationships (808); because these relationships are justified ('on secure ground'), so too must our responsibility practices (812).

Coates' interpretation of Strawson's argument can, for my purposes, be roughly reconstructed 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.

There is much to say about whether this argument (TA) is indeed a transcendental argument and about whether it is valid. I will assume that the argument is transcendental and valid. Because I agree with Coates' defence of P2 (see section six of his paper), I will focus on P1.

3 A Problem, a Wittgensteinian Response, Another Problem

Would Strawson subscribe to the idea that our ordinary interpersonal relationships are justified (P1)? Coates indicates that, for Strawson, a defence of this point will inevitably point us to Hume and to Wittgenstein's *On Certainty* (818). The main point of *On Certainty* is that there is an important distinction between (1) what can be justified or unjustified and (2) what is beyond being justified or unjustified (§359). What belongs to the second class cannot be sensibly questioned, affirmed or denied, 'underlies all questions and all thinking' (§415), is 'exempt from doubt' (§341), is 'not founded' (§253) but 'belongs to the foundations' (§411), to our 'frame of reference' (§83) or 'the scaffolding of our thoughts' (§211). Wittgenstein's distinction plays a crucial role in Strawson's *Scepticism and Naturalism*. In fact, all my quotations from *On Certainty* occur in Strawson's text (*Scepticism and Naturalism*, 12).

What about our interpersonal relationships? Do they belong to (1) or to (2)? Coates' reference to *On Certainty* only makes sense if our interpersonal relationships (or our commitment to these relationships, or their value) belong to what is beyond justification. At several points, Coates suggests precisely this. He says that 'the general activity of relating to others in reciprocal ways itself seems to be immune from worries about justification' (812), that our relationships with others 'provide a framework within which we are able to make sense of ourselves and of the world' (819), 'that we cannot sensibly deny the value of this framework' (819), and that 'our general engagement with others [...] does not stand in need of further justification' (820). Other passages point in the opposite direction: 'of course our friendships

and love relationships [...] are legitimate' (818), and they are 'widely agreed to be on secure ground' (812).

Coates faces a dilemma. On the first horn, he sticks with P1. The advantage here is that the argument retains its validity (assuming that the argument presented in section two is valid). The disadvantage is that Coates' reference to the influence of *On Certainty* on Strawson's approach to sceptical arguments becomes inexplicable. Because this influence is undeniably present in *Scepticism and Naturalism* (as Coates admits), a denial of it in 'Freedom and Resentment' would require a story as to how and why Strawson's approach radically changed. Moreover, if our relationships are justified, Coates will have to say something about what justifies them.

I believe that there is no way of overcoming the problems with the first horn of the dilemma. On its second horn, P1 is reformulated. The advantage here is that a reformulated version can respect the main insight of *On Certainty*, and that it is in accordance with some of Coates' claims about our interpersonal relationships ('immune from worries about justification', '[do] not stand in need of further justification'). However, this approach, though in line with some of Coates' own claims, seems to cause a problem for Coates' transcendental argument. It now becomes:

- (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 justified.

This argument does not seem valid (even if we assume the validity of the original argument in section two). I do not want to claim that the argument cannot (be made to) work, but as it stands, it is at least incomplete. How could the fact that *x* is a necessary condition of something that is

beyond being justified or unjustified make it the case that x is justified (and not beyond being justified or unjustified)? If this is the argument that Coates endorses, he owes us an explanation of how it can be made to work. If we assume that TA is valid, the most obvious way to get a valid argument with premises P1* and P2 is to adapt C. We then get:

- (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.
- This argument (TA*) is structurally similar to Coates' original argument (TA) and equally valid. Moreover, the idea that our responsibility practices are beyond being justified or unjustified is, in my view, exactly Strawson's point in 'Freedom and Resentment'.

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

This section started with a problem: P1 is incompatible with Coates' claims about the influence of *On Certainty* on 'Freedom and Resentment'. A Wittgensteinian response to the problem consists in adapting P1, but P1* causes a problem with the validity of Coates' argument that can only be solved (or, to formulate it more cautiously, that can clearly be solved) by adapting its conclusion. C*, however, is not the conclusion that Coates wishes to draw. Why not?

4 Justification in 'Freedom and Resentment'

Coates admits that, in 'Freedom and Resentment', Strawson 'seems to dismiss the import of justificatory questions' and that this seems to undercut the force of his interpretation, 'since on my view Strawson is providing a transcendental justification for our responsibility practices' (808). He explains:

There is undoubtedly a strand in 'Freedom and Resentment' that appears to dismiss the need for justification out of hand. For example, Strawson himself insists that 'the existence of the general framework of attitudes itself is something we are given with the fact of human society. As a whole, it neither calls for, nor permits, an "external" rational justification' (Strawson, 'Freedom and Resentment', 25). Strawson also goes on to remind us that attempts at "external" rational justification(s)' are precisely what get the optimist and pessimist into trouble. This suggests that Strawson's whole point is that no such justification is required. So is not any interpretation – naturalistic or transcendental – that seeks to understand Strawson as primarily concerned with the justification of our responsibility practices thereby flawed?

I doubt it. As Strawson says *immediately* preceding the potentially damning passage, 'questions of justification are internal to the structure [of "the web of attitudes and feelings which form an essential part of moral life as we know it"] or relate to modifications internal to it' (Strawson, 'Freedom and Resentment', 24-5). And this suggests that Strawson thinks there *is* a place for justification, so long as the kind of justification on offer is one that is 'internal'. Accordingly, if my proposed interpretation of the justification Strawson offers turns out to be internal to the structure of our practices in the relevant sense, then this interpretation is not dead on arrival, and can be evaluated on its philosophical merits. So the question becomes: is the proposed transcendental justification an 'internal' one? Or, like the optimist and the pessimist, have I also tried to justify our practices by appealing to considerations that are external to those practices? (808-809)

Coates' point is that Strawson does not want to do away with *all* kinds of justification: he allows for internal, but not for external justification. Coates thinks that an adapted transcendental

argument, with C* as its conclusion, does not make room for this distinction, and sees it as an advantage that his approach does.

Let us have a look, then, at how Coates explains the distinction between internal and external justification. According to Coates, an internal justification of a set of practices is one that does not abstract away from the context that gives them their significance (802, 810). In the case of our responsibility practices, the context that gives them their significance, 'the larger practice' of which they are 'an essential element' (810), is the context of ordinary interpersonal relationships, the 'wider framework of common forms of human relations' (817), and even 'our way of life more generally' (812). An external justification is one that does abstract away from this context.

There are some problems with this way of distinguishing between internal and external justification. First, if a justification of some practice p refers to a *larger* practice q of which p is an element, the term 'internal justification' seems ill-chosen. Intuitively, we would understand an *internal* justification of some practice p to be a justification that does *not* refer to elements 'outside' of p (external to p), to elements of a larger practice of which p is a part. Rather, an internal justification would be a justification that refers *only* to elements of p. The latter seems also to be Strawson's view: in the case of internal justification, questions of justification (of p) are internal to the structure (of p). Questions of justification of our moral responsibility practices are internal to the web of reactive attitudes ('inside the general structure or web of human attitudes and feelings of which I have been speaking', 'Freedom and Resentment', 25). Coates might argue that 'the general structure or web of human attitudes and feelings of which I have been speaking' does not refer to the structure of our moral responsibility practices, but to the 'larger' framework of interpersonal relationships. But how would he then read the following lines: 'The existence of the general framework of attitudes

itself is something we are given with the fact of human society. As a whole, it neither calls for, nor permits, an external "rational" justification' ('Freedom and Resentment', 25). Coates would have to understand this passage as being about the framework of interpersonal relationships. He would have to say (see his reading of the passage quoted above): this framework does not call for external justification, but it does call for internal justification, and internal justification of a practice is justification that connects that practice to a larger framework. But nothing is said about the larger framework by reference to which *this* practice (the larger practice of interpersonal relationships) could be internally justified.

Second, Strawson says that the optimist (who justifies our responsibility practices by reference to their utility) and the pessimist (who justifies them by reference to metaphysical facts) make a common mistake: they overintellectualize the facts because they look for an external justification of our responsibility practices. It is not clear, however, that the optimist's justification of our responsibility practices is external in Coates' sense. Suppose, for example, that the optimist says: our responsibility practices are justified because they are useful, in the sense that they enhance the quality of our interpersonal relationships (according to Strawson, the optimist refers to 'the efficacy of these practices in regulating behaviour *in socially desirable ways*', 'Freedom and Resentment', 4, emphasis added). This kind of justification seems both optimist and internal in Coates' sense, because it does not abstract away from the context of interpersonal relationships. If the optimist's justification is internal, however, it cannot be the case that the optimist and the pessimist make the common mistake of looking for an external justification of our responsibility practices.

In order to avoid these problems with the distinction between internal and external justification, I propose to understand the distinction as follows. An external justification of a practice is a justification of that practice *as a whole* or *in general*. Strawson's point is that

neither the framework of our moral responsibility practices, nor the framework of ordinary interpersonal relationships, can be externally justified. This just means that calls for the justification of the very practice of holding people morally responsible, or calls for the justification of the very practice of engaging in interpersonal relationships, are misguided. An internal justification occurs within a practice; it is the justification of a specific action, attitude or sentiment within that practice. With regard to the practice of engaging in ordinary interpersonal relationships, this means that questions such as 'Why do we have interpersonal relationships?' or 'Why is there such a thing as gratitude or love?' are questions for external justification, while questions such as 'Why are you grateful to him?' or 'Why do you care about her?' are questions for internal justification. With regard to the practice of holding people morally responsible, a question such as 'Why do we resent people?' is external, while 'Why do you resent her?' is internal.

There is a passage in which Coates seems to suggest this understanding of the distinction:

Consider: if having eyebrows came to be widely associated with support for a genocidal dictator, then one could certainly ask for justification from those who do not shave their eyebrows. [...] So even though growing eyebrows is in need of no general justification, in the circumstances under consideration, some specific justification must be offered on behalf of the hair above our eyes. (805)

Coates rejects this understanding, however, for the following reason. The association of having eyebrows with support for a dictator creates a specific context in which the question for a justification of having eyebrows is in place. Similarly, the truth of causal determinism would, according to the pessimist, provide a specific context in which the question for a justification of our holding another person responsible would always be in place. The pessimist could argue

that, if determinism were true, our holding another person responsible would in every case be unjustified. In other words, the pessimist could say: even if I concede that the practice of holding responsible cannot be unjustified in general, I can still maintain that holding another person responsible is unjustified in every specific circumstance. And at this point, Strawson would have to admit, according to Coates, that the distinction between external and internal justification is not effective against the pessimist.

There is much to say about this argument. I can only make two points here. First, it is true that, in specific or abnormal circumstances, a justification for a practice that normally does not require justification, such as the practice of having eyebrows (if having eyebrows can be called a practice), may be called for. The dictator case highlights such a specific circumstance, but the truth of determinism does not. Strawson's position is that *no general theoretical conviction* ('Freedom and Resentment', 12) could require us to justify a practice to which we are naturally committed (see also Magill, 'Blaming, Understanding and Justification', 190-191). The reign of a genocidal dictator is not a theoretical conviction, but the truth of determinism is. Moreover, a specific or abnormal circumstance is a circumstance that does not obtain always and everywhere. If determinism is true, however, it has always been true, it will always be true, and it is true everywhere. So it would not be a specific or abnormal circumstance. This is, basically, Strawson's point that abnormality cannot be the universal condition ('Freedom and Resentment', 12). This point has often been dismissed, although I think that it can be defended. What matters for our purposes is not whether it can be defended or not, but that it is Strawson's point.

Second, the pessimist holds that, if determinism is true, he can provide an *internal* justification for not holding others responsible in *every specific circumstance*. This 'internal justification' could be the same in each of these circumstances, something like 'It would not be

fair to hold her responsible. Because determinism is true, she could not have done otherwise.' If the justification is the same in all circumstances, the specificity of the circumstance does not count. In fact, what we have here is an external justification (holding responsible is unjustified because determinism is true) cut into ready-made pieces for every specific situation: whatever the situation is, the justification is already there. Because this is a kind of justification that is not tied to the specific situation in which it is asked (it does not appeal to the specificity of the circumstances, but to what all circumstances have in common), it is not an internal justification in Strawson's sense, but rather an external justification in disguise (see also Magill, 'Blaming, Understanding and Justification', 192, footnote 12).

In short, I believe that, *pace* Coates, my reading of Strawson on justification is effective against the pessimist (or at least as effective as Coates'). Moreover, it has several advantages over Coates' account. First, it provides a more consistent reading of the passages from 'Freedom and Resentment' (24-25) quoted by Coates in the beginning of this section. Second, it avoids some problems with Coates' reading of the internal-external distinction that I have mentioned. Third, it avoids a distinction that Coates needs, but for which he does not provide textual support: the distinction between the 'larger' practice of interpersonal relations and the 'smaller' sub-practice of moral responsibility. It is not clear that this whole-part conception is Strawson's way to think about the relationship between interpersonal relations and moral responsibility. Fourth, it allows for an analogous treatment of justificatory issues in (1) the case of moral responsibility practices and (2) the case of our ordinary interpersonal relationships. This analogy (either they are both justified, or they are both neither justified nor unjustified) is suggested both by TA and TA*, and seems needed for Strawson's argument to work. The basic analogy is this: in both cases, external justification is impossible, while internal justification is

_

¹ I would like to thank an anonymous referee for suggesting this point.

not. For Coates, it is difficult if not impossible to treat justification in (1) and (2) analogously. While he claims that our moral responsibility practices can be justified internally but not externally, he does not explain what, according to his understanding of 'internal', an internal justification of our ordinary interpersonal relationships (one that locates them in a larger framework?) would be like.

5 Advantages of TA*

I suggest that Strawson's argument is TA* rather than TA. TA* has several advantages over TA, some of which have been mentioned: it respects an insight that is central to Wittgenstein's On Certainty and prominent in Strawson's work on scepticism (section three), and it is in line with Strawson's treatment of justification (section four). Another advantage is the following. Coates recognizes that what Strawson has to say to the epistemological sceptic in Scepticism and Naturalism can be offered as a response to the moral responsibility sceptic as well (819). What Strawson has to say to the epistemological sceptic, however, is that we should not attempt to support, 'with rational justifications' (note that he does *not* say here: with rational *external* justifications), what the sceptic denies or doubts (Scepticism and Naturalism, 15). The sceptic 'puts forward his doubts by way of a challenge [...] to show [...] that the beliefs put in question are justified', and Strawson's preferred response to the sceptic is one 'which does not so much attempt to meet the challenge [that is, one which does not try to show that what the sceptic questions is justified] as to pass it by' (Scepticism and Naturalism, 2-3). What we should not do then, according to Strawson, is try to meet the moral responsibility sceptic's challenge by justifying our moral responsibility practices, while that is exactly what Coates' TA (and any argument with C instead of C* as its conclusion) does.

This point makes for a better understanding of the mistake that optimists and pessimists have in common. According to Coates, their mistake is that they ask for an *external* justification of our moral responsibility practices, one that is not connected to the 'larger' practice of which they are an essential part. Their mistake is *not*, he suggests, that they ask for a general justification of these practices. He refers to the following passage in 'Freedom and Resentment':

But it is here that the lacuna in the optimistic story can be made to show. For the pessimist may be supposed to ask: But *why* does freedom in this sense justify blame, etc.? [...] the only reason you have given for the practices of moral condemnation and punishment in cases where this freedom is present is the efficacy of these practices in regulating behaviour in socially desirable ways. But this is not a sufficient basis, it is not even the right *sort* of basis, for these practices as we understand them. ('Freedom and Resentment', 4)

Coates comments: 'In other words, Strawson claims that the optimist appeals to the wrong set of facts in her attempt to justify our responsibility practices' (801). But, as should be clear from the text, it is not *Strawson* who claims this, but the *pessimist*. And, according to Strawson, the pessimist is mistaken. Just like the optimist, she is looking for a justification of our responsibility practices. She accuses the optimist of not having provided the right sort of basis for our responsibility practices, but she does not see that she shares something with the optimist, namely the mistaken view that some sort of basis *has to be provided*. Thus, in asking for a justification of our responsibility practices, Coates makes the mistake that optimists and pessimists make. Coates' reading of Strawson is based on the idea that there is an important distinction in Strawson between an *external* rational justification of our responsibility practices and an internal rational justification of these practices, a kind of justification of our practices as a whole that we can still reasonably ask for. However, Strawson uses 'rational justification' and

'external rational justification' interchangeably, both in 'Freedom and Resentment' (14) and *Scepticism and Naturalism* (11, 15, 22). This is because Strawson's distinction is not between external and internal justification of our practices as a whole, but between external justification *as* justification of our practices as a whole, and internal justification as justification *within* the practice.

Apart from the fact that TA* has several advantages over TA, it also retains the advantages that Coates attributes to the original TA. First, because TA* is still a transcendental argument, it explains our commitment to moral responsibility practices as a rational rather than a naturalistic commitment. From the practical perspective of agents engaged in interpersonal relationships, we are rationally committed to moral responsibility practices. In Coates' words, the idea is still that 'we cannot consistently regard ourselves as being engaged in meaningful relationships with others and not also take ourselves to be apt targets of the reactive attitudes' (817). Second, TA* is still anti-sceptical. While it does not support with rational justifications what the sceptic doubts, as TA does, it does show the sceptical doubt to be idle. Third, Strawson ultimately opts for a radically modified version of optimism, the idea that determinism is no threat to our moral responsibility practices. The way in which TA modifies optimism is as follows: it claims to provide an internal instead of an external justification of our moral responsibility practices. TA*'s modification is different and more radical: it rejects the sceptic's demand for justification as idle, because our moral responsibility practices are beyond justification.

I cannot, in this response, fully support the claim that TA* is valid or sound. I submit, however, that it is closer to Strawson's argument than TA is. A methodological note is in place here. Strawson embraces the project of investigating the connections between the major structural elements of our conceptual scheme (*Scepticism and Naturalism*, 17). TA* establishes

a tight connection between (1) our moral responsibility practices and (2) the practice of engaging in interpersonal relationships, and can therefore be said to have a place in Strawson's project. I do not know, however, whether transcendental arguments are the best tools to elucidate the connection. After all, TA and TA* highlight the dependence of the relationships practice on the responsibility practice, but there is also a clear converse dependence of the responsibility practice on the relationships practice: according to Strawson, it is not just the case that we cannot have relationships without responsibility, but also that we cannot have responsibility without relationships (see De Mesel, 'Is Moral Responsibility Essentially Interpersonal? A Reply to Zimmerman'), and Coates seems to recognize this (813). Thus, (1) and (2) are *mutually dependent*. According to Strawson, the major task of analytic philosophy is 'to establish the connections between the major structural features or elements of our conceptual scheme – to exhibit it, not as a rigidly deductive system, but as a coherent whole whose parts are mutually supportive and mutually dependent, interlocking in an intelligible way' (Scepticism and Naturalism, 18, emphasis added). There is no doubt that transcendental arguments may help to accomplish this task, but it is worth keeping in mind that a transcendental argument such as TA* emphasizes only one of the two dependence relations.

6 A Note on Justification

I conclude that Strawson's argument is TA* rather than TA. If it is, then the argument does not, pace Coates, provide a justification of our responsibility practices. However, my conclusion needs to be qualified.² First, Strawson does not provide a justification *if* being justified is understood as Wittgenstein uses it in *On Certainty*, that is, as distinct from being beyond justification. If, for example, a justification for x is nothing more than an explanation that

-

² I would like to thank two anonymous referees for pressing this point.

silences doubt about x, then TA* does provide a justification of our responsibility practices: they are beyond doubt because they are necessary conditions for the possibility of our responsibility practices, and the latter are beyond doubt. Second, Strawson does not provide a justification if justification is understood as Strawson uses it in 'Freedom and Resentment', that is, in light of the distinction between external and internal justification. There is no external or internal justification, as Strawson understands them, for our moral responsibility practices. That does not mean that there is no justification in another sense of the term. If, for example, a justification for x is an explanation that makes rational (as opposed to merely natural) our commitment to x, then TA* provides a justification of our responsibility practices (see previous section). Thus, my disagreement with Coates is not so much over the claim that our moral responsibility practices are, according to Strawson's argument, justified in *some sense* of the term. Rather, I disagree with Coates over his claim that the notion of justification used in Wittgenstein's *On Certainty* and Strawson's 'Freedom and Resentment' is the notion of justification according to which our moral responsibility practices are justified if the transcendental argument is sound.

Bibliography

Coates, D. Justin. 'Strawson's Modest Transcendental Argument'. *British Journal for the History of Philosophy* 25, no. 4 (2017): 799-822.

De Mesel, Benjamin. 'Is Moral Responsibility Essentially Interpersonal? A Reply to Zimmerman'. *The Journal of Ethics* 21, no. 3 (2017): 309-333.

Magill, Kevin. 'Blaming, Understanding and Justification. A Defence of Strawson's Naturalism about Moral Responsibility'. Chap. 9 in *Free Will and Reactive Attitudes. Perspectives on P.*

F. Strawson's 'Freedom and Resentment', edited by Michael McKenna and Paul Russell, 187-

200. Abingdon: Routledge, 2016.

Strawson, P.F. 'Freedom and Resentment'. Chapter 1 in Freedom and Resentment and Other

Essays, 1-28. Abingdon: Routledge, 2008.

Strawson, P.F. Scepticism and Naturalism. Some Varieties. Abingdon: Routledge, 2008.

Wittgenstein, Ludwig. On Certainty. Oxford: Blackwell, 1975.