"Freedom and Resentment" and Consequentialism: Why: 'Strawson's Point' is not Strawson's Point

Dale E. Miller

Old Dominion University, demiller@odu.edu

Follow this and additional works at: https://digitalcommons.odu.edu/philosophy_fac_pubs

Part of the Ethics and Political Philosophy Commons

Original Publication Citation

“Freedom and Resentment” and Consequentialism: Why ‘Strawson’s Point’ Is Not Strawson’s Point

by Dale E. Miller
“Freedom and Resentment” and Consequentialism: Why ‘Strawson’s Point’ Is Not Strawson’s Point

Dale E. Miller

EARLY IN THE SECOND-PERSON STANDPOINT, Stephen Darwall articulates a thesis that he takes P. F. Strawson to advance in “Freedom and Resentment.” He maintains that Strawson asserts, and is right to assert, that good consequences are the wrong kind of reason to justify “practices of punishment and moral responsibility”:

Strawson argued that social desirability is not a reason of “the right sort” for practices of moral responsibility “as we understand them.” When we seek to hold people accountable, what matters is not whether doing so is desirable, either in a particular case or in general, but whether the person’s conduct is culpable and we have the authority to bring him to account. Desirability is a reason of the wrong kind to warrant the attitudes and actions in which holding someone responsible consists in their own terms.2

Darwall labels this thesis ‘Strawson’s Point.’

I will argue for a different interpretation of Strawson, one according to which it is not entirely true that he considers socially desirable consequences to be the wrong kind of reason to justify practices of punishment and moral responsibility and, more generally, one according to which he is not the unequivocal critic of consequentialism that Darwall takes him to be. In fact, I will contend that the account of the moral reactive attitudes that Strawson first presents in “Freedom and Resentment” may be a valuable resource for consequentialists. Because I will be challenging only Darwall’s reading of Strawson, my discussion will leave his arguments in The Second-Person Standpoint that build on Strawson’s Point intact (except insofar as Strawson’s imprimatur lends them force). I will begin by recapitulating Darwall’s objections to consequentialism and showing just how closely he takes Strawson to anticipate them.

1. Darwall’s Critique of Consequentialism and Reading of Strawson

a. Darwall’s critique of consequentialism

Darwall’s overarching aim in The Second-Person Standpoint is to establish that moral reasons are “second-personal,” by which he means that they stem from claims or demands that we make upon each other. One helpful example that he uses to illustrate the concept of a second-personal reason involves a person’s stepping on your foot. Your pain certainly gives him a reason to

move his foot, but according to Darwall this reason is “third-personal” (in virtue of being agent neutral). Everyone has reason to relieve the pain in your foot if they can; the oaf at fault is special only in how easily he can do so. When you demand that he move his foot, in contrast, this gives him a reason of a very different sort.

The reason would not be addressed to him as someone who is simply in a position to alter the regrettable state of someone’s pain or of someone’s causing another pain. If he could stop, say, two others from causing gratuitous pain by the shocking spectacle of keeping his foot firmly planted on yours, this second, claim-based (hence second-personal) reason would not recommend that he do so. It would be addressed to him, rather, as the person causing gratuitous pain to another person, something we normally assume we have the authority to demand that persons not do to one another.4

The notion of authority is central to Darwall’s conception of morality, since your making a claim or demand on someone gives her a reason only if you have the authority to make it. Authority is one of four concepts that constitute a circle that he takes to characterize the second-person standpoint. These are:

(a) the authority to make a claim on or demand or expect something of someone,
(b) an authoritative (legitimate) claim or demand,
(c) a (second-personal) reason (for complying),
(d) being accountable (to someone with the requisite authority) for complying.5

Darwall insists that “there is no way to break into this circle from outside it,” since “Propositions formulated only with normative and evaluative concepts that are not already implicitly second-personal cannot adequately ground propositions formulated with concepts within the circle.”6 He takes this stricture, which he summarizes with the slogan “second-personal authority out, second-personal authority in,”7 to undermine most consequentialist accounts of moral responsibility. The typical consequentialist strategy would be to move from the proposition that holding people accountable for a particular line of conduct would produce optimal outcomes to the proposition that they are accountable for it, but this would mean inferring that they have a second-personal reason to adhere to this line of conduct from the fact that we have a third-personal reason to treat them as if they do. It is just this sort of inference that Darwall means to block; a reason to desire the authority to make a claim is not a reason to believe that one has it.

7 Ibid.: 59.
Moreover, Darwall further argues, this same restriction also tells against most consequentialist theories of moral obligation. This is because there is a close conceptual connection between being under a moral obligation to perform or omit a particular action and being responsible or accountable for doing so. Darwall credits J. S. Mill with the best-known description of this connection (while remarking on the irony of a consequentialist’s being the author of the canonical statement of a point so damaging to consequentialism). Mill writes:

We do not call anything wrong, unless we mean to imply that a person ought to be punished in some way or other for doing it; if not by law, by the opinion of his fellow-creatures; if not by opinion, by the reproaches of his own conscience. This seems the real turning point of the distinction between morality and simple expediency.8

If a wrong action is necessarily one that we are accountable to others for not doing, then the concept of morally wrong action is implicitly second-personal, as are related concepts like that of moral obligation. This in turn means that we can infer conclusions about the moral standing of actions or our moral obligations only when some of our premises themselves invoke second-personal concepts. According to Darwall, this point is as telling against “indirect” consequentialist moral theories like rule consequentialism as it is against “direct” act consequentialism. He writes that “the only support [indirect consequentialism] allows for claims of obligation are instrumental considerations regarding how a practice of accountability itself, structured by some candidate rule, serves to advance an external goal,” and concludes that this “just seems to postpone the difficulty.”9

On the whole, then, it seems fair to describe Darwall as a critic of consequentialism in The Second-Person Standpoint, although this assessment should be qualified in two ways. First, while he takes consequentialist accounts of moral responsibility and obligation to be among the most prominent offenders, some non-consequentialist views will also attempt to break into the circle of second-personal concepts from the outside. In fact, Darwall says that Kant’s ethics do so as well, albeit from the direction of first- rather than third-personal deliberation.10 So consequentialism is not his only target. Second, Darwall is arguing for a meta-ethical restriction on how criteria for making judgments about whether someone is morally accountable for a line of conduct or what her moral obligations are can be justified, not a restriction on the criteria themselves. What this stricture prohibits are arguments for these criteria that fail to include a premise that is at least implicitly a claim about our second-personal reasons: second-personal authority out,

---

10 Ibid.: 213-42.
second-personal authority in. While Darwall concludes that contractualism is the most promising approach to normative ethics in light of his arguments, he entertains the possibility that a contractualist justification might be given for consequentialist views.11 So it is not consequentialism per se that Darwall is criticizing, but rather the arguments that most consequentialists make.

b. Darwall’s reading of Strawson

On Darwall’s reading, “Freedom and Resentment” criticizes consequentialism along nearly identical lines. At the center of Strawson’s essay is his account of the “reactive attitudes.” These are attitudes that we take toward people whom we recognize as agents in response to the use they make of their agency. For instance, we resent individuals who choose to injure or show ill will toward us (or toward people who are near and dear to us). And resentment has a “sympathetic or vicarious or impersonal or disinterested or generalized” analogue in the moral indignation or blame that we feel toward people who choose to mistreat anyone, even when we have no particular connection with their victims. “What we have here is, as it were, resentment on behalf of another, where one’s own interest and dignity are not involved; and it is this impersonal or vicarious character of the attitude, added to its others, which entitle it to the qualification ‘moral.’”12 Resentment also has “self-reactive” moral analogues, according to Strawson, including “such phenomena as feeling bound or obliged (the ‘sense of obligation’); feeling compunction; feeling guilty or remorseful or at least responsible; and the more complicated phenomenon of shame.”13

As is already apparent from the passage that I quoted in my introduction, Darwall takes himself to be following Strawson’s lead in arguing that consequentialists commonly give the wrong sorts of reasons in favor of whatever accounts of moral responsibility and obligation they endorse. The moral reactive attitudes, Darwall asserts, are inherently second-personal:

Reactive attitudes invariably concern what someone can be held to, so they invariably presuppose the authority to hold someone responsible and make demands of him. … It follows that the reactive attitudes are second-personal in our sense, and that ethical notions that are distinctively relevant to these attitudes – the culpable, moral responsibility, and I argue, moral obligation – all have an invariably second-personal aspect that ties them conceptually to second-personal reasons.14

11 Darwall (2006): 78, n. 32; 310-13. Darwall singles out John Harsanyi – who argues that the parties in the original position would favor rule consequentialism – as one consequentialist who might give the right sort of argument for the theory.
13 Ibid.: 15. While it is not entirely clear within “Freedom and Resentment” itself that Strawson means to include the self-reactive attitudes within the scope of moral attitudes, see (1980) “Reply to Ayer and Bennett,” Philosophical Subjects: Essays Presented to P. F. Strawson, Z. van Straaten, ed., Clarendon: Oxford University Press, p. 266.
What Darwall seems to mean by this is that, when you experience one of the moral reactive attitudes toward someone, you have the conviction that you had the authority to demand that she do or not do something and therefore that she had a second-personal reason. “For example,” he writes, “in feeling resentment or moral blame toward someone for stepping on your feet, you implicitly demand that he not do so … . It is to you, moreover, as if he had a reason … to avoid your feet owing to your legitimate demand, a second-personal reason.”

Darwall takes Strawson to argue in “Freedom and Resentment” that consequentialist accounts of moral responsibility are inadequate in virtue of offering the wrong kind of reason for holding people accountable for their actions. I have already noted that Darwall asserts that “Against these approaches, Strawson argued that social desirability cannot provide a justification of ‘the right sort’ for practices of moral responsibility ‘as we understand them’” and that Darwall labels this argument “Strawson’s Point.” Per Darwall, then, on Strawson’s account of the moral reactive attitudes the fact that our having some particular experience of these attitudes would produce optimal consequences does nothing to justify our having that experience, which is why it is impossible to cause oneself “to feel guilty or to resent a wrong by reflecting on the desirability (personal, social, or moral) of having these feelings.”

Similarly, Darwall claims that Strawson recognizes the conceptual connection between moral obligation and moral responsibility.

One way to see this is to note that Strawson includes a “sense of obligation” as a (reflexive) reactive attitude. … What we are morally obligated to do, he seems to be thinking, is what members of the moral community can appropriately demand that we do, including by responding with blame or other reactive attitudes if we fail to comply without adequate excuse.

On Darwall’s reading, therefore, Strawson takes consequentialist accounts of moral obligation to rest on the wrong kind of reason as well.

2. Strawson’s Two Standpoints

On Darwall’s interpretation, then, Strawson believes that consequentialist accounts of moral responsibility and moral obligation founder on the “wrong kind of reason” problem. But Darwall’s reading of “Freedom and Resentment” is somewhat selective. When taken as a whole, the essay is more con-
genial to consequentialism than Darwall acknowledges, as is Strawson’s later work on the same topics.

Strawson intends “Freedom and Resentment” as a contribution to the “free will” controversy; his ambition is to resolve the debate between “optimistic” compatibilists and “pessimistic” libertarians. What compatibilists are optimistic about is our ability to reconcile determinism with practices of moral responsibility and punishment; they believe that we can find satisfying consequentialist justifications of these practices that are fully compatible with the thesis that all human behavior has antecedent causes. Libertarians, pessimistic as they are about the compatibilist project’s prospects of success, believe that if determinism is true then we must on pain of inconsistency abandon these practices.

Strawson appreciates that this debate is of the sort in which “the conflicting doctrines, instead of being one true and the other false, share the truth between them.” On the one hand, he acknowledges that there is a serious “lacuna” in the optimist’s explanation of how our practices of moral responsibility and punishment are justified, in that it takes no account of the fact that we approve of these practices in part because they serve as an outlet for our reactive attitudes. “Only by attending to this range of attitudes can we recover from the facts as we know them a sense of what we mean, i.e. of all we mean, when, speaking the language of morals, we speak of desert, responsibility, guilt, condemnation, and justice.” On the other hand, however, the pessimist overlooks something, too, which is that attending to these practices’ role as expressions of the reactive attitudes is all that we need to do to “fill” the optimist’s lacuna. What the pessimist fails to grasp, according to Strawson, is what might be called the “resilience” of the reactive attitudes. So deeply rooted are these attitudes in human social life that they are in no way imperiled by our coming to accept some new philosophical doctrine, even determinism. This is not to deny that we might over time undergo changes in our dispositions to experience these attitudes – a possibility to which I shall return in the next section – but the pessimist’s fear that they might disappear altogether is unfounded.

So while Strawson finds something to criticize in the optimist’s stance, he does not think that the pessimist has all of the answers either. This suggests that he might not believe that the optimist’s consequentialist justification of our practices of moral responsibility and punishment is entirely misguided. One may be inclined to interject here that he must consider it completely misguided. After all, we have already seen that Darwall quotes him saying so: “… Strawson argued that social desirability cannot provide a justification of ‘the right sort’ for practices of moral responsibility ‘as we understand…’

21 Ibid.: 23, 25.
22 Ibid.: 11; see also 18.
stand them.’’ Yet it is important to look at the phrases Darwall quotes in their original context. Strawson writes:

[T]he pessimist may be supposed to ask: But why does freedom in this sense justify blame, etc.? You turn towards me first the negative, and then the positive, faces of a freedom which nobody challenges. But 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.23

Strawson is not speaking in his own voice here. Instead, he is relating what he imagines might be said by one of the contending parties in the debate that he is trying to resolve. He unquestionably means to agree with the pessimist that there is a gap in the compatibilist’s “optimistic story”; he is writing in his own voice when he calls the compatibilist’s utilitarianism “one-eyed” in virtue of ignoring the moral reactive attitudes altogether.24 Nonetheless, we should not take for granted that he wholeheartedly endorses the pessimist’s criticism. Strawson might believe that there is much more truth in the optimist’s position, including its consequentialism, than the pessimist admits. And there are clear textual grounds for thinking that he believes precisely this. In the final paragraph of “Freedom and Resentment,” he asserts that the optimist’s position is “the right one,” although it stands in need of being “radically” modified through the addition of some recognition that our practices of moral responsibility and punishment are in part expressions of our moral reactive attitudes. He then adds a conclusion that deeply problematizes any interpretation of the essay according to which it argues that consequentialist considerations are an entirely wrong kind of reason for these practices:

It is far from wrong to emphasize the efficacy of all those practices which express or manifest our moral attitudes, in regulating behaviour in ways considered desirable; or to add that when certain of our beliefs about the efficacy of some of these practices turn out to be false, then we may have good reason for dropping or modifying those practices.25

Understanding how Strawson could reach this conclusion in spite of his critique of the optimist requires recognizing that he distinguishes between two standpoints or perspectives from which we can view the world. Returning to the themes of “Freedom and Resentment” in his book Skepticism and Naturalism, he says that when we view the world from the first of these standpoints,

---

23 Ibid.: 4.
24 Ibid.: 23.
25 Ibid.: 25. Thanks to Michael McKenna for encouraging me to single out this passage for attention.
the standpoint that we naturally occupy as social beings, human behavior appears as the proper object of all those personal and moral reactions, judgments and attitudes to which, as social beings, we are naturally prone; or, to put the same point differently, human actions and human agents appear as the bearers of objective moral properties.

When we view the world from this standpoint of “participation” or “involvement,” Strawson believes, consequentialist considerations do indeed appear to be the wrong kind of reason for us to hold people morally responsible for their actions or, a fortiori, to punish them. He says little about the nature of the “objective moral properties” that he claims we will take people and (at least some of) the things they do to bear when we contemplate them from this perspective, but he does seem to be thinking primarily of properties that are second-personal in Darwall’s sense; blameworthiness and praiseworthiness are his first examples. Strawson does allow that the fact that our practices of moral responsibility and punishment play a role as expressions of our reactive attitudes does not mean that “we should be ready to acquiesce in the infliction of injury on offenders in a fashion which we saw to be quite indiscriminate or in accordance with procedures which we knew to be wholly useless” when we occupy the participant standpoint, since “savage or civilized, we have some belief in the utility of practices of condemnation and punishment.” So consequences will not appear to us to be entirely irrelevant when we see the world from this standpoint, but their relevance may seem to extend no further than to our decisions about how precisely to act on our moral reactive attitudes, e.g., what specific punishments we mete out for what offenses. From this perspective, then, the efficacy of our practices of moral responsibility and punishment does not bear emphasizing.

We can also see the world from a second standpoint, however, one that is available to us due to our ability to take what Strawson describes as the “objective attitude.” Taking the objective attitude toward an individual involves putting your reactive attitudes toward her in abeyance. “To adopt the objective attitude to another human being,” Strawson writes, “is to see him, perhaps, as an object of social policy; as a subject for what, in a wide range of sense, might be called treatment … to be managed or handled or cured or


trained…?” We naturally take the objective attitude toward people when we see that they lack something that we take to be a precondition of ordinary human agency. This may be because of some temporary disruption to their normal patterns of thought and feeling, such as great stress or post-hypnotic suggestion, or some more enduring condition, such as mental illness or immaturity.30

While Strawson thinks that we more or less automatically take the objective attitude toward people who we do not take to be exercising ordinary agency, he observes that we can also do so in other circumstances by choice. The objective attitude is not only something we naturally tend to fall into in cases like these, where participant attitudes are partially or wholly inhibited by abnormalities or by immaturity. It is also something which is available as a resource in other cases too. … [W]e can sometimes look with something like the same eye on the behaviour of the normal and the mature. We have this resource and can sometimes use it; as a refuge, say, from the strains of involvement; or as an aid to policy; or simply out of intellectual curiosity.31

It is the possibility of our taking the objective attitude through an act of will, and toward people generally rather than one person at a time – as we must when we use it, for instance, as an “aid to policy” – that allows us to view the world from what Strawson calls a “detached,” “objective” or “naturalistic” perspective (with this last term being used in one of two senses that he distinguishes, about which more below).32 Yet no one can see the world from this perspective only, taking the objective attitude toward it all of the time; the reactive attitudes are more resilient than this. We “naturally tend to fall into” the objective attitude only when we are interacting with someone who in some respect falls short of being a normal agent, and we can take it up through an act of will only for limited periods of time. Strawson describes the objective attitude as “a resource that we can sometimes temporarily make use of,” then adds, “I say ‘temporarily,’ because I do not think that it is a point of view or position which we can hold, or rest in, for very long.”33 We cannot hold this point of view for very long because we can step back from our roles as participants in social life only temporarily.

31 Ibid.: 9-10; see also Strawson (1985): 34.
32 Strawson (1985): 35-36. Darwall briefly mentions the objective attitude in The Second-Person Standpoint (2006: 69), but he does not discuss what Strawson says about our being able to adopt this attitude by choice. See also Darwall (2013): 65.
33 Strawson (1985): 34; see also Strawson (1974): 10. The pessimist’s mistake lies in thinking that our intellectual acceptance of determinism would somehow lead to our always taking the objective attitude toward everyone. This is a mistake because, given the resilience of the reactive attitudes, we are incapable of taking an objective attitude toward everyone all of the time. Embracing a new philosophical theory, even one that causes us to reconceptualize our normal social interactions, would not allow us to live the remainder of our lives taking an objective attitude toward ourselves and everyone with whom we interact.
We can, however, view the world from the detached standpoint sometimes. When we do, we will see “human beings and human actions … simply as objects and events in nature, natural objects and natural events, to be described, analyzed, and causally explained in terms in which moral evaluation has no place ….” From this standpoint, then, we will see human behavior in the optimist’s deterministic terms. We will also see the world as denuded of objective moral properties. Indeed, if someone could see the world from this perspective only, “the notion of ‘objective moral properties,’ would for him lack significance ….” Why? Viewing the world from this perspective involves, Strawson says, “the partial or complete bracketing out or suspension of reactive feelings or moral attitudes or judgments.” And for Strawson, it appears, our moral reactive attitudes are not merely responses to the objective moral properties that we usually perceive in the world. They are the source of these perceptions; we spread the attitudes on the world.

When we bracket or suspend or set on one side our moral reactive attitudes and look at the world from the detached standpoint, we may find that we have a very different conception of our reasons for action from the one that we have when we occupy the participant standpoint. No longer seeing human behavior “as the proper object of all those … moral reactions, judgments and attitudes to which, as social beings, we are naturally prone,” no longer taking “human actions and human agents” to be “the bearers of objective moral properties,” we will no longer see ourselves as having any distinctly moral reasons whatsoever. What sorts of reasons will we take ourselves to have? One obvious possibility is that, just as we share the optimist’s determinism when we view human behavior from this perspective, so too might we share his view that consequentialist considerations are our only reasons for action. We might, in other words, embrace consequentialism as a theory of practical reason. Interestingly, Strawson himself apparently takes for granted that we should do so. Consider his answer to the question of how we could deliberate about whether to rid ourselves of the moral reactive attitudes entirely, if – counterfactually – we had it in our power to do so: “[I]f we could imagine what we cannot have, viz. a choice in this matter, then we could choose rationally only in the light of an assessment of the gains and losses to human life, its enrichment or impoverishment ….” It is clear that he envisions this decision’s being made from the detached standpoint, for otherwise the most salient consideration would necessarily be that to lose the reactive attitudes would be to become indifferent to weighty reasons for or against different lines of conduct. And from this standpoint, he seems to assume that practical reasoning would naturally take the form of something

36 Ibid.: 40.
38 Ibid.: 13.
like a utilitarian calculus. It is from this standpoint that the consequences of our practices of punishment are something to emphasize.

Darwall’s reading of Strawson is therefore one sided. It uncritically identifies Strawson’s account of what reasons we take ourselves to have for holding people accountable while we view the world from our participant perspective as social beings with his account of what reasons we have tout court.39 This complaint might look like no more than a mere quibble, given that Strawson thinks that the detached standpoint is one that we can occupy only occasionally. Yet we should not understate the importance that he attaches to this standpoint. After all, it is the standpoint from which the optimist, whose position Strawson describes in “Freedom and Resentment” as “the right one,” argues. This seems to imply that it is from the detached standpoint that we see the world, including the normative order, most clearly.40 And while his position admittedly does seem to shift somewhat in the later Skepticism and Naturalism, even here he holds that the view of the world that we have from this perspective is no less accurate than that we have from the participant perspective.

39 Nor is he the only commentator on Strawson’s work to assume that Strawson himself means to endorse unreservedly the pessimist’s criticisms of consequentialism. For example, while K. E. Boxer’s recent analysis of “Freedom and Resentment” is in many respects instructive, she too describes him as “having seconded incompatibilists’ desert-related objections to accounts of moral responsibility based on efficacy … .” (2013) Rethinking Responsibility, Oxford: Oxford University Press, p. 75.

40 In a footnote to “Freedom and Resentment” Strawson may seem to go even further. Here he writes:

[M]ight it not be said that we should be nearer to being purely rational creatures in proportion as our relation to others was in fact dominated by the objective attitude? I think this might be said; only it would have to be added, once more, that if such a choice were possible, it would not necessarily be rational to choose to be more purely rational than we are (1974: 13, n. 1).

Prima facie, it is tempting to read this passage as suggesting that when we take the objective attitude we are in a better position to see what reasons we have. This in turn might suggest that what we take to be objective moral facts when we view the world from the standpoint of social beings are illusory, and so they cannot give us genuine reasons for action, even if we are emotionally unsatisfied with justifications for practices of moral condemnation and punishment that do not advert to them. But it is perhaps best not to put too much weight on this passage, given that in response to criticism Strawson later clarifies that

When I said that the surrender of the reactive attitudes would bring us nearer to being “purely rational creatures,” I did not mean that some irrational elements would disappear from our lives; for I am not committed to the view that reactive attitudes are irrational. I meant only that some affective elements, elements of feeling or emotion, would disappear, leaving us more exclusively ratiocinative creatures than before (1980: 261).

I am grateful to Michael McKenna for steering me toward Strawson’s further explication of the phrase “purely rational creatures.”
In *Skepticism and Naturalism*, Strawson raises the question “Which is the standpoint from which we see things as they really are?” He first sets out what looks to be at stake:

If it is the standpoint of participation and involvement, to which we are so strongly committed by nature and society, which is correct, then some human actions really are morally blameworthy or praiseworthy, hateful or admirable, proper objects of gratitude or resentment ... If, on the other hand, it is only from the so-called “objective” standpoint that we see things as they really are, then all our moral and quasi-moral reactions and judgments, however natural they may be and however widely shared, are no more than natural human reactions; no question of their truth or falsity arises, for there is no moral reality for them to represent or misrepresent.41

But Strawson’s way of dealing with this question is to reject it, writing that there is no error in either position but that the error lies instead “in the attempt to force the choice between them.”42 Against those who would argue that we see the world as it is only from the detached viewpoint, he embraces what he calls a “catholic,” “liberal” or “non-reductive” naturalism that he associates with one strand of Hume’s thought and with the “common sense” philosophy of Reid.43 Non-reductive naturalism stands in opposition to the “strict” or “reductive” naturalism that Strawson associates with the objective attitude and the detached standpoint. According to Strawson’s non-reductive naturalism, the resilience of the reactive attitudes, their inescapability, means that it is a mistake to think that we need to give some justification of our usually being subject to them and viewing the world through their lens. At the same time, though, he adds that

It is perfectly consistent with the adoption of the thoroughgoing or non-reductive naturalist’s way with moral skepticism – his way with the reductive naturalist – to allow validity to the purely naturalistic view of human behavior. This can be done without prejudice to the general validity of moralistic views of the same thing, so long as we are prepared to acquiesce in the appropriate relativizations of our conception of the realities of the case.44

So while he denies that the optimist’s view of human behavior and the social world is any more accurate or valid than that of the pessimist, he also denies that it is any less accurate or valid.

Strawson of course realizes that many will find implausible the notion that two ways of seeing the world that are so radically different can both be equally valid, but his commitment to the claim that they can is unambiguous.

But surely, it may be said, two contradictory views cannot both be true; it cannot be the case both that there really is such a thing as moral desert and that there is no

---

42 Ibid.: 37.
43 Ibid.: 38-41.
44 Ibid.: 41.
such thing, both that some human actions really are morally praiseworthy or blameworthy and that no actions have these properties. I want to say that the appearance of contradiction arises only if we assume the existence of some metaphysically absolute standpoint from which we can judge between the two standpoints I have been contrasting. But there is no such superior standpoint … .45

If the competing views of the world that are afforded by the detached and participant standpoints are equally valid, and it is a mistake to think that we must choose between them, then there can be no way for us to choose between the theories of practical reason that seem to us to be most warrant ed from each of these perspectives. Thus it seems we will be left with an ineliminable duality of practical reason (albeit one rather different from Sidgwick’s). From the participant perspective, it will seem clear to us that second-personal reasons, including moral reasons, figure prominently among our reasons for action. From the detached perspective, it will seem equally clear to us that no such reasons exist. The point here is not that on Strawson’s view what reasons we have depends on what perspective we are taking; the theory that seems on reflection to be most justified when we consider the matter from a given perspective will no doubt purport to offer the best account of our reasons regardless of what our circumstances are or how we happen to be looking at the world at that moment. The point is rather that from the two perspectives we will reach different answers to the question of what theory of practical reason we are most justified in accepting and that this is all that we can say about what account of our reasons for action is true or best.

So Darwall’s omission of the detached standpoint from his discussion of Strawson is not a minor one. The existence of this standpoint opens up a space in which it might be legitimate for us to employ consequentialist reasoning even if we cannot justify such reasoning on second-personal grounds. Of course, we would have to justify it on some grounds. While Strawson himself seems to take for granted that we would find something like a utilitarian account of practical reason most compelling when we have put our reactive attitudes in abeyance, our doing so hardly seems so inescapable that his catholic naturalism would license our saying that no argument for a consequentialist theory of practical reason is needed. But this argument will not need to include any second-personal premises. From the detached standpoint, we will

45 Strawson (1985): 37-38. In the interest of simplicity, I have here treated the distinction between the participant and detached standpoints as if it is strictly binary, i.e., as if at any one time we occupy one standpoint fully and the other not at all. Strawson, who does the same, points out that this is an exaggeration, although not a great one:

Standpoints and attitudes are not only different, they are profoundly opposed. One cannot be whole-heartedly committed to both at once. It will not do to say that they are mutually exclusive; since we are rarely whole-hearted creatures. But they tend in the limit to mutual exclusion” (1985: 36).
not judge any argument that depends on premises that refer to second-personal concepts sound, because we will not think that anything in the world answers to these concepts. Besides, a consequentialist account of practical reason will not issue any conclusions about who deserves blame or punishment. It will say only that we have the most reason to do those actions that will have, or can be expected to have, the best consequences (including, of course, acts of punishment and expressing blame). And, as Darwall should agree, if a theory entails no claims about second-personal authority, then an argument for it should not require any second-personal premises: no second-personal authority out, no necessity for any second-personal authority in.

In sum, then, the thesis that Darwall names ‘Strawson’s Point’ accurately reflects only how Strawson thinks things will appear to us from the participant standpoint. From this standpoint, consequences will strike us as the wrong kind of reason for our practices of moral responsibility and punishment. From the detached standpoint, though, they may seem to be precisely the right kind of reason for these practices. And, in “Freedom and Resentment,” Strawson shades toward suggesting that it is from the detached standpoint that we have the best appreciation of our reasons; it is from this standpoint that the optimist argues for determinism and for a consequentialist justification for practices of punishment and moral responsibility, after all, and Strawson says that the optimist’s position is ultimately correct and that it is appropriate both to emphasize the desirable consequences of these practices and to revise the practices when their consequences turn out to be less than desirable. Even in Skepticism and Naturalism, where Strawson is somewhat more evenhanded in his treatment of the optimist and pessimist than in “Freedom and Resentment,” he maintains that neither standpoint is epistemically privileged relative to the other.

3. An Indirect Consequentialism of the Moral Reactive Attitudes

In the previous section, I showed that Strawson is less critical of consequentialist thinking than Darwall takes him to be. When we take the detached perspective, consequentialism may seem to us to offer the best account of practical reason. Strawson himself seems to take for granted that it will, although some argument is needed to establish that this is the case. While I have not offered an argument for this claim here, nor said anything positive about what such an argument might look like, I have shown that it will not

46 What specific version of consequentialism might emerge as the most strongly justified theory of practical reason when we view the world from the detached standpoint is a difficult question, one that I cannot consider here, although I am inclined to think that it would turn out to be a “subjective” or “prospective” form of act consequentialism. See E. Mason (2014) “Objectivism, Subjectivism, and Prospectivism,” in B. Eggleston and D. E. Miller, eds., The Cambridge Companion to Utilitarianism, Cambridge: Cambridge University Press, pp. 177-98.
need to rest on second-personal premises. When we take the detached perspective, with our reactive attitudes held in abeyance, we will not think in terms of what Darwall describes as the second-personal “ethical notions that are distinctly relevant to these attitudes,” such as “the culpable, moral responsibility, and … moral obligation.” So we will neither believe that claims invoking these notions can be true nor think that the best account of practical reason is one that incorporates second-personal reasons.

But even someone who grants me all of this may think that I have implicitly conceded a considerable victory to Darwall. I have said that there is room in Strawson’s account of the reactive attitudes for consequentialism as a theory of practical reason, but I have not claimed to find room there for consequentialism as a theory of morality. In this section, as a way of meeting this objection, I will show how an indirect consequentialist moral theory might not only be consistent with but draw on Strawson’s account of the moral reactive attitudes. I will not be able to do more than sketch a bare outline of what I will call an “indirect consequentialism of the moral reactive attitudes” here, but it should be possible for me to convey at least a general sense of what such a view might look like. While most of this section is intended as a self-standing contribution to the consequentialism literature, as opposed to an answer to Darwall, near the end I will take up the question of whether this form of indirect consequentialism violates his restriction against arguments that attempt to reach second-personal conclusions without starting from second-personal premises.

In the broad sense in which I will use the term here (broader, I think, than Darwall’s), all contemporary consequentialist moral theories are indirect. In this sense, “indirect consequentialism” refers to consequentialist moral theories that tell agents to make little or no use of “calculation,” i.e., explicit consequentialist reasoning. Such theories come in two main varieties. First, there are indirect act-consequentialist theories that maintain that right action will always yield the best consequences – this is the theories’ “moral standard” or criterion for distinguishing right actions from wrong ones – yet tell agents to decide what to do by following some “decision procedure” other than expressively trying to work out which of the actions open to them this is. This decision procedure might, for instance, involve applying certain rules or expressing particular character traits. For these theories, the fact that an action was chosen via the recommended decision procedure is no guarantee of its rightness. Second, there are versions of indirect consequentialism other than act consequentialism, such as rule consequentialism and virtue consequentialism.

---

47 A point driven home to me by Susan Castro.
Theories of this second variety have decision procedures that involve little or no calculation, too, but they typically hold that actions chosen via the proper application of the recommended decision procedure are right (or at least are what ought to be done, in the case of theories like virtue consequentialism that eschew deontic categories like right and wrong action); this means that there is little distinction between their decision procedures and their moral standards. The consequentialist element of these theories comes in how the specifics of their decision procedures and moral standards are determined. A rule consequentialist, for example, may say that the “authoritative” moral code, by which I mean the collection of rules that spells out what our moral obligations are, comprises the set of rules whose general acceptance would yield better consequences than the general acceptance of any other set.

Any indirect consequentialist moral theory, of either variety, will hold that we ought to engage in consequentialist calculation sometimes. Even if such a theory denies that anyone should ever calculate “in the heat of the moment,” when we must decide between competing actions, it will still require that we occasionally engage in calculation “in a cool hour,” when no action is imminent, in order to look for ways in which our decision procedure might be improved in consequentialist terms. Indirect consequentialist moral theories of both varieties can therefore be described as “two-level” theories, inasmuch as they say that there are two different ways of deciding what to do that ought to be employed in different circumstances: application of a decision procedure on the one hand, calculation on the other.

So there is at least a superficial isomorphism between indirect consequentialist moral theorizing and Strawsonian moral psychology. Both distinguish between two different kinds of practical thinking, one that is to be used the great majority of the time and the other that is to be used only occasionally. This suggests the possibility that an indirect consequentialism might superimpose its two levels of practical thinking on Strawson’s two standpoints, saying that we should engage in higher-level thinking, i.e., calculation, when (and only when) we take the detached standpoint. When we occupy the participant standpoint, in contrast, our moral decision making will necessarily be guided by some other decision procedure.

What will this decision procedure be? It will be helpful for me to introduce some terminology here that Strawson does not and refer to our dispositions to experience the moral reactive attitudes. By this I mean no more than our propensities to experience them in particular circumstances, e.g., to feel guilty when we believe that we have performed a particular type of action. (I will concentrate in what follows on the negative or punitive moral reactive

---

52 Hare uses the phrase “cool hour” in this context in (1981): 52. An indirect utilitarian theory might call for a social division of labor, in which only some people ever engage in calculation, as in Sidgwick’s so-called “Government House” utilitarianism. For simplicity, I will ignore this possibility here.
attitudes, like “vicarious” blame or moral indignation and “self-reactive” guilt.) These dispositions would serve as the decision procedure of an indirect consequentialism of the moral reactive attitudes. Most if not all of our day-to-day decision making would be guided by them. This obviously sounds much like R. M. Hare’s “intuitive” moral thinking, which involves applying rules or “prima facie moral principles” that we have internalized in the sense of feeling compunction when we contemplate violating them and guilt when we actually do so.53 But it is unclear whether Hare supposes that these unpleasant feelings merely steer us away from certain courses of action in a more or less mechanical way or whether he would say with Strawson that they fundamentally transform the way we see the world, leading us to take actions and agents to bear objective moral properties. An indirect consequentialism of the moral reactive attitudes would affirm the latter possibility. So the dispositions that make up the decision procedure of this indirect consequentialism are more than just an algorithm or even a set of motivations, which makes them an especially robust decision procedure, one that is psychologically richer and probably also truer to our experience than some of the familiar alternatives. On this view, we really can reason about what to do from within the participant perspective, and when we do second-personal reasons will be included among the considerations that factor into our deliberations.

When we reason about what to do from the detached perspective, in contrast, we will not take ourselves to have any second-personal reasons. When we take this perspective we can certainly deliberate about when people ought to be held accountable, punished and so on, although we will approach these questions differently than we do from the participant perspective. I am presupposing that, from the detached perspective, we will answer these questions—like all practical questions—via consequentialist calculation.

We have seen that Strawson himself says that we can deliberate from the detached standpoint about whether our practices of moral responsibility and punishment stand in need of revision, and he too presupposes that we will deliberate about this in terms of the practices’ efficacy. But an indirect utilitarianism of the moral reactive attitudes will go a step further, by proposing that when we take the detached perspective we can also consider how we ought to think about these sorts of questions from the participant perspective. Suppose that our dispositions to experience the moral reactive attitudes are somewhat plastic, so that over time we can effect some changes to, e.g., the sorts of actions we feel guilty about performing. This might mean either that we can make this change on an individual basis, over the course of our lives, or that we can alter what dispositions are prevalent within our moral communities, over the course of generations. If our dispositions are to some degree plastic, then when we occupy the detached perspective we might con-
sider what dispositions to experience the moral reactive attitudes we should try to inculcate in ourselves and others. Considered from this perspective, the experiences of these attitudes that we have while we occupy the participant standpoint will be justified or warranted only insofar as they result from the dispositions that we have the most reason to instill. And if a consequentialist theory of practical reason would seem to us to offer the most compelling account of our reasons for action when we make this assessment from the detached standpoint, then these would be the dispositions that it would best promote the good for us to instill. What I am proposing, obviously, is that from the detached standpoint we might take a frankly instrumentalist view of our dispositions to experience the reactive attitudes, regarding them – like our practices of punishment and moral responsibility – as means of regulating behavior. This is to say that we might treat the question of what dispositions we should have as a matter of policy.

Strawson compares the justification of the reactive attitudes and the justification of induction:

The human commitment to inductive belief-formation is original, natural, non-rational (not irrational), in no way something we choose or could give up. Yet rational criticism and reflection can refine standards and their application … 54

Analogously, the indirect utilitarianism of the moral reactive attitudes that I am proposing calls for using rational criticism and reflection to refine not just the practices of moral responsibility and punishment that express our moral attitudes but also the very attitudes themselves, or rather our dispositions to have the attitudes. I do not mean to suggest that merely deciding from the detached perspective that it would be desirable to have a certain set of dispositions to experience the reactive attitudes would be sufficient to give one those dispositions; Darwall is surely right that it is impossible to make oneself “feel guilty or to resent a wrong by reflecting on the desirability … of having these feelings.” But you might reflect instead on the desirability of becoming a person who would feel guilty about engaging in some line of conduct yourself and resent others for doing the same. And while merely reflecting on this will almost certainly not be enough by itself to actually make you that sort of person, it might be enough to make you resolve to adopt the project of becoming such a person. Over time you might be able to pull this off – watching enough videos produced by PETA might cause many unconflicted omnivores to begin to feel some pangs of guilt while tucking into a steak, for instance. And given the way that Strawson takes our experiences of the reactive attitudes to color our perceptions, to cause us to see agents and actions as bearers of objective moral properties, this would mean changing one’s judgments from the participant standpoint about what objective moral properties various acts and agents bear and about what second-personal moral reasons you have.

Empirical questions abound here, such as that of how much power we might have to revise these dispositions. Strawson does not tackle questions like these himself, although he stresses that he does not mean to deny “the possibility and desirability of redirection and modification of our human attitudes” in light of empirical research, particularly research in psychology.\footnote{Ibid.: 24-25.} He also notes our “increased historical and anthropological awareness of the great variety of forms which these human attitudes may take at different times and in different cultures.”\footnote{Ibid.: 24. See also (1985): 46-48.} Certainly nothing he says suggests that he considers it impossible for a person to make specific retail changes in her dispositions to experience these attitudes; it is only the wholesale jettisoning of the reactive attitudes altogether that he claims is beyond our power.

What I have sketched so far might be called an indirect utilitarian theory of moral responsibility, but it is not yet a theory of right. But now suppose that Darwall is right (as I think he is) to endorse the Millian conception of moral obligation that he describes, according to which being under a moral obligation is a matter of being appropriately subject to some of the negative moral reactive attitudes (and possibly to punishment) if one fails to perform or omit some line of conduct. A Strawsonian who takes consequentialism to offer the best account of practical reason from the detached perspective will hold that, considered from that perspective, our dispositions to experience the moral reactive attitudes are appropriate just if it would be optimific for us to instill in ourselves dispositions to blame others for actions of that sort, to feel guilty for doing them and so on. So an indirect consequentialism of the moral reactive attitudes will say that an action is wrong, and hence that we have an obligation to omit it, just if a person with the optimific dispositions to experience the moral reactive attitudes would tend to feel guilty about doing the action herself, to blame others for doing it, etc.

Given this, such a theory will almost certainly be an indirect consequentialism of the second of the varieties that I distinguished previously. That is, it will not be a version of act consequentialism. There is very little reason for us to believe that it would maximize value for us to be disposed to blame others for every failure to maximize value, to feel guilty about every such failure of our own, etc. The theory \textit{might} be a version of rule consequentialism; that would depend on whether our dispositions to experience the moral reactive attitudes, or at least the best such dispositions that it is in our power to instill in ourselves, can be adequately captured in terms of a set of rules that we would blame others for violating, feel guilty about violating ourselves and so on.

One might wonder at this juncture what reason a Strawsonian consequentialist would take us to have for considering from the detached perspective what dispositions to experience the moral reactive attitudes we should have.\footnote{Noell Birondo pointed out the necessity for me to address this question.} To be sure, the liberal or catholic naturalism that Strawson espouses...
in *Skepticism and Naturalism* would not permit an individual to say that the specific dispositions to experience the moral reactive attitudes that she has at a given time stand in need of no further justification simply in virtue of the fact that she has them. Nor would it let the members of a given society say this about dispositions that are widely shared among them. In (very) loose terms, the sort of Reidian common sense philosophy to which Strawson appeals says that we are entitled to trust first principles of the human mind. These are principles that are inherent in our nature; they are not the products of enculturation and they cannot be altered by any experiences that we might have. So the only dispositions to experience the moral reactive attitudes that the view could encourage us to regard as needing no further justification are those that will invariably be shared by everyone, i.e., those that are not at all plastic. But the indirect consequentialism of the reactive attitudes that I have sketched only calls on us to reflect from the detached standpoint about whether those dispositions of ours that we believe are plastic are less well justified than others that might be inculcated in their place. Hence Reidian liberal naturalism does not directly entail, as it were, that the sort of “detached assessment” of our dispositions to experience the moral reactive attitudes for which this moral theory calls is otiose.

But this does not by itself gainsay the possibility that it might be possible to find sufficient justification for these dispositions from within the participant standpoint, so that reflecting on them from the detached standpoint would not be necessary. For instance, some interpretations of “reflective equilibrium” might imply that our dispositions are justified as long as the moral judgments that we would make on the basis of them cohere closely enough with each other and with our other beliefs. Darwall’s sophisticated positive argument in *The Second-Person Standpoint*, while not cast in terms of reflective equilibrium, can also be regarded as an example of thinking from within the participant standpoint about what dispositions to experience the moral reactive attitudes we are justified in having. (While Darwall’s intricate argument resists easy summary, in the proverbial nutshell he argues that by making any second-personal demands on one another at all we commit ourselves to viewing one another as free and equal, which then has implications for what demands we have the authority to make.) Strawson’s liberal naturalism might, in virtue of suggesting that there is nothing rationally defective about the participant standpoint, offer indirect support to the notion that we should assess dispositions to experience the moral reactive attitudes from the detached standpoint only after concluding that there is no satisfactory way to assess them that remains internal to the participant standpoint. So while my argument suffices to show that there is nothing irrational from a Strawsonian point of view about the detached assessment of dispositions to experience

the moral reactive attitudes, far more work might be needed to show that doing so is rationally obligatory.

While this rough sketch of an indirect consequentialism of the moral reactive attitudes stands in need of extensive further development, I have said enough, I think, for us to see that Darwall’s stricture against arguments for second-personal conclusions that have no second-personal premises poses no problem for the view. The theory requires two general sorts of arguments, and neither runs afoul of this stricture. First, the view cannot get off the ground unless some argument can be given that will convince us that a consequentialist theory of practical reason offers the best account of our reasons for action, when we consider the question from the detached standpoint. That argument will have no second-personal premises, for reasons that I have already noted, but its conclusion is not second-personal, either. From that conclusion plus empirical propositions we would draw further conclusions about which dispositions to experience the reactive attitudes we have the most reason to instill in ourselves, but these further conclusions are also not second-personal.

When we occupy the participant standpoint, in contrast, we will make and accept arguments with second-personal conclusions, e.g., about who has the authority to make what claims or demands and about when moral condemnation or punishment is deserved. But from this standpoint, we will have second-personal beliefs about the objective moral properties of acts and agents upon which these arguments can be premised. Of course, what specific second-personal propositions we believe will depend upon what dispositions to experience the moral reactive attitudes we have, and so what conclusions we reach when we occupy the detached standpoint may make a difference to what arguments we accept when we occupy this one. Nevertheless, the arguments that this view has us making and accepting from the detached standpoint are not arguments for second-personal propositions. Rather they are, at least in part, arguments for why we have reason to try to make ourselves into people who accept certain second-personal propositions. Without attempting to break into the circle of second-personal concepts from outside, these arguments may still bear on how we should think from within that circle.

4. Conclusion

Darwall’s reading of “Freedom and Resentment” makes Strawson out to be more of an opponent of consequentialism than he in fact is. Indeed, it is not clear that he is an opponent at all. Granted, Strawson does think that consequentialism is inconsistent with our ordinary moral experience, since we ordinarily perceive acts and agents as having objective moral properties and take those properties to ground second-personal moral reasons. Nonetheless, it is a mistake to conflate the view of the world that he takes us to have from the participant standpoint – the standpoint that we occupy when we have our
ordinary moral experience – with his view of the world. He credits us with being able, at intervals, to put our reactive attitudes in abeyance in a fairly thoroughgoing way. When we do this, and so look at the world from the detached rather than the participant standpoint, we will neither perceive any objective moral properties nor countenance any moral reasons. This means that, when we occupy this detached standpoint, it is at least possible that some version of consequentialism will seem to us to offer the most compelling theory of practical reason, as Strawson himself seemingly takes for granted that it will. And, importantly, he judges the views of our behavior and our reasons that we have from this perspective to be, at worst, no less warranted than those that constitute our ordinary moral experience. Contra Darwall, then, ‘Strawson’s Point’ is a misnomer as a label for the claim that desirability is the wrong sort of reason for practices of moral responsibility and punishment; this claim does not reflect the totality of Strawson’s view. Strawson’s account of the moral reactive attitudes actually suggests the possibility of an indirect consequentialism of the moral reactive attitudes, which may turn out to be a useful way of framing an indirect consequentialist moral theory. Such a view would, in contrast to the simple one-eyed utilitarianism that Strawson criticizes, have both eyes open (albeit perhaps not at the same time); it would be fully cognizant of the fact that our practices of moral responsibility and punishment are in part expressions of our moral reactive attitudes. Part of the interest of this view derives from the fact that it may be possible to defend such a moral theory in a way that circumvents Darwall’s requirement that arguments for theories of moral responsibility and obligation incorporate second-personal premises.59

Dale E. Miller
Old Dominion University
Department of Philosophy and Religious Studies
demiller@odu.edu

59 In addition to those whom I have already named, Ben Eggleston, Christopher Freiman, audiences at Wichita State University and the Universität Hamburg and an anonymous reviewer for this journal all provided valuable feedback and have my gratitude. Thanks are also due to Old Dominion University for the research leave during which this paper was written.