

# STRAWSON'S MODEST TRANSCENDENTAL ARGUMENT

## 1. Introduction

Although Peter Strawson's "Freedom and Resentment" (1974) has been widely discussed, its main argument is still notoriously difficult to pin down.<sup>1</sup> The most common interpretation of Strawson's argument takes him to be providing a thoroughly naturalistic framework for our responsibility practices. On this interpretation, Strawson argues that the legitimacy of our responsibility practices is grounded in the fact that it would be *psychologically* impossible for us to abandon the practices of praising and blaming. But as has been frequently suggested, it is implausible that psychological necessities can serve as a *justification* for our responsibility practices.<sup>2</sup> And no doubt this is correct, as far as it goes. However, as we will see, there are good reasons for thinking that this kind of naturalistic interpretation, which is so entrenched in discussions of "Freedom and Resentment," is mistaken and obscures the importance of Strawson's paper.

To rectify this mistake, I offer an alternative interpretation of Strawson's argument. As I see it, the core argument of "Freedom and Resentment" is rooted in Strawson's Kantian commitments.<sup>3</sup> In framing the issue, Strawson structures the debate between optimists and pessimists in the same way that Kant structures the antinomies in *Critique of Pure Reason*. And in motivating his rejection of these views, he apparently relies on the methodological principle (the so called "Principle of Significance") that Strawson attributes to Kant in *The Bounds of Sense*. So, in these respects, the negative task of "Freedom and Resentment" is clearly Kantian.

---

<sup>1</sup> For example, see Fischer (2015), who claims that "it is perhaps an unfortunate thing that this essay does not wear its meaning on its sleeve, so to speak; indeed I find it quite challenging," (93).

<sup>2</sup> Russell (1992) offers an especially powerful statement of this objection.

<sup>3</sup> As a biographical point, this isn't surprising. "Freedom and Resentment" was published between *Individuals* (1959) and *The Bounds of Sense* (1966). In the former, Strawson offers a transcendental argument for the existence of other minds, and in the latter, he endorses key features of Kant's critical philosophy.

More ambitiously, I'll also argue that Strawson's positive project is Kantian in spirit. Rather than offering a thoroughly naturalistic framework for moral responsibility, Strawson actually offers a *transcendental* argument for the legitimacy of our responsibility practices. Specifically, Strawson argues that these practices are grounded in an antecedent *practical* commitment to participate in ordinary human relationships—a commitment that Strawson takes to be practically (rather than psychologically) inescapable. As a result, Strawson takes our responsibility practices to be justified because they are presupposed by the very relationships that identify us as human agents. This positive account of the justificatory status of our responsibility practices, then, is also Kantian in spirit, though as we'll see, it is also more modest (and as a result, perhaps more plausible) than other Kantian accounts of human agency.

However, the aims of this essay are not just scholarly. The modest transcendental argument that I take Strawson to be developing in "Freedom and Resentment" has considerable merit. Moreover, I'll argue that it has important implications for contemporary debates about freedom and responsibility. In particular, I think that Strawson's argument, at least as I reconstruct it here, puts significant pressure on those who take moral responsibility skepticism to be *safe*—i.e., those who, like Derk Pereboom (2001, 2014), think that the truth of moral responsibility skepticism has no worrisome implications for the meaningfulness of our lives with others.

## 2. Diagnosing the Problem

"Freedom and Resentment" begins by introducing two accounts of the connection between causal determinism and the justification of our responsibility practices: optimism and pessimism. As Strawson defines these views, the optimist claims that the truth of causal determinism would not threaten the legitimacy of our responsibility practices, and the pessimist holds that the truth of causal determinism would threaten the legitimacy of our responsibility practices. Despite the fact that these views are clearly at odds with one another, Strawson interestingly takes his project to be that of reconciliation. To achieve this, he seeks, "a formal withdrawal [from the optimist] in return for a substantial concession [from the pessimist]" (Strawson

1974, 2). But what must the optimist withdraw from her account of responsibility? And what must the pessimist concede?

## 2.1. Optimism

According to Strawson, the optimist takes our responsibility practices to be justified because they're instrumental in creating a happy social order. That is, the optimist seeks to justify our responsibility practices on consequentialist grounds, by appealing to their *usefulness* in securing a well-ordered society. As Strawson puts it: "optimists about determinism point to the efficacy of the practices of punishment, and of moral condemnation and approval, in regulating behavior in socially desirable ways," (Strawson 1974, 2). But Strawson thinks that this is no basis for justification. Indeed, Strawson thinks that because they appeal to the efficacy of our responsibility practices in regulating behavior, the optimists miss something vital. As he says, "[utility] is not a sufficient basis, it is not even the right *sort* of basis, for these practices as we understand them," (Strawson 1974, 4). In other words, Strawson claims that the optimist appeals to the wrong set of facts in her attempt to justify our responsibility practices.<sup>4</sup>

So understood, Strawson accuses the optimist of *over-intellectualizing* our responsibility practices, since she seeks to justify those practices with a calculating appeal to their instrumental value in bringing about a well-regulated social order. But this is tempting only because the optimist fails to take into account *all* of the facts as we know them—most notably, she ignores the fact "our practices do not merely exploit our natures [for the good of social regulation], but express them," (Strawson 1974, 25). And in so doing, the optimist fails to appreciate the real significance of our responsibility practices.

Consequently, the optimist cheapens both the practices themselves as well as the status of our interpersonal relationships, which Strawson takes to be intimately tied to "non-detached" (and so, non-calculating) responsibility-entailing attitudes like resentment and gratitude, etc. If the justification of our

---

<sup>4</sup> Stephen Darwall (2006) actually identifies the point of Strawson's essay with this very idea. As Darwall understands Strawson, then, the "desirability [of a particular social order] is a reason of the wrong kind to warrant the attitudes and action in which holding someone responsible consists in their own terms," (Darwall 2006, 15).

responsibility practices rests solely on their utility, then we cannot understand ourselves as legitimately involved in *directly* interpersonal relationships. After all, if the optimist is correct, the legitimacy for such relationships would also depend on some utility calculus. But it's absurd to think that it would be appropriate to be grateful to or to love another merely because these attitudes satisfy some utility function. It can be appropriate for us to love or care about another simply because they are our friend, our brother or sister, our mentor, or that they are just simply lovely.

## 2.2. Pessimism

The pessimist rejects the optimist's naïve consequentialist justification for our practices. Indeed, she "recoils"—and *rightly* so, according to Strawson—at the notion that our responsibility practices are nothing more than mere instruments for social regulation. By contrast, the pessimist takes seriously the idea that backward-looking considerations have a significant and indeed, an *indispensable* role in the justification of our practices. And specifically, for the pessimist, it is only freedom of the will and basic desert that can serve to justify our practices.

However, the falsity of causal determinism is not among the "facts as we know them," and so the pessimist wrongly (at least according to Strawson) takes the facts as we know them to be insufficient as a justificatory basis for our practices. For the pessimist, then, we must go beyond the facts as we know them if we're to really see how our responsibility practices are warranted. However, because she goes beyond the facts as we know them in her attempt to justify our practices, Strawson claims that the pessimist fails to adequately recognize and appreciate her "own humanity" (Strawson 1974, 24). And this is problematic for Strawson because he takes it that our humanity is precisely the thing upon which our responsibility practices should be grounded. So as it turns out, the pessimist also *over-intellectualizes* our responsibility practices; she explains their significance by appealing to (putative) facts that abstract away from our humanity. As Strawson puts it, the pessimists' appeal to unknown (and perhaps even unknowable) metaphysical facts is nothing more than "a pitiful intellectualist trinket" that they might "wear as a charm against the recognition of [their] own humanity" (Strawson 1974, 24).

Behind Strawson's stinging rhetoric, there's an important point lurking. For whereas "the optimist's style of over-intellectualizing the facts is that of a characteristically incomplete empiricism, a one-eyed utilitarianism" (Strawson 1974, 23), the pessimist's style of over-intellectualizing the facts is that of (to borrow from a line from elsewhere in Strawson's work) a "dogmatic rationalism [that] exceeds the upper bounds of sense" (Strawson 1966, 12). And according to Strawson, pessimists make this latter mistake because they fail to recognize that the agential capacities that underwrite participation in meaningful interpersonal relationships are sufficient to underwrite our status as morally responsible agents.

This objection to the pessimist comes into even sharper focus when we consider Strawson's subsequent endorsement of Kant's so-called "Principle of Significance" in *The Bounds of Sense*. The Principle of Significance, which is the methodological principle that underlies Kant's rejection of "transcendent" metaphysics, holds that "if we wish to use a concept in a certain way, but are unable to specify the kind of experience-situation to which the concept, used in that way, would apply, then we are not really envisaging any legitimate use of that concept at all" (Strawson 1966, 16). This principle—at least as it was wielded by Kant—implies that transcendent metaphysical concepts (e.g., freedom of the will, God, the world-whole, etc.), which abstract away from the intuitive contexts that typically give concepts their significance, have no justification. Following Kant, it seems that Strawson also wields the Principle of Significance against what he calls the "panicky metaphysics" of pessimism. For Strawson, inflated metaphysical notions of freedom and desert cannot justify our practices of praising and blaming because they, like God and the world-whole, abstract away from the context that gives them significance—viz., the context of ordinary interpersonal relationships. What matters in these relationships are simply the capacities to show proper respect and concern for others, not the ability to do otherwise (holding fixed the distant past and the laws of nature) or sourcehood, as so many contemporary pessimists insist.

### 2.3. Strawson's Antinomy

It seems, then, that Strawson's diagnosis of the problems of optimism and pessimism is roughly analogous to Kant's diagnosis of the problems found in the Empiricist and Rationalist traditions. In the Third Antinomy,

Kant argues that the Empiricists, who defend the “Antithesis”—i.e., the view that “there is no freedom; everything in the world takes place solely in accordance with laws of nature” (CPR A445/B473)—fail to take seriously our humanity because their view cannot accommodate the practical perspective, which Kant takes us to be committed to *qua* agents.

In the same passage, Kant also criticizes the Rationalists, who defend the “Thesis”—the view that “causality in accordance with laws of nature is not the only causality... [and that] it is necessary to assume that there is also another causality, that of freedom” (CPR A444/B472). On Strawson’s interpretation of Kant (1966), the Rationalists also fail, in some sense, to take seriously our humanity, since in their zeal to posit transcendental freedom as an independent species of causation, they abstract away from the conditions required for possible experience—the very conditions that give our status as persons in a causal order its significance. In short, both the Empiricist and the Rationalist over-intellectualize the facts in ways that are parallel to the ways in which Strawson takes optimists and pessimists to over-intellectualize the facts.

Moreover, in his discussion of the Third Antinomy in *The Bounds of Sense*, Strawson concludes: “it seems obvious what the correct “critical” solution of this conflict should be. ... Every member of the series [of causes] which is actually “met with” in experience, however, may and must, be taken to have an antecedent cause. The thesis, then, is false, the antithesis true” (Strawson 1966, 209).<sup>5</sup> In other words, according to Strawson, although there are problems with both the thesis and the antithesis, the right way to resolve the antinomy is not by appeal to transcendental idealism (here, Strawson rejects Kant’s own solution), but simply to accept (a suitably modified) form of the antithesis (i.e., empiricism). Importantly, this will be *precisely* parallel to the move that Strawson makes with respect to the debate between pessimist and optimist, when he ultimately defends a suitably modified version of optimism. For Strawson, optimism is only acceptable when it is “radically modified” to respect the pessimists’ insight that backward-looking considerations are indispensable for the justification of our responsibility practices. So this tells us that not only is Strawson modeling his objections to optimism and pessimism on Kant’s objections to Empiricism and Rationalism, but that he’s also going to offer a positive account of how to reconcile these two views that is

---

<sup>5</sup> Henry Allison (1990) thinks this is a disastrous interpretation of Kant.

modeled on how (he thinks) Kant (should have) reconciled Empiricism and Rationalism. Recognizing this thus help us to better appreciate the positive account of the justification of our responsibility practices that Strawson introduces later on in “Freedom and Resentment.”

### 3. Freedom, Resentment, and Naturalism

Strawson’s positive project in “Freedom and Resentment” aims at bringing together the virtues of optimism and pessimism without their accommodating vices. As we’ll see, Strawson does in a way that, if not Kantian, is at least inspired by Strawson’s thinking on Kant. But before we can get to the Kantian elements of Strawson’s conciliatory project, I first want consider one popular, but to my mind mistaken, interpretation of Strawson positive project.

#### 3.1. A Justification for Moral Responsibility?

In a crucial and much discussed passage, Strawson gestures towards the justificatory basis of our moral responsibility practices. Specifically, he claims that:

I am strongly inclined to think that it is, for us as we are, *practically inconceivable* [that the truth of determinism could lead us to abandon the reactive emotions]. The human commitment to participation in ordinary inter-personal relationships is, I think, too thoroughgoing and deeply rooted for us to take seriously the thought that a general theoretical conviction might so change our world that, in it, there were no longer any such things as inter-personal relationships as we normally understand them ... (Strawson 1974, 11, emphasis added).

As I understand him on this point, the justification Strawson advances on behalf of our responsibility practices is that it would be “practically inconceivable” that we could abandon the responsibility-entailing reactive attitudes that we are prone to in virtue of our status as social agents—attitudes like gratitude, resentment, esteem, indignation, pride, guilt, love, and hurt feelings. Consequently, Strawson takes us to be

justified in holding ourselves and others morally responsible because the reactive attitudes that are constitutive of our responsibility practices are themselves grounded in our inescapable commitment to ordinary interpersonal relationships.

But *how* does the fact that it is “practically inconceivable” that we could abandon the reactive emotions (or ordinary interpersonal relationships more generally) justify our moral responsibility practices? Indeed, what does it even mean to say that it is “practically inconceivable” that we could abandon these attitudes? These questions become even more pressing when we follow the standard interpretative assumptions and adopt a relentlessly naturalistic reading of Strawson’s claim. On this interpretation, “practically” is read as modifying the relevant degree of inconceivability—it’s practically inconceivable that we could abandon our responsibility practices in just the same way it’s practically impossible that I could still run a 5 minute mile. It’s not *literally* impossible of course—just really, *really* hard given my current level of fitness, how little I’ve trained over the past few years, injuries I’ve sustained, my fondness for beer, etc. So too, on the preferred naturalistic reading of Strawson, it’s not literally impossible that we could abandon our responsibility practices; it would just be really hard given our psychological make-up.

Of course, if this is the correct way to interpret Strawson, it’s not obvious that he can answer the above challenge. How does the fact that it would be really hard for us to abandon our responsibility practices provide *any* justification whatsoever for these responsibility practices? It’s not clear that it can. To more fully assess the justificatory prospects of this interpretation, then, I turn to the gold standard for naturalistic readings of this sort: Paul Russell’s (1992) interpretation of Strawson.

### 3.2. Type-Naturalism and Justification

By claiming that our inescapable commitment to the reactive attitudes justifies our responsibility practices, Strawson might be endorsing what Paul Russell has called “type-naturalism.” For Russell, type-naturalism is the thesis that our proneness to the reactive attitudes “is natural to humans and requires no general justification of any sort,” (Russell 1992, 294). In other words, the type-naturalist claims that because we are



naturally disposed to feel attitudes like gratitude and resentment, the warrant for such attitudes cannot be questioned.

On this interpretation of Strawson's claim the thoroughgoingness of our commitment to moral responsibility puts worries about insufficient justification for our practices to rest. After all, other facts about human nature, e.g., that we have eyebrows, are not in need of justification. So why would we need to justify our natural and inescapable proneness to hold others morally responsible? In this way, type-naturalism doesn't explain the justificatory basis for our practices; instead it dispenses with the question of justification altogether, explaining why such justification is ultimately unnecessary.

As Russell notes, however, this won't undercut pesky questions about justification that pessimists are likely to raise. After all, even the pessimist can agree that natural facts about humans don't require any *general* justification. But it's a further thing to therefore conclude that our practices do not need justification in *specific* circumstances. That is, even though the type-naturalist might be right in claiming that our natural disposition to feel the reactive emotions is not in need of some general justification, it doesn't follow that any instance of a reactive emotion similarly escapes the need of justification. 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 don't shave their eyebrows. And simply pointing out most humans' *natural* disposition to grow eyebrows does no justificatory work in this context. So even though growing eyebrows is in need of no general justification, in the circumstances under consideration, it seems that some justification must be offered on behalf of the hair above our eyes.

And plausibly, says the pessimist, the truth of causal determinism similarly provides a context in which we might lack justification for engaging with others through blaming attitudes like resentment and indignation, natural though such engagement may be. As Russell puts it:

The Pessimist does not (or need not) claim that we are capable of suspending our disposition or liability to the reactive attitudes—much less that the thesis of determinism requires us to do so. ...

Rather, the Pessimist claims only that we can and must cease to entertain reactive attitudes insofar as

we have reason to believe that *everyone* is incapacitated in the relevant ways. If the thesis of determinism is true, the Pessimist argues, then we are, indeed, all morally incapacitated (Russell, 269).

In other words, the pessimist can accept type-naturalism since she can admit that, of course, we would be thoroughly and deeply *disposed* to engage with others via the reactive attitudes even if determinism were true. However, as we've seen, this does not thereby commit the pessimist to the further claim that any particular instance of resentment or indignation is *fitting*. Strawson, then, cannot appeal to the truth of type-naturalism to undermine the pessimist's position.

### 3.3. Token-Naturalism and Justification

Because type-naturalism turns out to be a non-starter, Russell claims that if Strawson is to undercut the pessimist's position, he must do so by embracing "token-naturalism." For Russell, token-naturalism is the claim that our proneness to attitudes like resentment and indignation is so deeply rooted in human nature that "we will inescapably or inevitably continue to entertain or feel [these attitudes], whatever reason suggests to us," (Russell 1992, 295). So unlike type-naturalism, which as we saw, is too weak to threaten the pessimist, token-naturalism commits Strawson to an implausibly strong claim. After all, on token-naturalism, the natural human proclivity to engage with others through the reactive attitudes is one that is *wholly insensitive* to reason.

In other words, if token-naturalism were true, then even if we were to recognize undeniably weighty reasons to refrain from regarding others from this perspective—reasons that would undercut the legitimacy of our emotional engagement with others via the reactive emotions—it wouldn't matter: our commitment to reactive emotions like resentment and indignation would be so unshakeable that we would ineluctably find ourselves resenting or getting indignant. Because of this inevitability, if token-naturalism is true, our engagement with others via blaming attitudes can't be any more rationally objectionable than other given facts of human nature, since our nature entails that we are often exercised by these emotions, even when they are not called for. And if we *can't* refrain from feeling resentment and indignation, then we can't be obligated to do so, since, after all, 'ought implies can'. Thus, on Russell's token-naturalist interpretation of Strawson, our

responsibility practices are justified in virtue of the fact that no alternative set of practices could possibly be required.

But despite the fact that token-naturalism can answer the pessimist's challenge, it is incredibly implausible as a thesis about human nature. Russell makes this point forcefully:

...it seems clear why the Pessimist finds Strawson's [token] naturalistic reply both misguided and disturbing. What is particularly disturbing about Strawson's naturalistic strategy, expressed in more general terms, is that it casts doubt on our ability or capacity to curb or control our emotional life according to the dictates of reason. More specifically, it seems clear that, despite disclaimers to the contrary, Strawson's naturalistic strategy invites us to accept or reconcile ourselves to reactive attitudes (and their associated retributive practices) even in circumstances when we have reason to repudiate them. Given this, it seems evident that we have good reason to reject Strawson's suggestion that we dismiss the Pessimist and refuse to take his arguments seriously. We have, on the contrary, every reason to take the Pessimist seriously (Russell 1992, 297-298).

In other words, because token-naturalism commits us to an implausible conception of human nature—one in which we genuinely are slaves to the destructive and rationally indefensible forces of our reactive emotions—we have good reason to doubt its truth. After all, we certainly *seem* to be capable of mollifying or withholding the reactive attitudes when the situation calls for it.

On this point, I heartily agree with Russell. Token-naturalism seems demonstrably false. We are, as a matter of fact, able to refrain from resenting when we take ourselves to have reason not to. This is especially true in circumstances in which exempting or excusing conditions are present. But it's also true in circumstances in which we think that perhaps it would be best to just "keep our mouths shut," as it were.

But notice these points pretty clearly square with the argument that Strawson develops in "Freedom and Resentment." In fact, Strawson himself is very sensitive to the role that exempting and excusing conditions play in our ordinary responsibility practices. For example, he notes that our reactive emotions *are*

sensitive to considerations that undermine an agent's capacity to fully participate in ordinary interpersonal relationships. That an agent wasn't himself, that he was pushed, that he was coerced, that he was under significant duress, etc.—these are all facts that are relevant to apt application of the reactive emotions.

Concerning this, Strawson claims that:

Inside the general structure or web of human attitudes and feelings of which I have been speaking, there is endless room for *modification, redirection, criticism, and justification*. But questions of justification are internal to the structure or relate to modifications internal to it. The existence of the general framework of attitudes itself is something we are given with the fact of human society" (Strawson 1974, 23).<sup>6</sup>

It seems, then, that Strawson would agree that the blaming attitudes are not wholly insensitive to reason. Indeed, by Strawson's lights there *is* room for rational revision of the sort Russell thinks would be impossible given the truth of token-naturalism. And so, we should not understand Strawson's attempt to justify our practices as relying on the truth of token-naturalism. As a result, his insistence that our commitment to interpersonal relationships and their attendant attitudes is practically inescapable should not be interpreted naturalistically, at least if "naturalistically" is meant to describe a kind of psychological, biological, or sociological necessity. If the inescapability of our commitment to the reactive attitudes is to serve as a justification for our responsibility, it can't be because such commitment is naturally necessary for humans.

### 3.4. Towards a Transcendental Justification

---

<sup>6</sup> Strawson also notes that sometimes we refrain from engaging with someone via the reactive emotions altogether, particularly in cases in which the "strains of involvement" are too great. This suggests that in addition to seeing these emotions as being sensitive to reasons of the sort mentioned above, they are also sensitive to pragmatic reasons of the sort that sometimes arise when we find another person too much to deal with in the moment.

Given his sensitivity to the role that excusing and exempting conditions have in our responsibility practices, it seems strange to saddle Strawson with a token-naturalistic argument. Consequently, I think we should reconsider what Strawson might have meant when he claimed that our commitment to our responsibility practices is *practically* inescapable.

Strawson's claim that it would be "practically inconceivable" that we could abandon our responsibility practices is not, at bottom, a psychological claim at all. Indeed, in my view, to say that it is "practically inconceivable" is to say that it is inconceivable from the *practical point of view of agents engaged in ordinary interpersonal relationships* that we abandon the attitudes that are essential to such relationships. In other words, Strawson seems to be claiming that the reactive attitudes are in some way a condition for the possibility of our ordinary interpersonal relationships, and as such, they cannot be questioned from the perspective of agents who are engaged in such relationships. Thus, our commitment to regarding one another via the reactive attitudes is *practically* inescapable because it is presupposed by the very activity that Strawson identifies with our humanity: the activity of relating to others as friends, lovers, partners, associates, and so on. Indeed, how else to understand the last bit of the preceding quote: that "the existence of the general framework of [reactive] attitudes itself is something we are given with the fact of human society," [Strawson 1974, 23]? Most naturally read, Strawson just seems to be committing himself to the idea that insofar as we take ourselves to be involved in the sorts of relationship that make up human society, we are practically committed to the reactive attitudes. So as I understand him, for Strawson, our responsibility practices are justified because these practices are conditions for the possibility of the ordinary interpersonal relationships that we are committed to from our practical perspective.

Of course, this interpretation is not perfectly explicit in the text, and in fact there are two significant hurdles to this reading of Strawson. The first is interpretative: elsewhere in "Freedom and Resentment," Strawson seems to dismiss the import of justificatory questions. This point in Strawson seems to significantly undercut the force of my interpretation, since on my view Strawson is providing a transcendental justification for our responsibility practices. The second hurdle is more substantive. Derk Pereboom (2001) and Tamler Sommers (2011), to name two examples, have argued very forcefully that ordinary interpersonal relationships

are not responsibility-entailing relationships. So it's not enough for Strawson to simply assert that these relationships are, in fact, responsibility entailing—it is, on its face, question-begging. In what follows, then, I will fill in a Strawsonian argument, showing that the possibility of our ordinary human relationships (as we know them to be) does depend on such relationships being responsibility-entailing relationships. Strawson does not explicitly run this argument in “Freedom and Resentment,” but it is one that is very much in the spirit of Strawson. So I think it is legitimate to reconstruct Strawson's project with this argument in mind. Once we see the way in which the reactive attitudes are presupposed by ordinary interpersonal relationships, we'll be able to see why Strawson thought that this might plausibly serve as a justification of these attitudes and their attendant practices.

#### 4. The Nature of Transcendental Justification

There is undoubtedly a strand in “Freedom and Resentment” that appears to dismiss the need for justification out of hand. 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 2008, 25). Indeed, Strawson 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 isn't 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 modification internal to it,” (Strawson 2008, 24, 25). 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.” This means that 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?

To answer these question, let’s begin by considering the nature of transcendental justification itself.<sup>7</sup> On one way of understanding such arguments, they are usually directed at those who are skeptical of some practice *p*, and they seek to show that some aspect of *p*-skepticism is problematic on grounds that (i) motivate *p*-skepticism or (ii) would be accepted by *p*-skeptics. And of particular relevance to the question at hand, it would seem that refuting *p*-skepticism on the grounds of (i) or (ii) would serve as the basis for an *internal* refutation of *p*-skepticism.<sup>8</sup> After all, if the *p*-skeptic must presuppose some element of *p* in order to motivate her skepticism, then she has no legitimate grounds for her skepticism. Thus, the successful transcendental argument will provide anti-skeptical justification for *p* on grounds that are internal to *p*.

The most ambitious form of transcendental argument will attempt to show that *p*-skepticism is incoherent. Recent attempts to revive Kantian transcendental arguments in metaethics have taken this as their goal: to show the practical reasons skeptic that her skepticism involves some kind of practical contradiction. One sort of practical reasons skeptic can’t challenge the authority of reason without presupposing it, since her challenge—*why must I act on reasons?*—derives its force only if reasons have authority. After all, to ask why you *must* act in some way is simply to ask what reasons you have.<sup>9</sup> Thus, this ambitious transcendental argument putatively provides an internal justification for the authority of practical reasons because of (i)—the basis for this kind of practical reasons skepticism seems to be problematic on the basis of the very grounds that might motivate it.

Of course, not all forms of practical reasons skepticism are so benighted. Nor do all ambitious transcendental arguments rely on showing (i). Another form of practical reasons skepticism questions the special, putatively overriding authority of moral reasons. In an impressive series of books and articles,

---

<sup>7</sup> For more on this issue, see Pereboom (2016).

<sup>8</sup> Of course, whether any transcendental argument is successful in this way is another matter. For more on this, see Stroud (1968) and Brueckner (1983).

<sup>9</sup> David Velleman (2006) offers this kind of reply on behalf of Kant.

Christine Korsgaard (1996, 2009)—again in a Kantian spirit—has attempted to offer an ambitious transcendental justification of moral reasons that, if successful, would show skepticism of this sort to be incoherent. In particular, she has argued that when we are careful in our investigations of *what it is* to act, we will discover that activity as such is governed by the Categorical Imperative. And because activity is practically inescapable for us—i.e., from the first-person practical perspective, rational creatures like us can't fail to act—all of our actions are therefore governed by the Categorical Imperative. So insofar as the skeptic takes herself to be a rational agent, she is bound by the constitutive standards of rational agency, and some of those standards are moral standards. Thus, by Korsgaard's lights, even the skeptic is committed to moral reasons.

If Korsgaard's argument is successful, it will also provide an internal transcendental justification for moral reasons. But unlike the above internal justification for practical reasons, which showed practical reasons skepticism to be problematic because of (i), Korsgaard's internal justification for moral reasons attempts to show that moral reasons skepticism is problematic because the skeptic must think of herself as a rational agent, and so should already take herself to be committed to the standards of rational agency—including moral standards. This justification relies on (ii) since there is no contradiction to be found in the grounds that motivate the moral reasons skepticism (as is the case with naïve practical reasons skepticism considered above).

Note, however, that not all attempts at internal justification—even those that rest on transcendental arguments—need to be so ambitious. There is room for more modest transcendental arguments—arguments which can serve as an internal justification for *p* even though they do not logically refute *p*-skepticism. A modest transcendental argument, if successful, would show that *p*-skepticism does not undermine the justificatory basis for *p*—not because *p*-skepticism is somehow incoherent, but because *p* itself is an essential element of a larger practice whose justificatory status is not in question. This form of transcendental argument would provide internal justification for *p* since it does not appeal to considerations (like *p*'s utility) that abstract away from the context in which *p* has its significance. And as I see it, Strawson has provided us with precisely this form of modest transcendental argument.



But before I get to the details of Strawson's argument, and in particular, how it could demonstrate our responsibility practices to be justified even if causal determinism obtains, I want to simply to note that insofar as modest transcendental arguments—like transcendental arguments more generally—seek to provide *internal* justifications of some practice, they are immune from the interpretative worry that Strawson himself rejects the need for justification. For whether or not Strawson's argument as I reconstruct it is ultimately sound, what is important to note at this stage is simply that it does not run afoul of the constraints that Strawson himself puts on a framework for moral responsibility—viz., that an attempt at justification not be 'external' to our responsibility practices themselves.

A final point before turning to Strawson's argument. In *Skepticism and Naturalism* (1983) Strawson does apparently endorse a kind of naturalism. But it's obvious that the kind of naturalism that Strawson endorses is neither type- nor token-naturalism. Moreover, it's not clearly a view that many today would label as naturalism. Instead, the sort of naturalism that Strawson's seems to endorse is one that welcomes modest transcendental argument. As Strawson puts it there:

... the naturalist philosopher will embrace the real project of investigating the connections between the major structural elements of our conceptual scheme. If connections as tight as those which transcendental arguments, construed as above, claim to offer are readily available, so much the better, (Strawson 1983, 22).

The idea of connections is key for Strawson, in a way that he makes clear in *Individuals*, again in *Skepticism and Naturalism*, and later in *Analysis and Metaphysics* (1992). The work of the philosopher—particularly the descriptive metaphysician (as Strawson takes himself to be)—is not to reductively analyze concepts or seek fundamental justifications; rather it is to demonstrate connections that exist between things, and show that at least in some cases, in light of these tight connections, if one relata cannot be questioned, then neither can the other. This, of course, is precisely what modest transcendental arguments aim for. They do not purport to show the skeptic to be refuted, but instead show why the skepticism cannot be maintained in one domain

only. And where skeptical creep is not plausible, one has reason to doubt the skeptical position in the initial domain.

Strawson's endorsement of modest transcendental arguments renders him a naturalist of sorts. But surely, he is not "relentlessly" naturalistic. Nor is he in any way suggesting that psychological facts serve as the justificatory basis for our responsibility practices. Instead, as I'll argue below, Strawson connects our responsibility practices with our way of life more generally. With some of the background clarified, I want to now turn to Strawson's argument.

## 5. Strawson's Argument

According to Strawson, the framework of reactive emotions and feelings are essential constituents of ordinary interpersonal relationships. That is, our responsibility practices are not a set of practices that are independent from social life more generally, and are instead internal to the more general set of practices that constitute our engagement with others as friends, lovers, partners, associates, etc. In particular, for Strawson it seems that our responsibility practices are conditions for the possibility of such relationships. As a result, from the practical perspective of those who find themselves engaged in interpersonal relationships, the reactive emotions and their attendant practices are inescapable. But no one has ever claimed that causal determinism constitutes a threat to our ordinary interpersonal relationships. In fact, we'll see that even moral responsibility skeptics are quick to deny that determinism itself constitutes a threat to the legitimacy of our relationships. Thus, because our interpersonal relationships are on secure ground, so too must our responsibility practices.

This constitutes a modest transcendental argument for the legitimacy of our responsibility practices. Notice that it does not logically refute the moral responsibility skeptic, since she could, after all, argue that ordinary human relationships are not legitimate modes of social interaction. But by connecting our responsibility practices with the larger human activity of engaging in interpersonal relationships, this argument does provide us with a justificatory basis for our responsibility practices that is immune from the threat of causal determinism. For even if it is possible, and indeed sometimes required, for us to doubt our justification

in being involved in some particular interpersonal relationship, the general activity of relating to others in reciprocal ways itself seems to be justified.

But it's not enough to leave it at that, since whether Strawson's modest transcendental argument is ultimately sound will depend on a number of contentious claims. In particular, it will depend on whether or not there really is an essential connection between our responsibility practices and our participation in ordinary interpersonal relationships. I turn to this claim below.

## 6. Responsibility and Relationships

At this point, I want to turn away from purely interpretative issues and instead focus on the philosophical merits of the modest transcendental argument that I am attributing to Strawson. This argument crucially depends on the premise that our ordinary interpersonal relationships are responsibility-entailing relationships. But there are significant reasons to doubt this. Recently, moral responsibility skeptics like Derk Pereboom and Tamler Sommers have argued that morally responsible agency is not a condition on the legitimacy of ordinary interpersonal relationships, since even if we are not morally responsible, it's still possible to love others, to be friends with them, and more generally, to engage with them in the ways that are characteristic of human activity.

However, it seems to me that Strawson provides us with the resources to see where these skeptics have gone wrong. In particular, Strawson's argument crucially turns on what he takes to be a "commonplace" of human life: our participation in meaningful forms of reciprocal relations. This fact about us, Strawson thinks, isn't merely important for our standing as responsible agents—it's ineliminably *vital* (indeed, this is what the optimist and the pessimist miss). After all, it is through interpersonal relationships that we, "as sharers of a common interest; as members of the same family; as colleagues; as friends; as lovers; as chances parties to an enormous range of transactions and encounters" (Strawson 1974, 6), are meaningfully engaged with one another *directly*, and reciprocally, through "non-detached" attitudes and emotions. And Strawson (plausibly) thinks that it is only because we engage in these kinds of relationships that we are morally responsible to one another. In other words, for Strawson, moral responsibility (or the property of *being*

morally responsible) is a property instantiated by human agents only because we are creatures who participate in these sorts of relationships.

This is one of Strawson's most important insights. And it's one that was largely ignored prior to Strawson and only haltingly heeded since. To see this, compare Strawson's conception of moral responsibility with those of Nietzsche and of Galen Strawson. For Nietzsche, the *causa sui*, i.e., freedom of the will in the "metaphysical superlative sense," and the kind of moral responsibility that such freedom underwrites, is grounded in "the desire to bear the whole and sole responsibility for one's actions and to absolve God, the world, ancestors, chance, and society from responsibility for them," (Nietzsche, BGE 21). No doubt we are not free and responsible in this sense. Yet, this conception of freedom and responsibility is a far cry from Strawson's idea that morally responsible agency is the glue that holds together ordinary human relationships.

Similarly, Galen Strawson (1986) identifies the property of *being morally responsible* with the property of *being an apt target of divine damnation* (or *beatification*, as the case may be). This, too, takes us away from (Peter) Strawson's more modest, but no less significant, conception of moral responsibility as being grounded in participation in distinctively human forms of engagement. For even if Galen Strawson is correct to think that it is impossible for us to be ultimately morally responsible in the sense that it would be fitting for God to eternally damn us, why should it thereby follow that we are not morally responsible in the sense that it is fitting for others to engage with us via the reactive emotions? Moreover, other than "metaphysical megalomania," why should we even care about a sense of moral responsibility that abstracts away from our ordinary interpersonal relationships?<sup>10</sup>

Of course, in light of the emphasis Strawson places on ordinary interpersonal relationships, it's natural to wonder about their import. Undoubtedly, I stand in such relationships: I am a partner, a child, a friend, a colleague, a sibling, etc. But so what? That is, what's so special about these relationships? And why should we think that they can serve as the justificatory basis of our responsibility practices?

---

<sup>10</sup> Following Daniel Dennett (1984), *ultimate moral responsibility* is not a "variety of [moral responsibility] worth wanting." I borrow "metaphysical megalomania" from Fischer (2006).

Well, for Strawson, these relationships are special because they manifest our commitment to the values of good will and respect—values located at the heart of morality. In fact, Strawson identifies this class of relationships exclusively in terms of a general concern for good will (a concern he calls “the basic demand”), where this concern is reflected in the “non-detached” attitudes and emotions that play a role in structuring reciprocal interpersonal engagement. These emotions are essential to reciprocal interpersonal relationships (i.e., relationships that require each party will regard the other with good will and respect) because our proneness to emotions like resentment and gratitude, indignation and esteem, and guilt and pride, is itself, according to Strawson, “the making of the demand [for good will and respect]” (Strawson 1974, 22). In other words, if I’m not prone to respond to your actions with gratitude and resentment or esteem and indignation, then I am not really regarding you as being subject to the basic demand that we show others good will and respect. And if I fail to regard you as subject to this demand, then I can’t *consistently* engage with you in relationships that require you to be subject to this demand (relationships such as friendships, mature loves, etc.). Therefore, if I fail to hold you to this expectation, I am not justified in treating you as a friend.

More baldly, Strawson’s argument for thinking that engagement with others through the reactive attitudes is a condition for the possibility of genuinely reciprocal interpersonal relationships is as follows. First, to engage in reciprocal relations of the sort Strawson discusses (e.g., friendship, mature love, etc.), we must, *inter alia*, expect of our friend (lover, neighbor, co-worker, etc.) that she will show us good will and respect. In other words, we must hold our friend to the basic demand of good will and respect. Undoubtedly, this is an aspirational ideal of sorts—quite often we know of ourselves and our friends that we will sometimes fail to live up to the basic demand.<sup>11</sup> But what’s important for Strawson is that, however imperfectly, we actually do hold ourselves and our friends to this demand. This suggests the first plank in Strawson’s argument:

---

<sup>11</sup> I am grateful to [omitted] for emphasizing this point to me.

1. When we interact with another agent *S* as a friend, we expect of *S* that she will show us good will and respect.

But as Strawson says, to demand or expect of another that she treats you with good will and respect *just is* the proneness to the reactive emotions. Building on this point, R. Jay Wallace argues that expectations of this sort—i.e., *normative* expectations—are constitutively tied to the reactive emotions and to the stance of *holding responsible* more generally.<sup>12</sup> And though I can't offer a full defense of this claim here, it is initially plausible. After all, if I'm, for example, playing Monopoly with a friend, I'm not really holding her to the expectation that she follow the rules of the game if I knowingly allow her to collect \$300 every time she passes GO. To genuinely *hold* her to that expectation, I must be prepared to address her unscrupulous playing. Similarly, in more straightforwardly moral domains, I'm not really holding someone to the expectation that they show their fellow persons good will and respect if I'm not disposed to address their wrongdoing. And because blaming attitudes like resentment and indignation are plausibly thought to be the most basic forms of moral address,<sup>13</sup> our proneness to these attitudes is inextricably tied to the activity of holding persons to moral norms. Hence,

2. Our expectations for good will and respect are constituted by our proneness to the reactive emotions.

And from (1) and (2), it seems to follow:

---

<sup>12</sup> See R. Jay Wallace (1994) for a full defense of this claim.

<sup>13</sup> Watson (1987) calls them “incipient” forms of moral address. Darwall (2006) takes it further, argument that the attitudes of resentment and indignation constitute forms of moral address.

3. Therefore, when we interact with another agent *S* as a friend, we are prone to engage with *S* via the reactive emotions.<sup>14</sup>

And it seems to me that an important upshot of (3) is:

4. *Justified* engagement with another agent *S* as a friend presupposes that *S* is an *apt* target of the reactive emotions.<sup>15</sup>

In other words, given (3) and the constitutive relationship between expectations of good will and respect and a proneness to the reactive emotions, it seems to follow that the *propriety* of the relationship will ultimately depend on whether *S* is a fitting or apt target of the attitudes that constitute the normative expectations that are characteristic of friendship.<sup>16</sup> After all, if friendship involves dispositions to engage with others via the reactive emotions and such individuals are not apt targets of those emotions, then it would seem that friendship itself might thereby be rendered inappropriate—something like a category mistake.

---

<sup>14</sup> Seth Shabo (2012) defends this claim on the basis of human psychology. To be friends with someone (or to love them) means you take them and their actions seriously. But when you do this, you can't help but to feel resentment (or other reactive emotions). So, for Shabo, the reactive emotions are entailed by love and friendship. Though I find much to recommend in Shabo, I do not think this reading of Strawson can make sense of the fact that he takes it to be “*practically* inconceivable” that we abandon the reactive emotions.

<sup>15</sup> I develop this argument in more detail in [omitted].

<sup>16</sup> To clarify: I don't mean for it to follow from (4) that we must always hold our friends responsible for their actions. After all, in many cases, there are good consequentialist reasons for refraining. But the presence of such reasons does not mean that the person is not an apt target of the reactive emotions. So too, sometimes the “strains of involvement” become too great, and we must temporarily alter our outlook. However, if this temporary alteration of one's outlook towards a friend becomes permanent, then it seems one no longer really relates to that person as a friend.

From here, there's not too much farther to go. For Strawson, and for many in his wake, the property of being morally responsible is understood in terms of the reactive emotions. Specifically, a Strawsonian theory of moral responsibility holds that:

5. To be an apt target of the reactive emotions just is to be a morally responsible agent.

So from (4) and (5), we can conclude:

6. Justified engagement with another agent *S* as a friend presupposes that *S* is a morally responsible agent.

This point generalizes, since it's not just friendship that seems to depend on seeing others' actions as manifesting good or ill will. Rather, this seems to be a fact of *all* reciprocal interpersonal relationships, since like friendship, all reciprocal relations seemingly require mutual expectations for good will and respect. Strawson then, shows that ordinary reciprocal interpersonal relationships essentially involve emotional engagement through the reactive attitudes.

At this point, Strawson concludes that:

7. Our ordinary reciprocal interpersonal relationships are responsibility-entailing relationships.

In other words, that we instantiate the property of being morally responsible is itself a condition for the possibility of our ordinary interpersonal relationships being fitting forms of engagement with other persons.

We can now appreciate Strawson's modest transcendental argument in defense of our responsibility practices. As I state it, (7) provides the key for unlocking what precisely Strawson means when he claims that it is practically inconceivable that we abandon our responsibility practices. Though this claim has been widely misinterpreted, it should be clear now that what Strawson means is that from *the practical perspective of agents*



*engaged in reciprocal interpersonal relationships*, we're rationally committed to the legitimacy of those responsibility practices. That is, 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. Strawson therefore leverages (7) in a way that provides an *internal* justification for our responsibility practices, since it locates the justification for praise and blame within the wider framework of common forms of human relation.

Of course, Strawson's conclusion is more modest than the sweeping anti-skeptical conclusions that some Kantians (e.g., Kant and, more recently, Velleman and Korsgaard) have offered. Those arguments try to show, e.g., that moral reasons are practically inescapable for any possible agent. But Strawson's argument, if successful, shows that moral responsibility is practically inescapable *only* for agents who are social *in a particular way*. As a result, one could doubt the justification of our responsibility practices by rejecting the legitimacy of our commitment to a way of life that involves reciprocal interpersonal relationships. Such a skeptic might accept (7) and on the basis of this reject our way of life. That is, this hardened skeptic can simply claim that if Strawson's modest transcendental argument is sound, it doesn't vindicate our responsibility practices; instead, it undermines our way of life.

But is it really credible that our friendships and mature love relationships are *per se* illegitimate? Or is it instead more plausible that such relationships *are* valuable, and so worthy of the import we naturally attach to them? Strawson never considers these questions, but the reason for this is, I suspect, simply that he takes the answer to be obvious: *of course* our friendships and love relationships (as a class of relationships) are legitimate. This suggests that Strawson is, in the end, endorsing a version of naturalism. But it is not token- or type-naturalism. Nor is it any kind of naturalism about our responsibility practices themselves. Rather, Strawson's cutting off the argument at our deep allegiance to friendships, mature love relationships, and reciprocal relationships more generally.<sup>17</sup>

---

<sup>17</sup> Watson (2015) offers a similar interpretation of Strawson on this point. However, Watson frames the issue in terms of human sociality rather than in terms of a specific variety of human sociality that prizes reciprocal relationships. It's not clear that Strawson's modest transcendental argument (at least as I've construed it here) will work for the former, though plausibly, it does for the latter. On this point I'm indebted to an anonymous referee.

Whether this move was dialectically felicitous in 1962, when Strawson published “Freedom and Resentment” is not entirely clear. To the extent the essay should be read as having anti-skeptical conclusions, we might think that he owes us (or his audience at the time) something less modest, something that could ward off *any* potential skeptical challenge. Perhaps this is so. But it’s also the case that, were Strawson’s essay to appear today, it would be significant. For as I’ve already said, contemporary skeptics regularly argue that moral responsibility skepticism is no threat to the legitimacy of our friendships or love relationships.<sup>18</sup> What this means is that they are *not* radical skeptics. And so, if Strawson’s modest transcendental argument is sound—i.e., if he succeeds in showing our responsibility practices to be inescapable from the practical point of view of agents engaged in reciprocal interpersonal relationships—then it follows that the varieties of moral responsibility skepticism currently on offer are not stable. Either, these views must admit that *some* (if not all) of our responsibility practices are defensible, or they must embrace a more radical form of skepticism.<sup>19</sup> The significance of Strawson’s modest transcendental argument, then, is not that it can parry *any* challenge to the legitimacy of our responsibility practices, but that it can parry an especially powerful and attractive one: that we can have the relationships that make our lives meaningful (e.g., friendships and love) without being committed to the nasty and brutish emotions of resentment and indignation that are at the heart of our responsibility practices.

## 7. Conclusion

I’d like to conclude by making two final points.

First, because all but the most hardened skeptics will admit that we already know our general engagement with others as friends, lovers, colleagues, and even parties to chance encounters does not stand in need of further justification, the legitimacy of these relationships cannot depend on the falsity of causal determinism, since we do not currently know causal determinism to be false. Indeed, the question of whether causal determinism is true is completely irrelevant to the question of whether you or I are fitting participants

---

<sup>18</sup> Cf. Pereboom (2001); Sommers (2011); Waller (2011); and Smuts (ms).

<sup>19</sup> I argue for this claim in more detail in [author a; author b].

in reciprocal human relationships. On this basis, we can conclude that causal determinism does not need to be false for it to be the case that our responsibility practices are, at least in general, justified by their place in a wider interpersonal framework. Strawson's compatibilism—his modified optimism—is thus vindicated.

Second, I want to note that the modest transcendental argument I attribute to Strawson secures the conciliatory aims Strawson sets for himself early in the paper. Like the optimist, Strawson concludes that causal determinism is no threat to the legitimacy of our responsibility practices. But like the pessimist insists, he does so in a way that does not instrumentalize those practices. However, unlike both optimist and pessimist, Strawson ultimately grounds the legitimacy of our practices not in some external fact, but in the practical point of view we take as creatures who relate to each other reciprocally. Strawson's positive argument thus provides us with the resources we need to solve the antimony that arises between optimist and pessimist, and it does so in a way that neatly parallels Kant's own solution to the Third Antinomy. Of course, the modesty of Strawson's argument means it is not fully Kantian: because a hardened skeptic can escape its conclusion, it fails as a "deduction." But by attending to the Kantian elements of "Freedom and Resentment," we can see the structure of Strawson's argument as it actually stands. In particular, we can see that the justification of our responsibility practices is grounded in the necessary conditions of possible human relationships.

## Works Cited

- Allison, Henry. *Kant's Theory of Freedom*, Cambridge University Press (1990).
- Brueckner, Anthony. "Transcendental Arguments I," *Nous* 17, (1983): 551-575.
- Darwall, Stephen. *The Second-Person Standpoint*, Harvard University Press (2006).
- Dennett, Daniel. *Elbow Room: The Varieties of Free Will Worth Wanting*, MIT Press (1984).
- Fischer, John Martin. "Playing the Cards That Are Dealt You," *Journal of Ethics* 10, (2006): 235-44.
- \_\_\_\_\_. "Peter Strawson and the Facts of Agency," *Oxford Studies in Agency and Responsibility, Volume 2*, eds. David Shoemaker and Neal A. Tognazzini (2015).
- Korsgaard, Christine. *The Sources of Normativity*, Cambridge University Press (1996).
- \_\_\_\_\_. *Self-Constitution: Agency, Identity, and Integrity*, Oxford University Press (2009).
- Nietzsche, Friedrich. *Beyond Good and Evil*, trans. R. J. Hollingdale, Penguin Books (1972).
- Pereboom, Derk. *Living Without Free Will*, Cambridge University Press (2001).
- \_\_\_\_\_. "Transcendental Arguments," *Oxford Handbook of Philosophical Methodology*, eds. John Hawthorne, Herman Cappelen, and Tamar Szabo Gendler, Oxford University Press (2016): 444-462.
- Russell, Paul. "Strawson's Way of Naturalizing Responsibility," *Ethics*. vol. 102, no. 2 (1992): 287-302.
- Smuts, Aaron, "Love and Free Will," unpublished manuscript.
- Sommers, Tamler. *Relative Justice: Free Will, Moral Responsibility, and Cultural Diversity*, Princeton University Press (2011).
- Strawson, Galen. *Freedom and Belief*, Oxford University Press (1986).
- Strawson, Peter. *Individuals*, Methuen & Co., Ltd., (1959).
- \_\_\_\_\_. *The Bounds of Sense*, Methuen & Co., Ltd., (1966).
- \_\_\_\_\_. "Freedom and Resentment," *Freedom and Resentment and Other Essays*, Methuen & Co., Ltd., (1974).
- \_\_\_\_\_. *Skepticism and Naturalism*, Columbia University Press (1983).
- \_\_\_\_\_. *Analysis and Metaphysics*, Oxford University Press (1992).
- Stroud, Barry. "Transcendental Arguments," *Journal of Philosophy* 65, (1968): 241-256.
- Velleman, J. David. "A Brief Introduction to Kantian Ethics," *Self to Self*, Cambridge University Press, (2006).
- Wallace, R. Jay. *Responsibility and the Moral Sentiments*. Harvard University Press (1994).
- Waller, Bruce. *Against Moral Responsibility*, MIT Press (2011).

- Watson, Gary. "Responsibility and the Limits of Evil," *Responsibility, Character, and the Emotions*, ed. Ferdinand Schoeman, Cambridge University Press (1987).
- . "Peter Strawson on Responsibility and Sociality," *Oxford Studies in Agency and Responsibility*, eds. David Shoemaker and Neal A. Tognazzini, (2015).