Responsibility and Normative Moral Theories

Stephen Darwall and R. Jay Wallace have independently argued that morality is essentially interpersonal by appealing to necessary connections between morality and responsibility. According to Darwall, morality is grounded in fundamentally second-personal accountability relations. On Wallace’s view, a normative moral theory must say that agents’ attitudes toward the moral properties of their actions are reasons for responsibility reactions, which only relational moral theories can do. If either argument succeeds, non-relational moral theories are flawed. I demonstrate that neither argument succeeds. First, I show that grounding morality in accountability relations is implausible. I then argue that the necessary connections that Wallace posits between responsibility and morality exist but need not be explained by moral theories. Finally, drawing on the objections to Darwall and Wallace, I show that, plausibly, no necessary connection between morality and responsibility exists that would rule out any non-relational moral theory. Hence Darwall and Wallace’s strategy does not work.

Keywords: responsibility, morality, consequentialism, second-personal reasons, reactive attitudes, fittingness

Stephen Darwall (2006) and R. Jay Wallace (2013) have independently argued that morality is fundamentally interpersonal by appealing to necessary connections between morality and responsibility. According to Darwall, morality is grounded in fundamentally second-personal accountability relations. On Wallace’s view, the correct normative moral theory must say that agents’ attitudes toward the moral properties of their actions are reasons to respond to the agents with ‘responsibility reactions,’ or reactions that hold them responsible. This requirement on a moral theory, Wallace claims, shows that morality is essentially interpersonal. If either argument succeeds, normative moral theories that are at bottom non-relational, such as theories that ground morality in impersonal value or in human flourishing, are flawed.

I aim to show that these arguments do not succeed. I then make the case that, plausibly, no other necessary connection between morality and responsibility exists that
would show that morality is fundamentally relational or that would rule out any non-relational moral theory. The upshot is that Darwall and Wallace’s basic strategy does not work.

The argument proceeds as follows. In Section 1, after describing Darwall’s view, I raise two independent objections to his idea that morality is grounded in fundamentally second-personal accountability relations. First, I argue that, contrary to this idea, some agents who lack the capacities required to participate in accountability relations can act wrongly. Second, I argue that accountability is not plausibly fundamentally second-personal, and so neither is morality. In Section 2, I make the case that, contrary to Wallace, the necessary connections between morality and responsibility that he posits do not undermine any moral theory. This is because a moral theory need not explain why agents’ attitudes toward moral properties are reasons for responsibility reactions, and it need not explain why we all have reason to care about morality. Finally, in Section 3, I argue that, plausibly, no other necessary connection between morality and responsibility exists that would undermine non-relational moral theories.

1. **Morality is not Grounded in Accountability Relations**

In this section, after presenting Stephen Darwall’s idea that morality is grounded in accountability relations, I raise two independent objections against it.
1.1 Darwall’s View

Darwall’s view in a slogan is morality as equal accountability. What we are morally obligated to do or not do, according to Darwall, just is what we have the authority to demand of one another that we do or not do. Since our authority to make demands of one another implies our accountability to one another for complying with those demands, we can put Darwall’s idea like this: what we are morally obligated to do or not do just is what we are morally accountable to one another for doing or not doing.

If Darwall is right, then the reasons to act provided by moral obligations are second-personal reasons. A second-personal reason is a normative reason for acting that derives from our authority to make claims and demands on one another’s conduct. Borrowing Darwall’s example, my authority to demand that you not step on my foot gives you a second-personal reason not to step on my foot. This second-personal reason contrasts with the agent-neutral reason that you have not to step on my foot that is grounded in the badness of pain rather than in my authority to demand this of you. Because our authority to make claims and demands on one another’s conduct implies our responsibility to one another for complying, we can equally say that our being accountable to one another gives us second-personal reasons to act in certain ways. For example, you have a second-personal reason not to step on my foot because you are accountable to me for not stepping on my foot.

Crucially, Darwall thinks that our authority to make claims and demands upon one another’s conduct is irreducibly second-personal rather than being grounded in our relationship to third-personal facts. Darwall therefore contrasts our authority to make claims and demands upon one another’s conduct with epistemic authority, which is
grounded in our relationship to third-personal epistemic facts. If Darwall were right, accountability and moral obligation would be irreducibly second-personal concepts, and non-relational moral theories would be wrong. Yet, I now argue, Darwall’s idea is flawed for two independent reasons.

1.2 The First Objection: Wrongness and Exempting Conditions

On Darwall’s view, an agent acts wrongly in virtue of acting in a way that he is accountable to others for not acting. Thus, for Darwall, an agent acts wrongly in virtue of being blameworthy for his action absent an excuse (Darwall 2010: 143; Darwall 2006: 93). Notice that Darwall’s idea reverses the common idea that an accountable agent is blameworthy for an action in virtue of acting wrongly without an excuse. As I now argue, Darwall’s idea, in contrast to the common one, mistakenly says that any agent who lacks the capacities required for accountability cannot act wrongly.

First distinguish excuses and exemptions. As P.F. Strawson (1962/2003) says, reactive attitudes – e.g. resentment, indignation, and gratitude – are reactions to the quality of will that agents express in their actions, and they hold agents responsible for those actions. Negative reactive attitudes, like resentment and indignation, are inhibited by excuses and exemptions.¹ Excuses – e.g. ‘he didn’t mean to’ or ‘he couldn’t have known’ – inhibit negative reactive attitudes consistently with viewing the agent as fully responsible. They do so, Strawson says, by showing that the agent in fact had sufficient good will toward others when he acted. Unlike excuses, exemptions – e.g. ‘he is a young

¹ Strawson (1962/2003) makes this distinction without these category labels, which I borrow from Wallace (1994).
child’ or ‘he has schizophrenia’ – invite us to see the agent as not a responsible agent (at least not fully so). They instead encourage us to take the ‘objective attitude’ toward him: to see him ‘as an object of social policy… to be managed or handled or cured or trained…’ (Strawson 1962/2003: 79). Therefore, a person with an exemption is not an appropriate object of reactive attitudes (at least not fully so).

Since being an accountable agent is consistent with having an excuse, Darwall builds excuses into the conceptual connection that he posits between wrongness and blameworthiness: an action is wrong if and only if an agent is blameworthy for doing it absent an excuse. However, this conceptual connection entails that agents who meet exempting conditions cannot act wrongly, since they are not blameworthy for their actions even when they lack an excuse. As I will show, that is the wrong result.

Yet first, I want to head off the concern that I am misinterpreting Darwall. You might worry that, by ‘excuses,’ Darwall means both excuses and exemptions. That is not so. Darwall understands morality in terms of accountability to one another, and agents who meet exempting conditions cannot participate in accountability relations. More specifically, Darwall claims a) that an agent must possess second-personal competence – i.e. the capacity to recognize and respond to second-personal reasons – in order to be subject to second-personal reasons and so moral obligations and b) that this capacity is required for accountable agency. Since agents who meet exempting conditions lack the competence required for accountable agency, Darwall is committed to saying that they lack second-personal competence and so are not subject to moral obligations. They therefore cannot act wrongly.

Darwall is mistaken on this point. Although some agents who meet exempting

---

2 See, for example, Darwall 2006: 21–3, 33–5, 75–9.
conditions – e.g. lions and infants – cannot act wrongly, other agents who meet exempting conditions can. Consider young children. When a young child shoves her brother to get a toy, her parent tells her that she acted wrongly, and that seems true. It is not that the young child would have acted wrongly had she been older and more mature and had thereby developed the competence required to be responsible. She acted wrongly in spite of lacking that competence. You might object that the parent is just training the child. However, we also say that people with certain mental illnesses act wrongly, even though we are not trying to train them. For example, a schizophrenic who shouts racial slurs at black people acts wrongly, even though his mental illness is an exempting condition. Further, psychopaths can behave in cruel and selfish ways, and they act wrongly when they do. Yet they plausibly meet an exempting condition because they lack the capacity for proper moral reasoning (Watson 2011).

Is this just a clash of intuitions? After all, Darwall might reject the intuition that these agents act wrongly. Fortunately, I can do more than appeal to intuition; I can explain why the above agents can act wrongly despite meeting exempting conditions.

To start, notice a major difference between lions and infants on the one hand and children, the mentally ill, and psychopaths on the other: only agents in the latter group express a quality of will in acting, making it appropriate to appraise them morally for their actions. The young child who shoves her brother to get a toy acts selfishly. When she instead shares with him, she expresses kindness. The schizophrenic who shouts racial slurs at black people expresses mean-spiritedness, and psychopaths express cruelty in manipulating others. The fact that these agents’ actions express a quality of will makes it plausible that they can act wrongly. After all, it would be odd to suppose that we can
appropriately appraise someone in terms of virtues and vices for an action that we cannot assess as morally right or wrong. For example, it would be odd to suppose that we can legitimately appraise a psychopath as cruel for some harmful action but cannot legitimately claim that he acted wrongly in doing it. If an agent is capable of acting cruelly, we should accept that he is capable of acting wrongly. In saying this, I mean to provide further intuitive support for the intuition that young children, the mentally ill, and psychopaths can act wrongly despite not participating in accountability relations.

Why can agents act wrongly when their actions express their quality of will? On my view, which I connect to the quality of will idea below, the explanation is this: to be capable of acting wrongly, it is sufficient to have the capacity to make and act on judgments about normative reasons. By this capacity, I mean the capacity to take certain considerations as counting for or against actions and to respond accordingly in action. To take a consideration as counting for or against an action, the consideration simply needs to play the role of counting for or against an action in the agent’s reasoning (including subconscious reasoning) about what to do. The agent need not conceptualize the consideration as a reason, and he may be wrong that it is a reason. For example, if I volunteer for a charity because I judge that its cause is worthwhile, I take its cause being worthwhile as a normative reason to volunteer, even if I do not conceptualize this as a reason and even if I am wrong that the cause is worthwhile. The capacity to make and act on judgments about normative reasons should not be confused with the capacity to reason instrumentally. Normative reasons are justifying reasons, and the capacity to make and act on judgments about normative reasons allows agents to reason about which actions are justified and which goals to pursue, not just about how to achieve their goals.
effectively.

Unlike agents who can only reason instrumentally, agents with the capacity to make and act on judgments about normative reasons can act wrongly. This is because, when agents make judgments about normative reasons and act accordingly, they and their actions can reasonably be evaluated based upon how well they do with respect to the actual reasons in play. To see this, consider an agent with badly mistaken judgments about normative reasons. We can reasonably say more about this agent than that he makes incorrect judgments, since his judgments about reasons at least partially constitute virtues and vices – i.e. they at least partially constitute his quality of will. For example, cruelty is partly constituted by, say, the normative judgment that someone else’s pain is insignificant. Thus an agent’s judgments about reasons, in conjunction with other features such as his desires, license appraisals of him in terms of virtues and vices. Now consider the agent’s actions. Because his judgments about reasons are badly mistaken, we can reasonably say that his resulting actions are badly mistaken. Since what makes them badly mistaken is that they are contrary to moral reasons, we should call these badly mistaken actions morally wrong. Lions and infants, on the other hand, do not make judgments about normative reasons. Hence they do not make and act on badly mistaken normative judgments, and so they do not have a quality of will and cannot act wrongly.

One might object that, to act wrongly, an agent must be subject to moral reasons, which requires the capacity to recognize and respond to moral reasons (or whatever capacity is required for accountability). This objection trades on an ambiguity about what it means to be subject to moral reasons. If being subject to moral reasons means that you and your actions can appropriately be evaluated according to moral reasons, then
the above argument shows that you need not have the capacity to recognize and respond to moral reasons (or whatever capacity is required for accountability) in order to be subject to moral reasons in this sense. Hence some agents who meet exempting conditions can be subject to moral reasons in this sense and so can act wrongly. On the other hand, if being subject to moral reasons means that others can appropriately hold you to complying with moral reasons, then the above argument shows that you can act wrongly without being subject to moral reasons in this sense. The same holds if you substitute ‘moral obligations’ for ‘moral reasons.’ The above argument demonstrates that some exempt agents, like psychopaths, are between lions and accountable agents in being subject to moral obligations in one sense but not another: they are subject to moral obligations in the sense that they and their actions can be evaluated with respect to those obligations, but they are not subject to moral obligations in the further sense of its being appropriate to hold them to fulfilling those obligations.

One might wonder whether Darwall could reply that he is giving an account of what it is to wrong another, not what it is to act wrongly. If so, Darwall could agree that psychopaths act wrongly in manipulating others but would claim that they do not wrong others, since wronging others requires being in accountability relations with them. I doubt that this reply would work. It seems reasonable for a victim of a psychopath’s manipulations to claim not only that the psychopath acted wrongly but also that he wronged her. Yet I need not press this point. Darwall is clear that he is giving an account of what it is to act wrongly. To see this, note that wrong actions violate moral obligations period, whereas wrongings violate ‘bipolar moral obligations’ to others. Darwall explicitly distinguishes his account of bipolar obligation from his view that we

3 I thank an anonymous reviewer for this objection.
are considering, saying ‘[i]n The Second-Person Standpoint and elsewhere I have argued that moral obligation period is a second-personal concept…’ (2012: 334). He then adds that ‘second-personal’ is consistent with impersonal reactive attitudes like indignation, and hence “[s]econd person” does not mean “second party” (2012: 336). Since Darwall’s view presented here is a view of moral obligation period, he is committed to claiming that agents who meet exempting conditions cannot act wrongly. Moreover, Darwall cannot simply shift to bipolar obligation to overcome the objection that I have raised, since the authority underlying his theory is, he says, a ‘representative authority’ that we have in virtue of being members of the moral community who can appropriately hold one another accountable to complying with moral obligations period (2012: 337).

The above argument is brief but serves its purpose. The goal is to support the intuition that some agents who meet exempting conditions, like young children, the mentally ill, and psychopaths, can act wrongly. The above reasonable explanation for why these agents, but not other exempt agents, can act wrongly gives us good reason to stand behind our intuition. Thus it is not merely a clash of intuitions if Darwall digs in his heels. He would need to show that these agents cannot act wrongly even though they can act cruelly and even though they are rational agents whose judgments about normative reasons and resulting actions can reasonably be evaluated relative to the actual reasons. It is doubtful that such an argument would work. We should therefore stand behind the intuition that some exempt agents can act wrongly and reject Darwall’s idea that morality is grounded in accountability relations.
1.3 The Second Objection: Accountability is Not Irreducibly Second-Personal

In this section, I aim to refute Darwall’s claim that accountability is irreducibly second-personal. If successful, the objection shows that moral obligation is not irreducibly second-personal either.

To start, recall that second-personal reasons are grounded in our authority to make claims and demands upon one another’s conduct. Of course, we lack the authority to make any claim or demand on someone’s conduct that we want to make. We have the authority to make certain claims and demands. Why do we have the authority to make some claims and demands but not others? For example, why do I have the authority to demand that you remove your foot from on top of mine but not the authority to demand that you buy me lunch? Darwall cannot plausibly say that it is just basic that I have the authority to make certain claims and demands but not others. Instead, I have the authority to make certain claims and demands but not others because, in part, reasons support certain claims and demands but not others. Further, the reasons supporting these claims and demands cannot themselves be second-personal. After all, second-personal reasons derive from our authority to make certain claims and demands, and so they cannot determine which claims and demands we have the authority to make. The reasons grounding legitimate claims and demands must therefore be third-personal reasons. Hence the second-personal reasons that Darwall says are entailed by moral obligation and wrongdoing must be grounded in third-personal reasons.

Darwall seems to agree. He says:

… many of the reasons that ground claims of wrong and obligation are not themselves second-personal. That an action would cause severe harm, or even pain to your bunions,
is a reason for someone not to do it, whether or not anyone has any standing to demand that he not, and it supports, moreover, a relevant demand. But the action cannot violate a moral obligation unless such a standing exists, so any reason that is entailed by the moral obligation must be second-personal. (2006: 99–100)

Here Darwall insists that moral obligations entail second-personal reasons. Yet he also says that third-personal reasons, such as that an action would cause severe harm, ground claims of wrong and obligation and support demands. Elsewhere, Darwall states that the reasons that make an action wrong – the wrong-making features – are not (typically) second-personal: ‘[w]rong-making features, such as that an action would cause harm…, themselves entail nothing about legitimate demands and thus are not second-personal reasons’ (2010: 151). It makes sense that wrong-making features are third-personal because, as just mentioned, third-personal reasons must support some claims and demands as legitimate and others as illegitimate. (Two clarifications are in order. First, Darwall says that some wrong-making features, such as that an action would violate someone’s rights, are second-personal. Yet as Darwall adds, second-personal wrong-making features are not conceptually independent of wrongness (2010: 151), and for the argument given above, they must themselves be grounded in third-personal reasons that support their associated demand as legitimate. I therefore set aside this nuance. Second, third-personal reasons may appeal to essentially relational facts. They are just not derived from our authority to hold others accountable.)

However, if the wrong-making features of actions are third-personal, as Darwall admits, why not think that wrongness is a third-personal concept? Darwall’s answer appeals to a conceptual connection between obligation and accountability: ‘[w]e hesitate to impute wrongdoing unless we take ourselves to be in the range of the culpable...’ (2006: 93). So Darwall’s idea is that, by tying obligation to second-personal reasons
rather than to the supporting third-personal reasons, we explain why only accountable agents can act wrongly. The first objection above shows the flaw in this reasoning: it is not the case that only accountable agents can act wrongly. Hence Darwall’s rationale for denying that wrongness is third-personal does not withstand scrutiny.

Yet even if we grant that only accountable agents can act wrongly, a second problem emerges. For moral obligation and wrongness to be irreducibly second-personal, accountability must be too. Hence, for Darwall’s argument to succeed, the capacity required for accountability must make ineliminable reference to our authority to make claims and demands upon one another. This means that our being accountable agents cannot be grounded in the capacity to recognize and respond to the third-personal reasons that ground claims of wrong and obligation and that support demands. After all, if accountability were so grounded, our authority to make claims and demands upon one another’s conduct would be like third-personal epistemic authority, which is rooted in our capacity to recognize and respond to third-personal epistemic reasons. I now argue that accountability is plausibly rooted in the capacity to recognize and respond to the third-personal reasons that ground claims of wrong and obligation, thereby undermining Darwall’s argument that morality is fundamentally second-personal.

First notice that, as accountable agents, we must have the capacity to determine which claims and demands we can legitimately make. Without such a capacity, we would surely lack the authority to make claims and demands at all. Thus being an accountable agent requires the capacity to recognize and respond to the third-personal reasons that ground claims of wrong and obligation and that support demands.
Does being accountable also require an irreducibly second-personal element? Darwall says yes. Pointing to Hobbes’s distinction between command and counsel, Darwall claims that the capacity to recognize and respond to third-personal reasons gives us the authority to counsel others about their conduct but not the authority to demand certain conduct from them (2006: 12). Epistemic authority, after all, is grounded in our capacity to recognize and respond to third-personal epistemic reasons, but only gives us the authority to counsel others. Thus, according to Darwall, accountability requires an irreducibly second-personal element to give us the authority to make claims and demands of one another.

Although this argument sounds plausible, it is incorrect. The capacity to recognize and respond to the third-personal wrong-making features of actions, along with the capacity to recognize this capacity in others, are plausibly sufficient to give us the authority to hold others accountable. We need not postulate an irreducibly second-personal element to explain that authority. Here is why: these third-personal capacities are plausibly sufficient to respond appropriately to others with reactive attitudes, and if we can appropriately have reactive attitudes towards others, then we can appropriately hold others accountable and so have the authority to hold others accountable. (After all, without such authority, we surely could not appropriately hold others responsible.)

To see that these third-personal capacities are sufficient for appropriately having reactive attitudes, first consider how we should understand the appropriateness of reactive attitudes. I claim that we should understand their appropriateness like the appropriateness of other emotions: as fittingness. An emotion is fitting if and only if it
accurately presents its target as having certain features. To illustrate, consider envy and amusement. Envy presents its target as enviable – (roughly) as having something good that the one envying lacks, and envy is fitting if and only if its target is enviable. Amusement presents a joke as amusing, and amusement is fitting if and only if the joke is amusing. As D’Arms and Jacobson (2000) argue, we must distinguish the moral rightness and fittingness of emotions. Even if it is morally wrong to envy or to be amused by offensive jokes, these emotions may nevertheless be fitting: the one envied may be enviable, and the offensive joke may be amusing.

Now consider reactive attitudes. As emotions, reactive attitudes also present their targets as having certain features: features that make them blameworthy in the case of negative reactive attitudes like resentment and features that make them praiseworthy in the case of positive reactive attitudes like gratitude. Reactive attitudes are appropriate in the sense of fitting if and only if they accurately present their targets as having those features.

Thus, to appropriately have reactive attitudes, we need only be capable of having our reactive attitudes accurately present their targets. To determine what this takes, consider resentment. My resentment presents its target as having acted wrongly from insufficient good will toward me (so that he lacks an excuse) while possessing the capacities, whatever they are, that make him an accountable agent (so that he lacks an exemption). Resentment may also present its target in other ways, such as having acted as the one resenting would not have acted, but we can set such features aside, as they are relevant to the resenter’s standing to resent, not to whether the agent is blameworthy.

---

4 See D’Arms and Jacobson 2000. I borrow the examples of envy and amusement from them.
Now, for the sake of argument, make this supposition: accountability not only requires but is grounded in the capacity to recognize and respond to the third-personal reasons that ground claims of wrong and obligation. Given this supposition, resentment would present its target as having acted wrongly from insufficient good will toward me while possessing the capacity to recognize and respond to the third-personal reasons that ground claims of wrong and obligation. On this supposition, for my resentment to present its target accurately, I just need the following capacities (in addition to the basic capacities to recognize other agents and their intentions): a) the capacity to recognize and respond to the third-personal reasons that ground claims of wrong and obligation, so that I can see that the agent acted wrongly from insufficient good will and b) the further capacity to recognize that capacity in him, so that I can see that he lacks an exemption. With just these capacities, my resentment can accurately present him and therefore be appropriate. The upshot is this: if accountability is grounded in the capacity to recognize and respond to the third-personal reasons that ground claims of wrong and obligation, we just need the capacity to recognize and respond to those third-personal reasons, and the capacity to recognize that capacity in others, in order to respond appropriately to others with reactive attitudes.

Hence, contrary to Darwall, the capacity to recognize and respond to third-personal wrong-making features (and the capacity to recognize this capacity in others) can give us more than the authority to counsel others; it can give us the authority to hold others accountable by giving us what we need to respond appropriately to them with reactive attitudes. There is therefore no need to postulate an irreducibly second-personal

---

5 We need not determine how resentment presents its target as having that capacity. All that matters is that it must, even for Darwall. After all, as noted above, this capacity is required for being an accountable agent, making resentment inappropriate if the agent lacks that capacity.
element to explain our authority to make claims and demands. It would be superfluous. We get a more straightforward and less mysterious picture by grounding accountability in the capacity to recognize and respond to the third-personal reasons that ground claims of wrong and obligation and that support demands.

Notice that our authority to make claims and demands upon one another’s conduct differs from the sergeant’s authority to make claims and demands upon her troops. The sergeant’s authority requires more than the capacity to recognize and respond to the third-personal reasons that support her claims and demands and the capacity to recognize this capacity in her troops. She must also be formally placed in a position of authority over her troops. This makes sense because the sergeant has an institutional authority, and there are good reasons for the institution to grant a sergeant’s authority only to some people with those capacities. Our authority to make claims and demands upon one another’s conduct is different. It is not an institutional authority, and there is no reason to limit it only to some people with the requisite third-personal capacities.

I have argued that postulating an irreducibly second-personal authority to make claims and demands upon one another’s conduct is superfluous. This is because, once we have the capacity to recognize and respond to the third-personal reasons that ground claims of wrong and obligation and that support demands, and the capacity to recognize this capacity in others, we have all that we need to appropriately hold others accountable for their conduct via reactive attitudes, and we therefore have the authority to hold one another accountable.

Perhaps surprisingly, Darwall seems committed to the premises of the above argument. To start, he adopts a fitting-attitude theory of normative concepts, which
explicates normative concepts in terms of the fittingness of the attitude(s) distinctive of the concept. For example, the concept of the desirable, he says, is that of being a fitting object of desire (2010: 139). Darwall then explicates the concepts of moral obligation and wrongness in terms of the fittingness of reactive attitudes. That is why, for Darwall, a wrong action is one for which an agent is blameworthy – and so appropriately subject to negative reactive attitudes – absent an excuse (2010: 142–3). Darwall then claims that the wrong-making features of actions make reactive attitudes fitting, not wrongness itself:

… (the wrong-making features [of an act]) consist in whatever reasons (of the right kind) there are for the reactive attitudes through which one holds someone responsible for performing an act of that kind. By analogy with our earlier point about the desirable, something does not become more blameworthy than its wrong-making features make it by virtue of being morally wrong; its being wrong just is its being blameworthy (if done without excuse) to whatever degree its wrong-making features make it. (2010: 143)

Now recall that the wrong-making features of actions are third-personal reasons. Hence Darwall is committed to the idea that third-personal reasons, not second-personal reasons, make reactive attitudes fitting. This is precisely my point. Yet Darwall does not notice what follows: if the third-personal wrong-making features of actions make reactive attitudes appropriate, then our authority to hold others accountable is not irreducibly second-personal. It rather depends just upon our capacity to recognize and respond to those third-personal wrong-making features of actions (and our capacity to recognize that capacity in others). Because accountability is not irreducibly second-personal, Darwall’s argument that moral obligation and wrongness are irreducibly second-personal fails.

To be clear, I do not deny the existence of second-personal reasons. We have second-personal reasons not to act in certain ways because people have the authority to hold us accountable for acting in those ways. The point is that wrongness itself does not
provide second-personal reasons. Instead, the fact that others can appropriately hold me accountable for performing wrong actions gives me a reason – a second-personal reason – not to act wrongly. Hence Darwall is right to say that the fact that others can blame me for acting in some way gives me a second-personal reason not to act in that way. He errs in claiming that this second-personal reason is entailed by the act’s wrongness.\(^6\)

Before concluding this section, consider a worry that brings us back to how this section began: if our authority to make claims and demands is grounded in our capacity to recognize and respond to third-personal reasons, why do I have the authority to demand that you remove your foot from mine but not that you buy me lunch? After all, good third-personal reasons support both demands.\(^7\) The answer is this. I have the authority to demand only the former because, according to the correct (third-personal) moral theory, you act wrongly and wrong me in refusing to remove your foot from mine but not in refusing to buy me lunch, and so resentment toward you is appropriate only for the former.

To summarize, I have presented two independent objections to Darwall’s idea that morality is grounded in accountability relations. First, I argued that some agents who meet exempting conditions, and so who cannot participate in accountability relationships, can nevertheless act wrongly. Second, I argued that Darwall’s idea that accountability is irreducibly second-personal is implausible because it is superfluous: we can reasonably ground accountability in our capacity to recognize and respond to the third-personal reasons that ground claims of wrong and obligation. Thus morality is not fundamentally

---

\(^6\) See Darwall 2010: 151. Using the conceptual possibility of supererogatory actions, Darwall argues that an act’s wrongness provides a reason not to do it that is additional to those provided by the act’s wrong-making features (Darwall 2010: 147–8). If his argument is correct, I have shown that the additional reason provided by the act’s wrongness is not second-personal.

\(^7\) I thank an anonymous reviewer for pressing this worry.
Responsibility and Normative Moral Theories

second-personal either. Because morality is not grounded in accountability relations, Darwall’s argument that morality is fundamentally interpersonal does not work.

2. Wallace on Reasons for Responsibility Reactions

In this section, I show that, contrary to R. Jay Wallace, the necessary connections between morality and responsibility that he posits do not rule out non-relational moral theories.

2.1. Internalism about Responsibility

Wallace (2013) argues for internalism about responsibility, the view that there is a necessary connection between morality and responsibility. To flesh it out, Wallace first discusses internalism about motivation, the view that there is a necessary connection between morality and motivation. Wallace claims that we should understand the necessary connection between morality and motivation normatively: moral properties provide reasons for action, and so agents are necessarily motivated to act morally insofar as they are deliberating correctly and are practically rational. Wallace then models internalism about responsibility after internalism about motivation, claiming that morality and responsibility are normatively connected in two respects. First, we all have reason to care about moral values, which makes us susceptible to reactive attitudes when those values are thwarted. Second, the moral properties of actions provide defeasible reasons for responsibility reactions.
The second respect is directly analogous to internalism about motivation. Wallace’s idea is that, just as moral properties provide reasons for action, they provide reasons for responsibility reactions. He says, ‘moral wrongness is not only a reason for action (a reason, specifically, to avoid actions that have the property) but also a reason for responsibility reactions (a reason, specifically, for blaming the agent who performs an action with this property)’ (2013: 234). The first respect, in contrast, fits the model of internalism about motivation only insofar as it postulates a necessary normative connection between morality and responsibility. After all, it postulates not a reason for responsibility reactions but rather a reason to care about moral values, which makes us prone to responsibility reactions.

According to Wallace, these normative connections have implications for determining the correct moral theory. Start with the first. As Wallace points out, utilitarianism cannot account for the fact that we all have reason to care about moral values, since it may not maximize well-being to care about utilitarian standards. Wallace takes this to be a problem for utilitarianism; if he is right, consequentialism generally faces the same problem. Yet as Wallace admits, a version of perfectionism can account for this connection. According to perfectionism, we should promote our individual flourishing, and if we stipulate that caring about morality is a virtue, this version of perfectionism would say that we all have reason to care about morality. Thus a moral theory need not be fundamentally interpersonal to account for the first normative connection.

Turn to the second: moral properties provide defeasible reasons for responsibility reactions. Wallace argues that this connection implies that morality is fundamentally
interpersonal. To understand his argument, we must first see how Wallace clarifies and refines this connection.

To clarify, when Wallace says that moral properties provide reasons for responsibility reactions, he means ‘reasons of the right kind’ – i.e. reasons that make responsibility reactions appropriate. After all, Wallace models internalism about responsibility on internalism about motivation, and regarding the latter, he says that moral considerations, being reasons for action, make motivation to comply with them appropriate (2013: 228). Further, Wallace claims that reactive attitudes are appropriate responses to wrong actions and that the reasons for them are like the reasons that we have to experience other emotions (2013: 232). Thus, for Wallace, moral considerations make reactive attitudes appropriate and so are reasons of the right kind for responsibility reactions.

Now for the refinement. Wallace admits that the idea that moral considerations are reasons for responsibility reactions is not quite right. To refine this idea, he looks again to internalism about motivation. By noting that reasons for action are reasons for intention, he refines internalism about motivation as follows: moral properties provide reasons for intention. For example, the wrongness of an action provides a reason to intend not to do it. Similarly, Wallace says that we should refine internalism about responsibility to say this: agents’ attitudes toward moral properties are reasons for responsibility reactions – e.g. an agent’s indifference toward wrongness is a reason for resentment. This refined version is superior because blame is properly responsive not to the moral qualities of actions but to agents’ reasons for acting, since blame is inappropriate toward someone who acts wrongly but has an excuse.
According to Wallace, individualistic moral theories are flawed because they cannot explain why agents’ attitudes toward moral properties are reasons for responsibility reactions. Individualistic moral theories cannot explain this because the reasons that they posit are reasons of the wrong kind for reactive attitudes. Consider consequentialism. On consequentialism, it is not agents’ attitudes toward the moral properties of their actions that gives us reason to respond to them with responsibility reactions but rather the desirability of those reactions. Thus consequentialism cannot explain the connection at issue. Or consider perfectionism. On perfectionism, it is not agents’ attitudes toward the moral properties of their actions that gives us reason to respond to them with responsibility reactions but rather the contribution that having those reactions makes to our flourishing. Hence perfectionism cannot explain the connection at issue either. As Wallace puts the point with respect to utilitarianism and perfectionism: ‘there is nothing in the nature of such requirements that would seem to explain why our attitudes toward them have the normative significance for responsibility reactions that internalism postulates’ (2013: 241).

In contrast, relational theories of morality can explain why agents’ attitudes toward moral properties are reasons for responsibility reactions. On a relational theory, other agents’ attitudes can figure into the reasons supplied by the moral theory, and so the moral theory can supply reasons of the right kind for reactive attitudes. For example, Wallace says that if moral requirements are grounded in the claims of other agents, then indifference to moral requirements is indifference to the claims of other agents, and reactive attitudes are appropriate responses to such indifference. Hence, Wallace claims,
the second normative connection supports a fundamentally relational understanding of morality.

Wallace is correct that consequentialism cannot explain the fact that we all have reason to care about moral values. He is also correct that individualistic moral theories cannot explain the fact that agents’ attitudes toward moral properties are reasons for responsibility reactions. However, Wallace is wrong to think that moral theories must explain these normative connections. I take each normative connection in turn.

2.2. The First Normative Connection

Grant that we all have reason to care about moral values, which makes us prone to responsibility reactions. As we have seen, if a moral theory must explain this normative connection, as Wallace thinks, consequentialism would be flawed because, according to consequentialism, we may not have reason to care about moral values, as it may be best not to care about them.

Contrary to Wallace, a moral theory need not explain this normative connection. Here is why: we all have reason to care about moral values in the sense that it is appropriate or fitting for everyone to care about, or value, moral values. This is because the nature of those features that make moral values valuable gives us all reason – reason of the right kind – to care about moral values.

Wallace seems to anticipate and reject this response. After claiming that we all have reason to care about moral values, he says that moral values are different than many other values in this respect, such as the values associated with philosophy and cabinetmaking, since the values of these activities are insufficient to give everyone
reason to care about them (2013: 231). He clearly takes this point to generalize: the value of something is insufficient to give everyone reason to care about it. If this generalization were correct, the nature of moral values would not give everyone reason to care about them, and it would seem that a moral theory must account for this reason.

The generalization is implausible. To see why, notice that some values provide interest- and skill-dependent reasons to care about them. In other words, they provide reasons to care about them only to those with the requisite interests and skills. The values associated with philosophy and cabinetmaking are of this type. If I have no interest in philosophical questions or struggle to operate a saw, the values of these activities give me no reason to invest emotionally in these activities. However, if I enjoy engaging with philosophical questions and have the aptitude for it, then the value of philosophy gives me reason to invest emotionally in philosophy. Other values provide interest- and skill-independent reasons to care about them – i.e. they provide reasons to care about them to everyone, regardless of interests and skills. Intuitively, moral values are of this type. It is implausible that moral values provide reasons to care about them only to those with certain interests and skills, just like it is implausible that moral requirements apply only to those with certain interests and skills.  

This distinction undermines Wallace’s generalization. The values associated with philosophy and cabinetmaking do not give everyone reason to care about those activities because those values provide interest- and skill-dependent reasons, not because values are generally insufficient to provide reasons to care about them. Because moral values

---

8 Of course, we must have certain basic rational competences in order for reasons and moral requirements to apply to us. I discuss what I think those rational competences are in Section 1.2. So we should read ‘skills’ here to mean skills beyond basic rational competences. Wallace too must accept such a caveat in saying that we all have reason to care about moral values.
provide interest- and skill-independent reasons, the nature of moral values gives everyone reason to care about moral values. Putting the point in terms of fittingness, it is fitting for those with the requisite interests and skills to care about philosophy and cabinetmaking, whereas it is fitting for everyone to care about moral values. Hence the fact that it is fitting for everyone to care about moral values explains why we all have reason to care about moral values; a moral theory need not explain this.

Notice that caring about moral values involves having either promotion or respect attitudes toward those values, depending upon which moral theory is correct. For example, on utilitarianism, someone who cares about moral value desires to promote well-being, and on deontology, someone who cares about moral value has respect attitudes toward rational agency. Hence, depending upon which moral theory is correct, it is fitting to have either promotion or respect attitudes toward what is of moral value. You may wonder what the relationship is between these fitting attitudes and the reasons for action postulated by moral theories.

One possibility is to define reasons for action in terms of the attitudes that it is fitting to have toward moral values.\(^9\) Hence, if it is fitting to have promotion attitudes toward moral values, then we have reason to perform some action insofar as it promotes moral values, and if it is fitting to have respect attitudes toward moral values, then we have reason to perform some action insofar as it respects moral values. On this option, the disagreement between consequentialism and deontology comes down to what it is fitting to care about – specifically, whether it is fitting to have promotion or respect attitudes toward what is of moral value – and reasons for action follow suit. This option accounts for the fact that, according to consequentialism, it is a contingent matter whether

---

\(^9\) I thank Michael Smith for raising this possibility and for helpful discussion.
we have reason to acquire the fitting promotion attitudes, as acquiring them may not promote moral values.

Another possibility is that our reasons for action are fixed independently of the attitudes that it is fitting to have toward moral values. Consequentialists think that our reasons for action are fixed by what promotes value, and therefore it is contingent whether we should have promotion attitudes toward moral values, since having those attitudes may not promote value. Deontologists, in contrast, think that we should do what respects value and have attitudes that respect value. Hence, for deontologists, it is not a contingent matter whether we should have fitting attitudes toward value.

On either option, the important point is this: consequentialism can reasonably say both that it is fitting to have promotion attitudes toward moral values and that it is a contingent matter whether we have (moral) reason to have promotion attitudes toward moral values. As a result, consequentialism need not and indeed should not explain why we all have reason to care about moral values. This is because, to explain this reason when it is already explained by the fittingness of caring about moral values would, for the consequentialist, be to conflate moral rightness and fittingness.

Of course, for consequentialism, a tension emerges when it is undesirable to care about moral values. In such cases, a rational agent would care about consequentialist moral values based on the reasons of the right kind to care about them but would also try to stop caring about those values, since caring about them is undesirable. Yet, because desirability is a reason of the wrong kind to care or stop caring about moral values, the rational agent could not stop caring about them on the grounds of desirability. Here consequentialism must be self-effacing. To stop caring about moral values, and so act
rightly, the agent must somehow induce in himself a false belief to the effect that those values are not worth caring about – i.e. he must find a way to believe falsely that the reasons of the right kind for caring about those values do not hold. The same applies to any emotion or attitude. Suppose that it would be desirable to envy someone who is unenviable. You could not envy him on those grounds, since the desirability of envy is a reason of the wrong kind for it. Thus, to do the right thing by consequentialist lights, you would somehow have to induce in yourself a false belief that, say, he had something desirable that you do not. Consequentialists are driven to be self-effacing in order rightly to avoid being revisionary about emotions and other attitudes, since as we have seen, emotions and attitudes present their targets in certain ways and so have internal correctness conditions that may be in tension with moral reasons.

Is it a problem for consequentialism to be self-effacing in this regard? I doubt it, but that is a separate issue. To show that a moral theory should not be self-effacing in this regard would require a distinct argument that, on the correct moral theory, the moral reason for caring about moral values must be able to ground caring about moral values.

2.3. The Second Normative Connection

Recall the second normative connection between morality and responsibility: agents’ attitudes toward the moral properties of their actions are reasons for responsibility reactions – e.g. an agent’s indifference toward the wrongness of his action is a reason for resentment. According to Wallace, this normative connection supports a fundamentally interpersonal moral theory. This is because individualistic moral theories like consequentialism posit reasons of the wrong kind for responsibility reactions and
therefore cannot explain why agents’ attitudes toward moral properties are reasons for responsibility reactions.

Wallace is right that individualistic moral theories cannot explain this normative connection. However, contrary to Wallace, this is not a problem for those theories. To show this, I focus on consequentialism. Other individualistic moral theories can make a similar argument.

To start, return to internalism about motivation. On that view, the moral properties of an action are reasons to perform or not to perform *that action* or, on the refined version, reasons to intend to perform or to intend not to perform *that action*. (For simplicity, I stick with the original version. You can substitute the refined one if you wish.) Granting internalism about motivation, consequentialism must say that the fact that an action A fails to maximize value is a reason to avoid performing A. Now apply this to responsibility reactions. According to internalism about motivation, the moral properties of *responsibility reactions* provide reasons for responsibility reactions – e.g. the wrongness of resenting someone provides a reason not to resent him. Consequentialism can account for this by saying that the fact that resenting someone fails to maximize value provides a reason not to resent him.

Now ask: must consequentialism *also* explain the fact that agents’ attitudes toward the moral properties of their actions are reasons for responsibility reactions? Importantly, such an explanatory demand does not follow from internalism about motivation. Internalism about motivation focuses on the moral properties of *responsibility reactions themselves* as reasons to have or not have them, not on the moral properties of the actions performed by the targets of responsibility reactions.
You might think that consequentialism must explain why agents’ attitudes toward moral properties are reasons for responsibility reactions on the grounds that a moral theory must account for all of the reasons for responsibility reactions. That is not so. Consequentialists, once again, can distinguish moral rightness and fittingness. They can say that desirability is a reason for responsibility reactions that makes responsibility reactions morally right and that agents’ attitudes toward the moral properties of their actions are reasons for responsibility reactions that make responsibility reactions fitting. The nature of reactive attitudes supports this idea. As discussed, reactive attitudes present their targets as having certain attitudes toward the moral properties of their actions by presenting their targets as having a certain quality of will, and hence the fittingness of reactive attitudes depends upon agents’ attitudes toward the moral properties of their actions. The upshot is that consequentialism need not explain why agents’ attitudes toward the moral properties of their actions are reasons for responsibility reactions. Indeed, consequentialism should not explain this. After all, the fittingness of responsibility reactions already explains it, and so to explain it otherwise, according to consequentialism, would be to conflate the fittingness and moral rightness of responsibility reactions.

You might worry that, if desirability is a reason to have reactive attitudes according to consequentialism, it must contribute to making reactive attitudes appropriate. Not so. According to internalism about motivation, the rightness of having a reactive attitude is a reason that makes appropriate the motivation to have it, and so for consequentialism, the desirability of having a reactive attitude makes appropriate the motivation to have it. But the appropriateness of this motivation is independent of the
appropriateness of the reactive attitude itself. A reactive attitude’s appropriateness is determined solely by the reasons of the right kind for it – i.e. by whether it accurately presents its target – not by the reasons of the right kind to be motivated to have it. Thus consequentialism can consistently say that it is appropriate to be motivated to have a desirable reactive attitude because it is morally right, even if that reactive attitude is inappropriate. (Of course, if a reactive attitude is appropriate, the reasons of the right kind to have it would also make appropriate the motivation to have it, but it would not be moral motivation according to consequentialism.)

Again, a tension exists for consequentialism. Because desirability is a reason of the wrong kind to have a reactive attitude, we cannot acquire a reactive attitude based on what consequentialism claims is a reason of the right kind to be motivated to have it. To account for this, consequentialism must be self-effacing. It must acknowledge that the moral motivation to have a reactive attitude that it posits cannot successfully produce the reactive attitude but must instead aim at inducing in the agent the belief that the reasons of the right kind to have the reactive attitude exist. The same holds for any emotion. As above, if a moral theory should not be thus self-effacing, that is a separate issue. To show it would require a distinct argument that, on the correct moral theory, the moral motivation to have a reactive attitude (or any emotion) must be capable of aiming at and producing that reactive attitude (or that emotion).

In this section, I demonstrated that, contrary to Wallace, consequentialism need not and should not explain a) why we all have reason to care about moral values and b) why agents’ attitudes toward the moral properties of their actions are reasons for responsibility reactions. We all have reason to care about moral values because it is
fitting for everyone to do so, and agents’ attitudes toward the moral properties of their actions are reasons for responsibility reactions in that they make responsibility reactions fitting. Hence these normative connections do not rule out consequentialism or any individualistic moral theory. They do not show that morality is fundamentally interpersonal.

3. The Strategy Does Not Work

So far, I have dismissed two arguments purporting to show that morality is fundamentally interpersonal by appealing to a necessary connection between morality and responsibility. From the failure of these arguments, we can reasonably conclude that the general strategy does not work. Because moral theories are not grounded in accountability relations (contrary to Darwall) and need not explain the reasons of the right kind for responsibility reactions (contrary to Wallace), it is not a strike against a moral theory if the reasons that it supplies are reasons of the wrong kind for responsibility reactions. Individualistic moral theories, whose reasons are reasons of the wrong kind for responsibility reactions, cannot therefore be ruled out on those grounds.

Further, it is implausible to suppose that another necessary connection between morality and responsibility rules out some or all non-relational theories. As discussed above, contrary to Wallace, moral theories need not explain why we all have reason to care about moral values, since the fittingness of caring about moral values explains this reason, and thus this necessary connection does not count against consequentialism. Further, it is difficult to see what other necessary connection between morality and responsibility could undermine a moral theory. After all, what stance besides caring
about moral values could make us susceptible to responsibility reactions? And if there were such a stance and we all had reason to adopt it, it is reasonable to think that we could again explain this reason by appealing to the fittingness of adopting that stance. Thus, moral theories need not explain the reasons of the right kind for reactive attitudes, nor the reasons that we all have to adopt stances that make us prone to reactive attitudes. We should therefore conclude that no necessary connection between morality and responsibility exists that would rule out any non-relational moral theory.

Of course, a moral theory must give the correct results, saying that an action is right if and only if it is right. You might wonder whether interpersonal theories better account for the rightness of responsibility reactions. After all, if a moral theory posits reasons of the right kind for reactive attitudes, it says that holding someone accountable is morally right if and only if it is appropriate, which intuitively seems correct. Yet some individualistic moral theories also give this result. For example, perfectionism could say that it is a virtue to hold someone accountable if and only if it is appropriate to do so. So any difficulty would be for theories, such as consequentialism, on which the fittingness and rightness of responsibility reactions may come apart.10

If this is a problem for consequentialism, it is a distinct one. Charging consequentialism with giving unintuitive results in particular cases is different than saying that it cannot account for necessary connections between morality and responsibility. After all, when giving unintuitive results about an action’s rightness, consequentialism can either be revisionary by claiming that we should jettison our intuitions or work to give the intuitive results.

---

10 I thank Tristram McPherson for pressing this worry.
For what it is worth, I doubt that consequentialism faces a problem in claiming that the rightness and fittingness of a responsibility reaction may come apart. Notice that consequentialism can credibly account for the following intuition: that participating in our moral responsibility practice is morally justified, in the sense that we are morally justified in following its standards. This is because participating in our responsibility practice plausibly maximizes value even though the practice has non-consequentialist standards, just as participating in friendship with its non-consequentialist standards plausibly maximizes value. After all, just as friendship expresses our social nature, so does our responsibility practice. As Strawson says, ‘our [responsibility] practice does not merely exploit our nature but expresses it…’ (1962/2003: 93). It would therefore be psychologically costly, perhaps impossible, not to engage in our responsibility practice, and thus participating in it plausibly maximizes value. I suspect that this is all that consequentialism needs to avoid being problematically revisionary. After all, if we buy into our responsibility practice, as we morally should, this will affect our intuitions about the moral rightness of following its standards in particular cases. Although this argument needs fleshing out, doing so is beyond the scope of this paper.

4. Conclusion

Stephen Darwall and R. Jay Wallace have independently argued that a necessary connection exists between morality and responsibility that demonstrates that morality is fundamentally relational. As I have shown, their arguments do not succeed. I gave two independent arguments against Darwall’s idea that morality is grounded in accountability relations. First, I argued that, contrary to that idea, some agents who lack the capacities
to participate in accountability relations can act wrongly. Second, I demonstrated that accountability is not plausibly irreducibly second-personal, since we can reasonably ground accountability in the capacity to recognize and respond to the third-personal reasons that ground claims of wrong and obligation and that support demands. As a result, moral obligation is not plausibly irreducibly second-personal either.

Next, I took up Wallace’s argument. I agreed that the following necessary connections exist between morality and responsibility: 1) we all have reason to care about moral values, which makes us susceptible to responsibility reactions, and 2) agents’ attitudes toward the moral properties of their actions are reasons for responsibility reactions. However, contrary to Wallace, I argued that a moral theory need not explain these connections, as they are explained by the fittingness of the responses in question.

Finally, I demonstrated that, plausibly, no necessary connection between morality and responsibility exists that would undermine any non-relational moral theory. The arguments presented against Darwall and Wallace show that moral theories need not supply reasons of the right kind for reactive attitudes. Further, as with 1) above, if there were some reason for adopting a stance that makes us prone to responsibility reactions, it is reasonable to think that it can be explained by the fittingness of adopting that stance. The upshot is that Darwall and Wallace’s strategy does not succeed. We cannot convincingly argue against non-relational moral theories by appealing to necessary connections between morality and responsibility.11

11 I am indebted to Michael Smith for very helpful conversations and comments on multiple drafts of this paper. I am also grateful to Eden Lin, Victoria McGeer, Tristram McPherson, and anonymous reviewers for valuable comments on earlier drafts. I presented earlier versions of this paper to the New York Early Career Ethics Workshop in summer 2014 and the NYU Ethics Forum in spring 2017, and I would like to thank the participants of both for insightful questions and helpful discussions.
References


