

The Supererogatory, and How to Accommodate It[†]

A traditionally noted feature of act-consequentialism is that it doesn't seem to leave room for the supererogatory. Trouble is, supererogatory acts seem to exist. Urmson writes:

We may imagine a squad of soldiers to be practicing the throwing of live hand grenades; a grenade slips from the hand of one of them and rolls on the ground near the squad; one of them sacrifices his life by throwing himself on the grenade and protecting his comrades with his own body. [Doing so] is clearly an action having moral status. But if the soldier had not thrown himself on the grenade would he have failed in his duty? Though clearly he is superior in some way to his comrades, can we possibly say that they failed in their duty by not trying to be the one who sacrificed himself? If he had not done so, could anyone have said to him, "You ought to have thrown yourself on that grenade"? Could a superior have decently ordered him to do it? The answer to all of these questions is plainly negative.¹

Urmson's case appears to illustrate that, at least on occasion, individuals can act in ways that go beyond the call of duty. The soldier sacrificed himself in a way that clearly is morally good, but—*contra* act-consequentialism—is certainly not his *duty*.

So far, so plausible. But the category of the supererogatory resists easy analysis. Traditionally, supererogatory actions are understood as actions that are morally good, but not morally required, actions that go "beyond" the call of our moral obligations. As I shall argue in this paper, however, the traditional analysis can be accepted only by a view with troubling consequences concerning the structure of the moral point of view. I propose a different analysis that is extensionally correct, avoids the problems of the traditional view, and, incidentally, also defuses any objection to act-consequentialism, or any other first-order moral theory, on grounds that it cannot accommodate the supererogatory.

1. *The Traditional View*

[†][*Acknowledgements*].

¹J. O. Urmson, "Saints and Heroes" in *Essays in Moral Philosophy*, ed. Melden (Seattle, WA: University of Washington Press, 1958), 202-3.

Discussions of the supererogatory have generally assumed three analytic focal points, the conjunction of which I shall call “the traditional view”. Though some analyses further limit actions that can properly be called “supererogatory”, rarely do any views not accept the basic principles captured in the following brief statement by Rawls: “[Supererogatory acts] are acts of benevolence and mercy, of heroism and self-sacrifice. It is good to do these actions but it is not one’s duty or obligation. Supererogatory acts are not required, though normally they would be were it not for the loss or risk involved for the agent himself.”² The first principle noted by Rawls can be captured as follows:

Permissible not Required: If an act ϕ is supererogatory, ϕ is morally permissible, but is not morally required.

In other words, *Permissible not Required* holds that for any supererogatory action ϕ , it is morally acceptable to ϕ , but ϕ -ing is not the subject of a moral requirement. One is never morally required to perform a supererogatory action.

However, *Permissible not Required* is clearly insufficient to capture what it means for an action to go *beyond* the call of duty. One would not, for instance, describe the action of double-knotting my shoes as supererogatory. But double-knotting my shoes is surely morally permissible and also not morally required. The difference between saving one’s friends by jumping on a grenade and double-knotting one’s shoes is that the former and not the latter appears to have *positive* moral valence. The former and not the latter is:

Morally Good: If an act ϕ is supererogatory, ϕ is morally good, meritorious, or heroic.

Morally Good holds that supererogatory actions are not simply permissible, but have a particularly positive moral status.

Consider now the third feature of the traditional view, also noted by Rawls. Many hold that one essential feature of the supererogatory is that supererogatory actions are supererogatory in part *because* they involve some non-trivial sacrifice to the agent. Though this claim is controversially strong, a somewhat weaker claim is surely an important aspect of the traditional view:

²John Rawls, *A Theory of Justice* (Cambridge, MA: Harvard University Press, 1971), 117.

But for Sacrifice: A subset (S) of supererogatory actions would have been morally required but for the fact that they require non-trivial sacrifice on the part of the agent.

Take Urmson's soldier. It seems quite sensible to say that one of the essential facts that renders his action supererogatory is that this action is sacrificial—indeed, ultimately sacrificial. Were it the case that this soldier could have saved his friends *without* sacrifice, he certainly would have been morally required to do so. If I am in a position to donate half my yearly salary to Oxfam International, but only at significant cost to my own well-being, doing so is supererogatory. If my donations fail to affect my well-being, or affect it only trivially, making these donations is morally required.³

A note on *But for Sacrifice*. S is a proper subset of all supererogatory actions. Some—like myself—believe that most supererogatory actions will involve some sacrifice on the part of the agent. My view is motivated by abstraction from, e.g., Urmson's soldier. Others believe that not all supererogatory actions would be morally required were it the case that there were no agential sacrifice.⁴ It might be that some supererogatory actions are merely morally *favoured*, but under no circumstances required, even when so acting is costless. This view is motivated by abstraction from cases that seem “smaller” in significance: the case of a person including a stranger in a fun outing, for instance.⁵ In the remainder of the paper, I will limit my discussion to cases that fall *within* S (or, at least, that seem to do so given my considered judgments). These cases generate serious problems with the traditional view, even if the relative size of S is very small.⁶

The traditional view might be supplemented by a number of additional

³An anonymous reviewer notes that some accounts of the supererogatory allow many kinds of sacrifice, rather than prudential sacrifice. Of course, this is quite correct; though I take prudential sacrifice to be the paradigmatic example of sacrifice, but for which supererogatory actions (at least those that fall within S) would be required, I do not mean sacrifice to be limited strictly to prudential sacrifice. For instance, one might regard the risk of prudential sacrifice as itself not a prudential sacrifice, but a sacrifice sufficient to render an act that would be morally required supererogatory. In addition, the loss of one's capabilities, autonomy, or other non-welfarist goods might very well constitute sacrifice sufficient to render an otherwise required action supererogatory. However, this is largely beside the point for present purposes; even if we focus, strictly speaking, on prudential sacrifice, *But for Sacrifice* still holds: there is surely *some* subset of supererogatory actions that would be required but for the prudential sacrifice involved.

⁴See, for instance, Terry Horgan and Mark Timmons, “Untying a Knot from the Inside Out: Reflections on the ‘Paradox’ of Supererogation” in *Social Philosophy and Policy* 27 (2010), 47-8. Thanks to [...] for calling my attention to this issue.

⁵Horgan and Timmons, 47-8.

⁶See Horgan and Timmons, 50-59.

principles concerning the structure of supererogatory action. For instance, some⁷ might argue that the idea of the supererogatory should be limited to acts that are in some way beneficial, or are particularly praiseworthy. But *Permissible not Required*, *Morally Good*, and *But for Sacrifice* seem to be the beating heart of any traditional analysis of supererogatory actions, whether further principles are required or not.

2. *The Puzzle*⁸

James Drier notices a problem for the traditional view:

But seen another way [the supererogatory] is puzzling. Morality, we are inclined to think, is a matter of what reasons one has *from the moral point of view*. When there is a supererogatory act available, it would be better for you to perform it. So surely you have a reason, from the moral point of view, to perform the act. You may have some reason not to perform it, but at least typically you will have no reason *from the moral point of view* to refrain from it (if you do have some such reason, then it will ordinarily be outweighed by the reason you have to perform, because by hypothesis it is better to perform). But now it is hard to see how it could be permissible, from the moral point of view, to refrain from doing something that you have an undefeated reason (from that very point of view) to do. Everything from the moral point of view speaks in favor of [performing a supererogatory act], and nothing at all speaks against it. In what sense it is “all right,” “permissible,” “not wrong” to fail to act? There seems to be no sense at all.⁹

One way to state this puzzle more precisely is to introduce a modicum of terminology. First, consider the distinction between a *supererogatory* action, and a *merely erogatory* action. Though both supererogatory and merely

⁷See, for instance, David Heyd, *Supererogation: Its Status in Ethical Theory* (Cambridge: Cambridge University Press, 1982).

⁸This puzzle is well-known. See, e.g., Horgan and Timmons, 36-38. See also Douglas Portmore, “Are Moral Reasons Morally Overriding” in *Ethical Theory and Moral Practice* 11 (2008), 378-381.

⁹James Dreier, “Why Ethical Satisficing Makes Sense and Rational Satisficing Doesn’t” in *Satisficing and Maximizing: Moral Theorists on Practical Reason*, ed. Byron (Cambridge: Cambridge University Press, 2004), 148. Dreier, it seems to me, oversteps. There could be *some* moral reasons against a supererogatory act; the general point, however, holds. A supererogatory is certainly morally better than a merely erogatory act.

erogatory actions are permissible, supererogatory action goes “beyond” one’s duty. Merely erogatory action does not. To see the notion of a merely erogatory action more clearly, consider the following case. Imagine that you can react in one of three ways to a person down on her luck. You can assist her by going out of your way to buy her a nutritious meal. Second, you can offer her one dollar. Finally, you can do nothing. Assume that the third option is morally disallowed. Also, assume that the first is supererogatory. One might say, plausibly, that the second is “merely erogatory”; it counts as the fulfillment of one’s moral obligations, but not in a way that is particularly morally special.

Second, consider the idea, mentioned by Dreier, of a *moral reason*. A moral reason is a fact that counts in favor of a particular action ϕ from the moral point of view. In this way, a moral reason is similar to other domain-specific reasons, such as prudential reasons, aesthetic reasons, legal reasons, etc. Domain-specific reasons help determine what one is required to do from the perspective of that domain. Further, in any domain an act ϕ can be supported by reasons that are weaker, stronger, or that are equivalent in strength to another act ψ . When an act ϕ is supported by stronger reasons than another act ψ , I shall say that ϕ is *decisive with respect to* ψ . So, for instance, because there is stronger prudential reason for me to eat cherry pie than apple pie (because, say, I prefer to do so), eating cherry pie is prudentially decisive with respect to eating apple pie. If there is stronger *moral* reason to ϕ than ψ , ϕ -ing is morally decisive with respect to ψ -ing in just the same way. If a particular act ϕ is decisive with respect to all other actions from the point of view of a particular domain, I will say that this act is *overriding*. ϕ -ing is morally overriding, for instance, if ϕ -ing is morally decisive with respect to any other act ψ .

Given the traditional view, if ψ is merely erogatory, and ϕ is supererogatory, ψ must not be morally decisive with respect to ϕ . If ψ is both merely erogatory *and* morally decisive with respect to ϕ , it would appear that ϕ is certainly not morally meritorious; one does not behave in a particularly morally good way if one behaves in a way for which there is less moral reason than the act that is merely erogatory. For the same reason, it cannot be the case that ϕ is supported by reasons that are equivalent in strength to ψ . This would entail that ϕ and ψ are of equivalent moral importance or quality. But this is incompatible with *Morally Good*. Supererogatory actions go *beyond* one’s duty. Hence it must be the case that supererogatory actions are morally decisive with respect to merely erogatory actions.

But this just is the puzzle. How could it be morally permissible to fail to perform a supererogatory act ϕ , but instead only to perform merely

erogatory act ψ , given that ϕ is morally decisive with respect to ψ ? How could performing an act that is morally worse than an open alternative be compatible with fulfilling one's moral obligations? In general, it seems wrong to say that we are morally permitted to act in ways that are morally defeated by other open alternatives. To put this point in a slightly different way, call a theory of the method by which moral reasons determine moral requirements a theory of "moral rationality". At first glance, it seems plausible to say that moral rationality is *optimizing*: if ϕ is morally decisive with respect to ψ , one is morally required to ϕ rather than ψ .

Of course, an approach to moral rationality is only as plausible as its verdicts. In this case, an optimizing view appears to leave us unable to say that supererogatory actions exist. So much, it would seem, for the optimizing view. But the traditional view must offer an alternative account of moral rationality that would plausibly allow for the existence of supererogatory actions. I consider three here; none, I argue, is successful.

3. Moral Satisficing

According to Dreier, "Ethical satisficing is structurally similar to rational satisficing. Ethical satisficing theory says that it can be morally right to choose an alternative that is good enough... even though there is a better alternative available."¹⁰ For a moral satisficing rationality, one is permitted to perform an act that is morally good enough; one needn't perform the morally optimal act. On a scheme of moral satisficing, one violates a moral obligation only if one acts in a way that is not morally good enough.

Moral satisficing rationality accommodates *Permissible not Required* and *Morally Good*. If we accept that one is morally required only to perform an action that is morally good *enough*, this can leave open the possibility of morally *better* actions that are not morally required. Hence it would be permissible, not required, to perform a supererogatory action despite the fact that this action is morally decisive with respect to a morally permitted act, ψ . Furthermore, because ϕ -ing is morally better than ψ -ing, and given that ψ -ing is morally *good enough*, moral satisficing can explain why ϕ -ing is plausibly morally heroic or "beyond" the call of one's duty.

Though moral satisficing has much to say for itself, at least when it comes to the supererogatory, it cannot accommodate the traditional view. To embrace moral satisficing, one must reject *But for Sacrifice*. To see this, consider the following. Assume that "morally good enough" is, say, 70% of

¹⁰Dreier, 142.

morally optimal. If we assume that ψ -ing is 70% of morally optimal, any act ϕ that is morally decisive with respect to ψ will be supererogatory. So far so good. But the traditional view requires a further claim: that *were* a supererogatory action (one falling within S , anyway) not of substantial sacrifice to the agent, this action would be morally required. But a satisficer cannot accept this; assuming that ψ is “morally good enough”, satisficing permits ψ -ing whether or not ϕ -ing involves any sacrifice at all, assuming, *ex hypothesi*, that it is at least 70% of morally optimal.

Of course, one might deny *But for Sacrifice* and claim that one needn't be required to behave morally optimally when doing so is of no sacrifice. But this claim is clearly implausible. Saving one's comrades when it is of no sacrifice is surely required; but is supererogatory when one must sacrifice one's own life to do so. Take another case. Assume that I am presented with a range of buttons: the first would be to perform an action that is 1% of morally optimal, the second, 2% of optimal, and so on. Pressing none of these buttons involves any sacrifice on my part. In this case, I am clearly required to press the morally optimal button. To behave in a way that is morally sub-optimal when doing so costs one nothing at all seems paradigmatically impermissible (at least in this case). Moral satisficing cannot make this claim.

4. *Justice and Beneficence: Dreier's Sophisticated Satisficing*

This simple form of satisficing cannot accommodate the traditional view. Dreier, however, advocates a form of moral rationality that may be more promising. Dreier refers to this account as a form of moral satisficing, though his view is somewhat more sophisticated than its previous incarnation. In considering the purportedly supererogatory action of helping to return a stranger's hat that has blown away in the wind, Dreier argues that there are, in fact, two moral points of view:

To borrow from virtue theory, one point of view we can adopt is the point of view of the perfectly virtuous agent...or, less ambitiously, just the beneficent agent. From this perspective, there is everything to be said in favor of fetching the stranger's hat and nothing to be said against it. Failing to fetch the hat is falling short of perfection (in this dimension) and not permissible at all. But we can also adopt a less ambitious perspective—that of the just person (maybe 'dutiful' would be a better word). From the point of view of justice, there isn't anything to be said

in favor of going to all that trouble to get a stranger's hat. If I do go to all the trouble, that doesn't make me more just. The suggestion is that judgments of wrongness are made from the point of view of this less demanding virtue, whereas judgments of what would be better or worse are made from the more ambitious point of view.¹¹

Dreier's account is interesting, and much more can and should be said about it than I can or will say here. For Dreier's view, moral rationality is a mix of two different moral points of view. The point of view of "justice" determines the moral permissibility and impermissibility of actions. For Dreier, however, there is no reason of justice to fetch the stranger's hat. If so, refraining from so doing is permissible, not required. The second point of view, the point of view of "beneficence", determines the relative goodness or badness of actions, but does not say anything about the permissibility or impermissibility of these actions. Reasons of beneficence rank-order actions, but that ϕ -ing is supported by greater reasons of beneficence can never count in favor of a requirement to ϕ . Putting these points of view together, one performs a supererogatory action if one performs an action that is better, i.e., decisive from the point of view of beneficence with respect to, the merely erogatory action (which one might interpret as the morally worst action that is compatible with the requirements of justice).

Dreier's view also fails to accommodate *But for Sacrifice*. Take Urmson's case. Dreier, to successfully claim that saving one's comrades at a sacrifice of one's own life is supererogatory, must say that, from the point of view of justice, "there isn't anything to be said in favor" of saving one's friends. Of course, there is something to be said in favor of rescuing one's comrades from the point of view of beneficence, and this is what renders one's action supererogatory. But Dreier's view cannot accept the further claim that were there no sacrifice involved, one is morally required to save one's comrades. If "there isn't anything to be said in favor" of so doing from the point of view of justice, and justice is the point of view that determines "judgments of wrongness", it seems difficult to see how Dreier might deliver the verdict that it is wrong not to save one's comrades under any circumstances at all. It would seem that any reason to save one's friends is a reason of *beneficence*, not justice, and hence this reason cannot influence the extent to which it would be wrong not to save one's friends when it is of no sacrifice.¹²

¹¹Dreier, 149.

¹²Incidentally, this is the chief difficulty for Horgan and Timmons' suggestion as well. Horgan and Timmons claim that the bulk of the supererogatory is explained by the fact

Of course, Dreier claims that the domain of justice is the domain of “right and wrong”. Hence one might be tempted to suggest that one’s requirement to save one’s comrades if it is of no sacrifice is a requirement of the point of view of justice. Doing so in this case *does* “make me more just.” But one cannot make this claim and maintain the integrity of Dreier’s proposal. To allow Dreier to claim that e.g., saving one’s comrades by sacrificing one’s life is supererogatory, he must say that the reason to save one’s comrades is a reason of beneficence: it is something that makes the act “better”, but not morally *required*. But to say that one is required to save one’s comrades when there is no sacrifice involved would seem to collapse the distinction between reasons of beneficence and reasons of justice. Such a position seems to require reasons of betterness to play a role in determining one’s moral obligations, after all. But if this is the case, it is difficult to see how such reasons wouldn’t play a role in the case in which saving one’s comrades *would* entail sacrifice. If so, we must now search for an alternative form of “justice rationality” that would allow one to perform a sub-optimal act from the point of view of justice when sacrifice is involved. But this just *is* the puzzle of the traditional view.

One might press the following response to my dismissal of Dreier’s view.¹³ We often say that “giving someone her due” is what is morally required; but that going beyond the obligation to give someone her “due” is supererogatory. And if this is the case, thinking about the supererogatory in terms of the two moral domains Dreier suggests is intuitive. After all, if we identify giving someone her “due” with the requirements of justice, wouldn’t we say that, no matter the cost involved, going beyond what is required by the domain of justice is supererogatory? I think not. If we wish to say that saving one’s comrades by throwing oneself on a grenade is supererogatory on this view, we must say that this action goes beyond giving these individuals their “due”—it is not required, strictly speaking, by the domain of justice. But if that is correct, we seem unable to say that, were it of no cost the soldier would have been required to save them (as he surely would have been). If, no matter what the cost, going beyond giving someone her

that an act ϕ is supported by moral *favoring* reasons, but not moral *requiring* reasons. However, the extent to which a particular action is supported by favoring rather than requiring reasons cannot be determined by the extent to which this action requires sacrifice. If so, morality must merely favor, e.g., my saving of my comrades, whether or not it is of substantial sacrifice. But this, it seems to me, is entirely mistaken. Horgan and Timmons, to their credit, admit that in certain cases a further account is required to cover cases falling within *S*. In particular, they seem to accept Portmore’s view, discussed in the following section. See Horgan and Timmons, 59-62.

¹³Thanks to an anonymous reviewer.

“due” is supererogatory, it is supererogatory to save these individuals even if there is no cost. Of course, one might be tempted to claim that justice demands saving someone from death when it is of no cost. But this would be to destroy the integrity of Dreier’s approach as explored in the previous paragraph.

The following point is critical. It could very well be that some supererogatory action would be supererogatory even if no sacrifice were involved. However, even if there is *one action* that falls within *S*, one and only one action of which can rightly be said that without sacrifice it would be required, the views so far considered fail.

5. *Non-Moral Moral Justification*

The failure of both versions of a moral satisficing rationality can be traced to the same source: neither form could allow that a lack of agential sacrifice would, of itself, render a previously supererogatory action morally required. To accommodate *But for Sacrifice*, however, one must allow that agential sacrifice (or lack thereof) can play this moral role.

Some forms of moral rationality allow agential sacrifice to play this role. For instance, one might claim that under conditions of non-trivial sacrifice, one has a certain degree of “moral permission” not to conform to what otherwise would be moral demands to assist others. One such possibility is explored by Joseph Raz. Raz holds that the supererogatory is the outcome of a set of reasons known as “exclusionary permissions”. Raz explains this idea as follows: “The permission to refrain from performing an act [of] supererogation is an exclusionary permission, a permission not to act on certain reasons. An act is a supererogatory act only if it is an act which one ought to do on the balance of reasons and yet one is permitted not to act on the balance of reasons.”¹⁴

For Raz, an exclusionary permission is a certain sort of reason (a “second-order” reason) that justifies one in ignoring certain other reasons (“first-order” reasons). Surely there are strong first-order moral reasons to save one’s comrades by jumping on a grenade. But the soldier is permitted, given his exclusionary permission, to exclude these reasons in his moral deliberation. Furthermore, this view may be designed to accommodate *But for Sacrifice*. For supererogatory actions falling within *S*, one might say that the exclusionary permission is *triggered* by the sacrifice involved; no

¹⁴Joseph Raz, *Practical Reason and Norms* (Oxford: Oxford University Press, 2nd ed., 1990), 94.

sacrifice, no permission.

This is an interesting proposal, and one worth consideration. But an immediate question arises. If agential sacrifice is to trigger an exclusionary permission, this sacrifice itself must have at least some form of normative significance. But what sort of normative significance might it have? One might say that the exclusionary permission is triggered by such sacrifice because this sacrifice, of itself, has *moral significance*, and hence allows one to ignore moral reasons that tell in favor of committing the act in question. But such a proposal seems not to capture *Moral Goodness*. After all, in this case the merely erogatory act is declared permissible by something that is itself morally significant: the sacrifice to the agent. But if there are specifically moral reasons with significance enough to justify a merely erogatory action ψ , it is hard to see how ϕ -ing could be morally *special*. After all, morality seems to care about ψ -ing, sufficient to render ψ -ing morally justified.¹⁵

However, we might interpret Raz's view differently. Why not say that these exclusionary permissions, and the moral justification they entail, come about given the *non-moral* significance of sacrifice to the agent? This view would, in essence, suggest that one has moral justification to perform a merely erogatory act ψ rather than a supererogatory act ϕ because certain non-moral reasons, e.g., reasons not to sacrifice one's self in certain ways, generate exclusionary permissions: permission not to conform to the balance of *moral* reasons. In this way, non-moral reasons are relevant in determining which actions are genuinely morally required.

This very position is developed in detail by Douglas Portmore, and his analysis deserves attention here. Portmore exploits an important ambiguity in the way in which reasons may be relevant from the moral point of view. From the point of view of any domain, reasons may have *requiring* strength, i.e., may count in favor of a requirement to ϕ , or *justifying* strength, i.e., may count in favor of the justification of ϕ -ing from this point of view. Non-moral reasons, though they do not possess moral requiring strength, *can* possess moral *justifying* strength. For Portmore, though non-moral reasons (including prudential reasons) have no moral requiring strength, they have moral justifying strength; they can count in favor of moral justification to ψ rather than ϕ , despite the fact that ϕ is morally decisive (i.e., supported by stronger moral reasons) with respect to ψ .¹⁶ Non-moral reasons (including prudential reasons) are all-things-considered morally justifying when the,

¹⁵See, for instance, B. C. Postow, "Supererogation Again" in *The Journal of Value Inquiry* 39 (2005), 248.

¹⁶Portmore, 372, 375n12.

say, prudential reason not to ϕ is of *sufficient* weight compared to the moral reason to ϕ . This proposal could—though needn’t—be expressed in terms Raz proposes. One might say that non-moral reasons have moral justifying strength insofar as they trigger certain second-order moral reasons, viz., exclusionary permissions, which justify morally sub-optimal acts. Henceforth, this is “Portmore’s view”.

Portmore’s view can accommodate the traditional view in a rather ingenious way. It can accommodate *Permissible not Required*, because the reasons that tell in favor of merely erogatory actions are of merely moral justifying, not moral requiring, strength. Were they of the latter, supererogatory actions would be morally impermissible. In addition, it can accommodate *Moral Goodness*. The moral justification allotted to a merely erogatory act is a result of non-moral reasons; this allows us to say that a supererogatory act is *morally* superior (since it is morally decisive), though not required (given that the merely erogatory action, though morally sub-optimal, is morally justified as a result of non-moral reasons). Further, it can accommodate *But for Sacrifice*. In cases of supererogatory acts falling within S , were it the case that there were no sacrifice involved, there would be no non-moral reasons of moral justifying strength that tell in favor of the merely erogatory act.

The traditional view is thus accommodated. Indeed, Portmore’s approach is required to accommodate the traditional view.¹⁷ This may seem surprising, but permits of a relatively straightforward argument. For any supererogatory action ϕ , falling within S , this act must be supported by reasons that would be sufficient to require ϕ , were it not for the sacrifice to the agent involved (by *But for Sacrifice*). But (by *Permissible not Required*) ϕ cannot be required as it stands. Hence the sacrifice *itself* must have some sort of force in determining our moral requirements. But this force is either the product of moral reasons or non-moral reasons. If this force is entirely the product of moral reasons, *Moral Goodness* fails: any moral requirement to ϕ is outweighed by the moral significance of ψ -ing, and hence we are unable to capture the moral “specialness” of ϕ -ing. Hence this force must, at least in part, be the product of non-moral reasons. But the only way to allow that non-moral force can determine moral requirements is to declare that non-moral reasons have moral justificatory (though, given *Permissible not Required*, not moral requiring) force. And this just *is* Portmore’s view.¹⁸

¹⁷See also Portmore, 378-381.

¹⁸One might critique Portmore’s view by claiming that it is conceptually confused. After all, if a reason is relevant from the moral point of view, doesn’t that make it a moral reason? In this vein, Shelly Kagan claims that the moral point of view should be

6. *The Traditional View and the Bounds of Moral Justification*

Because Portmore's view is required to accommodate all three tenets of the traditional view, any problem for Portmore's view is a problem for the traditional view. In this section, I argue that, given straightforward assumptions, Portmore's view implausibly expands the bounds of moral justification.

The best way to see this is to compare two cases. The first is:

Gus: Gus finds himself the recipient of an inheritance from a wealthy relative. This inheritance will allow Gus to buy a new car, which Gus desires to do, and which will allow Gus to see much more of his significant other, who lives in a distant town. Alternatively, Gus could donate his inheritance to Oxfam International, which would save ten people from death.

On reflection, it is plausible to say that Gus's donation would be supererogatory. Not only this, it seems plausible to say that were there no cost involved in saving ten people from death, Gus would surely be morally required to do so. Hence we should accept:

1. Gus's potential donation would be supererogatory, falling within *S*.

So far so good. But before I introduce the second case, let's shift gears slightly. There is surely some reasonably serious harm to which we may morally permissibly subject an individual for the purposes of saving some number of people (perhaps large) from death. Moral theories will differ on the *seriousness of harm-to-number of people saved* ratio; act-consequentialism, for instance, will claim that one person harmed will justify one person saved, for any harm short of death. Others will hold that there

concerned with moral reasons only, in a way that is simply a conceptual truth: "since we are concerned with what is required by *morality*, the relevant reasons—whether decisive or not—must be moral ones," (Shelly Kagan, *The Limits of Morality* (Oxford: Oxford University Press, 1989), 66). If only moral reasons are relevant to determining our moral obligations, Portmore's view is ruled out.

Though this objection is tempting, I think Portmore's view survives. Portmore claims that some reasons have moral requiring strength, others have moral justifying strength. But there is nothing conceptually confused in claiming that non-moral reasons can be among those reasons that possess moral justifying strength. This view would be conceptually confused only if we claimed that non-moral reasons help to determine which actions are *morally optimal*. But Portmore's view does not do this. Thanks to an anonymous reviewer for suggesting this objection.

are agent-centered restrictions against harm, and that any reason to harm can only be outweighed by a larger amount of good, a larger number of individuals saved. For any plausible moral theory, however, there is some number of people (n) we could save from death that would morally justify, say, torturing one person, even leaving other things equal (e.g., that the agent is not among those who would be saved). If so, the moral reason to save n from death is of (at least) equivalent moral strength to the moral reason not to torture. Further, with each harm of decreasing seriousness, the moral reason not to harm someone in that less serious way is of equivalent strength to the reason to save fewer individuals. If we are licensed to torture someone to save, say, 1000 people from death, surely we are morally justified in harming someone in a less significant way to save some lesser number of people. With all this in mind, the following seems plausible:

2. Other things (including non-moral reasons) being equal, one is morally justified in beating someone up to save ten from death.

However, if we accept (2), a problematic result arises. Consider now the second case:

Stan: Stan knows that his neighbor has just inherited a substantial amount of money from a wealthy relative. Were Stan to intimidate his neighbor into giving him the money as a result of beating his neighbor up, this would be a prudential benefit to Stan, given that this would allow him to buy a new car, from which he will derive pleasure, and which will allow Stan to see much more of his significant other, who lives in a distant town. (Assume that Stan would avoid punishment.)

If we accept (1) and (2), and Portmore's view, it would seem that Stan is perfectly justified—*morally justified*, mind you—in beating up his neighbor to buy a new car.¹⁹ As the non-moral reasons in Gus's and Stan's cases are of identical strength, they are, given Portmore's view, of equivalent moral justifying strength. But given that the non-moral reasons are identical in both cases, if Gus *is* justified in refusing to save ten, but Stan *isn't* justified in beating up his neighbor, it must be that refusing to beat someone up is morally more important than, and hence morally decisive with respect to, saving ten. But given (2), it must be that refusing to beat someone up is

¹⁹I'm here abstracting from the moral relevance of the "neighbor" relation. This is justified: it is no more plausible to say that Stan is morally justified in beating up a perfect stranger to buy a new car.

not decisive with respect to saving ten. Hence, given that Gus's and Stan's moral justifying reasons are of precisely the same strength, and Stan's moral reason not to harm is no stronger than Gus's moral reason to save, Stan is morally justified in beating up his neighbor to purchase a new car.

Care is required here. The mere fact that Stan could be morally justified in beating up his neighbor to buy a new car shouldn't by itself be regarded as a fatal result of the traditional view. After all, many views—including act-consequentialism, for instance—will hold that such an act can be morally justified, especially if Stan's well-being is furthered by a new car more than his neighbor is burdened by the beating. But the problem for Portmore's view is not this verdict *per se*, but rather the fact that Stan can be morally justified in beating up his neighbor even if his interests are, comparatively, of insignificant moral weight. Here's why. Focus on Gus. For Gus's potential donation to Oxfam to be supererogatory, it must be the case that Gus's refraining from donating to Oxfam is morally sub-optimal. But if Gus's interests in the new car are of greater *moral* weight than the interests of the people he could potentially help, buying a new car is not morally sub-optimal, but rather morally optimal. Hence for Gus, the moral justifying strength of his new car must outstrip the *per se* moral significance of his new car. Given the weight of the reasons involved, the same must be said of Stan: Stan is morally justified, on Portmore's view, *even if* Stan's improvement in welfare is less morally significant than the harm inflicted on his neighbor. *This* verdict is simply incredible, and must be rejected. If Stan is justified, this justification cannot be *moral* justification.

My critique here is complex; for clarity I'll state it in semi-formal terms. Assuming Portmore's accommodation of the traditional view, and taking (1) and (2) as previously stated:

3. Non-moral reasons, including prudential reasons, have moral justifying strength. (Portmore's view.)
4. Stan is not morally justified in beating up his neighbor to purchase a new car. (Assumption for *reductio*)
5. The moral justifying strength of Gus's prudential reason to purchase a new car is sufficient to outweigh Gus's moral reason to save ten from death. (1 and 3)
6. The moral justifying strength of Stan's prudential reason to purchase a new car is not sufficient to outweigh Stan's moral reason not to beat up his neighbor. (3 and 4)
7. The reason for Stan to purchase a new car, and the reason for Gus to purchase a new car, are of equal prudential strength.

(Assumption)

8. The reason for Stan to purchase a new car and the reason for Gus to purchase a new car are of equal moral justifying strength. (3 and 7)

9. The moral reason not to beat someone up is decisive with respect to the moral reason to save ten from death. (5, 6, and 8)

10. Other things (including non-moral reasons) being equal, one is not morally justified in beating someone up to rescue ten from death. (9)

11. Hence, one is morally justified in beating up one's neighbor to purchase a new car. (*Reductio*, 10 and 2)

Though one might critique the reasoning in this argument, it is, on its face, straightforward and compelling.²⁰ Given (1) and (3), (5) follows: because Gus's potential donation falls within *S*, there is a reason of moral requiring strength in favor of it, but Gus's non-moral reasons are of sufficient strength to justify not donating; this is precisely how Portmore's view accommodates the existence of supererogatory actions that fall within *S*. Given (3) and (4), we must hold the opposite in Stan's case: even though Stan has a moral justifying reason to beat up his neighbor, this reason not sufficient to morally justify beating up his neighbor, hence (6). (8) follows from the fact that the purchase of a new car, for both Stan and Gus, are of equal prudential

²⁰The most obvious targets are (9) and (10). (Thanks to an anonymous reviewer.) One might accept (5), (6), and (8), and deny (9) by holding, for instance, that the reason not to beat someone up has "extra" moral weight only against competitor *non-moral* reasons, rather than competitor *moral* reasons. This is a perfectly coherent view, but to deny (9) and accept (5), (6), and (8) would imply an unsavory intransitivity of moral comparatives (such as "better than", "ought to do rather than", etc.). Here the predicate "morally justified rather than" is intransitive: one is morally justified in saving ten rather than refusing to beat someone up ($\neg(9)$); one is morally justified in acting on one's non-moral reasons of strength *s* rather than saving ten ((5)), but one is not morally justified in acting on one's non-moral reasons of strength *s* rather than refusing to beat someone up ((6) and (8)). This spells trouble: the transitivity of moral comparatives seems to me essential bedrock in moral deliberation. Generally speaking, I'm with John Broome: "[a] comparative relation is necessarily transitive," (John Broome, *Weighing Lives* (Oxford: Oxford University Press, 2004), 50). Though some have denied the transitivity of moral comparatives (most commonly, "better-than"), as a defense of the traditional view, this strikes me as a scorched-earth strategy. Second, one might deny that (10) follows from (9), i.e., that the reason not to beat someone up is morally decisive with respect to the reason to save ten entails that one is not morally justified in beating someone up to rescue ten from death. However, to justify saving ten under such conditions, one must be justified in performing morally sub-optimal actions, *even if no sacrifice is at stake*. But this form of satisficing was rejected, for good reason, in §3.

significance; if so, their moral justifying strength must be identical.²¹ (9) also follows: if Stan's and Gus's non-moral reasons are of equivalent moral justifying strength, then for Stan to be *unjustified* in conforming to his non-moral reason, and for Gus to be *justified* in conforming to his non-moral reason, it must be that refusing to beat up one's neighbor is of greater *moral* significance than saving ten. Hence refraining from beating someone up is morally decisive with respect to saving ten. This implies (10), i.e., that—non-moral reasons being held equal—one is not morally justified in beating someone up to save ten. But (10) and (2) are inconsistent. Hence, *reductio*.²²

The reasoning on display is difficult to reject. The next step is to consider the premises. One might, for instance, reject (1) rather than (4). But (1) seems obvious; indeed, Gus's potential donation seems a paradigmatic instance of the supererogatory; furthermore, if we can save such individuals from death at no cost to ourselves and fail to do so, those whom we failed to save surely could complain on moral grounds. (3) is essential to the traditional view, as I have already had occasion to argue. (7) is a legitimate assumption for the purposes of argument. This leaves (2). One might complain that I have no genuine reason to believe (2); after all, that we are morally justified in beating up one to save ten, rather than, say, eleven,

²¹One might suggest that there could be additional *per se moral* significance for Gus to buy his new car, and hence, despite the fact that they are prudentially identical, they are of different moral justifying strength. But this can simply be assumed away; assume that there are no present factors that would render Gus's welfare morally more important than Stan's welfare.

²²An anonymous reviewer suggests the following: consider a quasi-Kantian approach of the following structure: Gus needn't donate because Gus isn't required to treat himself as a means to securing overall benefit, whereas Stan is positively required not to treat his neighbor as a means to securing overall benefit. Couldn't this proposal accept (1) and (2), but deny (3)? I think not. To accept (2), one must say that the moral reason not to beat someone up—*however this reason is understood*—can be balanced against the reason to save ten from death, *however this reason is understood*. In other words, one must say that beating someone up is not morally decisive with respect to saving ten. But once we accept the traditional view, we must hold that Gus's prudential reason, which is equal in terms of moral justifying strength with Stan's prudential reason, can justify failing to save ten. But we can only say (*modulo* note 20) that Stan is *not* morally justified in proceeding—given that Gus and Stan's non-moral reasons are of equivalent moral justifying strength—if refusing to beat someone up *is* morally decisive with respect to saving ten. But, given (2), it isn't. The *construction* of the moral reasons involved (whether they are Kantian, welfarist, agent-relative, agent-neutral, etc.) makes no difference. A Kantian approach might claim that one is never morally justified in treating someone else as a (mere) means. But this approach must deny (2), which, as I argue here, is no more plausible than accepting (11).

seems arbitrary. But the specific number identified by (2) is neither here nor there; the number can change without changing the seriousness of the problem for the traditional view. Imagine that Gus could save 100 people from death as a result of his Oxfam donation. Many would say that even under these conditions, Gus's donation remains supererogatory; his prudential reason is strong enough to morally justify his refusal to save 100 from death. But to deny that one could be morally justified in beating someone up to save 100 people from death is positively absurd. Hence Stan's prudential reason, which is identical to Gus's, must also have the power to morally justify beating up his neighbor, insofar as refraining from beating up one's neighbor is not morally decisive with respect to saving 100. One might put this in general terms: there is some number (" n ") of individuals Gus could save from death at the cost of his new car and new life with his significant other, the saving of which would no longer be supererogatory. For the traditional view to survive, we must say that we are unjustified in beating someone up to save *any* sub- n number of people from death. But this claim is worse than dubious.

Portmore explores one way to deny (2).²³ According to Portmore, Gus's reason to donate is "imperfect". Imperfect reasons "do not support performing any specific alternative, but instead support performing any of the alternatives that would each constitute an equally effective means of achieving the same worthy end."²⁴ Perfect reasons—like Stan's reason not to harm his neighbor—are reasons to perform a particular act at a particular time. According to Portmore, Gus's imperfect reason of beneficence is—roughly speaking—to be sufficiently beneficent over the course of a life, and hence the reason to be beneficent at any particular time (i.e., to donate now) is quite weak. Hence it might be that any reason Gus has to donate here and now is *non-fungible* with Stan's reason not to harm his neighbor. Though this response is available, it is no better than (11). First, this response would require us to deny not only (2), but that Gus's donation falls within S . If Gus would be required to donate *now* on the assumption that so donating is costless, this can only be the result of a *perfect* reason to do so. But that this donation falls within S , as we have seen, is very difficult to deny. If that's correct, there must be a perfect, rather than imperfect, reason for Gus to donate. (2) survives.

Portmore's account is a truly remarkable method by which to accommo-

²³Portmore has reported this proposal to me in personal communication.

²⁴Portmore, *Commonsense Consequentialism* (Oxford: Oxford University Press, forthcoming), ch. 6. See also Horgan and Timmons, 35-36.

date the traditional view. I have argued, however, that given straightforward assumptions, and straightforward reasoning, Portmore's view implausibly expands the bounds of moral justification. This is not to say, of course, that there is no way to avoid this conclusion. One might deny one or more of the assumptions or judgments that force the traditional view to accept (11). However, these judgments seem to me robust enough to warrant the search for an alternative to the traditional view. In what follows, I offer just such an alternative that is (a) satisfying in itself and (b) keeps the boundaries of moral justification right where we want them.

7. *The Supererogatory...*

To begin, consider again the notion of a moral reason. Moral reasons are reasons to act in a certain way from a specific point of view, i.e., the moral point of view. But moral reasons are not the only sorts of reasons to act in certain ways rather than others. There are reasons of etiquette, professional reasons, legal reasons, prudential reasons, etc., that are all relevant to their specified domain.

There's a funny thing about all these reasons, however. That there is a reason to ϕ from the perspective of, say, etiquette, professional norms, legal norms, etc., does not necessarily entail that *we* have any reason to ϕ . That Anton LaVey's *Satanic Bible* commands one to ϕ is surely a reason to ϕ from the *Satanist's* point of view. But the Satanist's point of view, of itself, generates no real reasons for us to conform to its demands. In this, the Satanist's point of view seems to differ from other points of view that are widely recognized to impact what we ought to do. Prudence, morality, associative norms, etiquette, all seem to generate reasons that not only determine what is required from that point of view, but that also have an impact on what we ought, actually, to do.

These reflections seem to indicate that there is a point of view separate from, e.g., morality, prudence, religion, etc., that determines whether, and to what extent, we should conform to the reasons or requirements offered by any of these individual domains. For the remainder of the paper, I call this the point of view of practical reasons, the requirements of this point of view *rational* requirements. I have a reason to conform, e.g., to the demands of prudence, morality, or any other domain only if the requirements of those domains generate *practical* reasons for me to act. Given the verdicts in the previous paragraph, moral reasons, prudential reasons, reasons of etiquette, etc., all seem to generate practical reasons. Anton LaVey-reasons do not. Furthermore, the practical importance of these domains is unequal. Though

I may have a practical reason to ϕ if ϕ -ing is *aesthetically* required, the fact that $\neg\phi$ -ing is *morally* required seems to indicate that I am *rationally* required not to ϕ . One has a stronger practical reason to conform to the requirements of morality than the requirements of aesthetics, or so it would seem.

I take this conceptual territory to be familiar. Furthermore, I take it that a standard assumption is that morality is the most important practical domain. In particular, a standard analysis of the relationship between practical and moral rationality holds that the requirements of morality are rationally *supreme*: if one is morally required to ϕ , one is thereby *rationally* required to ϕ . On this view, one has overriding practical reason to conform to the requirements of morality. But this view is substantive, and can be denied. For instance, whether one has overriding practical reason to conform to legal requirements is certainly a case-by-case matter.²⁵ So might it be with morality.²⁶

Assume, for the purposes of argument, that moral requirements do not always yield rational requirements. Now take the following case.

Rose: Rose is a retiree with a substantial pension, and lives comfortably. Rose could get by with less, but this would require her to give up some things she enjoys doing. Assume now that Rose is morally required to assist others with her resources rather than spending her resources on herself. Assume also that Rose is not rationally required to assist those whom she could assist. Imagine now that Rose dedicates substantial time and money to a local family, themselves down on their luck. Rose provides for their food, lodging, and child care which is burdensome, and leaves her unable to live the life she would otherwise want to.

How would we describe Rose's action? *Ex hypothesi*, Rose behaves in accordance with her moral obligations. But also, *ex hypothesi*, Rose is not rationally required to do so. One might, of course, simply describe Rose's action as an instance of action that conforms to a moral requirement. But that doesn't seem to say it all; Rose's action is not *just* morally required, Rose's action is morally required in a way that isn't *required of her*. This point is important to stress. Though it is morally required, it is "beyond the call of duty".

²⁵See Roger Crisp, *Reasons and the Good* (Oxford: Oxford University Press, 2007), ch. 1.

²⁶See, for instance, Phillipa Foot, "Morality as a System of Hypothetical Imperatives" in *Virtues and Vices* (Oxford: Oxford University Press, 1978).

The point of this story should by now be obvious. Rather than holding that supererogatory actions are those that are morally good but are not morally required, we should instead say—roughly—that supererogatory actions are those that are morally good, but that are not *rationally* required.

Insofar as my view rejects the traditional view, it offers a reinterpretation of the three focal points of the traditional view. My view can accept *Morally Good* as stated. Supererogatory action, on my view, is morally better than that which rationality commands. *Permissible not Required*, on the other hand, becomes

Permissible not Required II: If an act ϕ is supererogatory, ϕ is rationally permissible, but is not rationally required.

In addition, *But for Sacrifice* becomes

But for Sacrifice II: A subset (S) of supererogatory actions would have been rationally required but for the fact that they require non-trivial sacrifice on the part of the agent.

Though my account of the supererogatory, like the traditional view, permits of a number of potential additions and supplementary principles, these theses are the heart of what I call “the anti-rationalist view”.

8. ...and How to Accommodate It²⁷

To accept the existence of supererogatory actions on the anti-rationalist view, one can accept that intramoral rationality is optimizing; one is morally required to ϕ if and only if ϕ -ing is morally overriding. Given the troubles of alternative approaches, this is a point in favor of the anti-rationalist view. However, my view divorces practical rationality from moral rationality. If so, one must adopt a form of *practical* rationality that can accommodate the principles detailed above. Is this possible?

Obviously, to accommodate the existence of supererogatory actions on the anti-rationalist view, we must be anti-rationalists about morality; we must believe that one can be practically justified in refusing to perform morally optimal actions. However, some versions of anti-rationalism cannot accommodate the existence of supererogatory actions, others can. To see this, consider the various ways one might construe the relationship between

²⁷I argue for the claims in this section in much greater detail in [*Author’s Published Paper*].

moral reasons and practical reasons.²⁸ One common thing to say about this relationship runs as follows:

Authority: That ϕ -ing is morally decisive with respect to ψ -ing is a *pro tanto* practical reason to ϕ rather than ψ .

Authority holds that moral reasons are practical reasons in the following sense: that ϕ -ing is supported by the weight of morality over ψ -ing is at least *some* practical reason to ϕ rather than ψ . If *Authority* holds, moral requirements (i.e., actions that are morally overriding) “have rational authority such that it is *pro tanto* irrational to fail to act in accordance with them.”²⁹ But *Authority* is very weak. It says only that in the absence of any countervailing considerations, one ought to behave in accordance with the weight of moral reasons. However, one must obviously accept *Authority* to accommodate the existence of supererogatory actions on the anti-rationalist view. Without doing so, any practical reason whatsoever not to ϕ renders ϕ -ing practically irrational even if ϕ -ing is morally optimal. This would be tantamount to the rejection of *Permissible not Required II*.

However, to accommodate the supererogatory, one must deny a stronger association between moral reasons and practical reasons. One must deny:

Supremacy: If ϕ -ing is morally decisive with respect to ψ -ing, ϕ -ing is practically decisive with respect to ψ -ing.

Supremacy holds that moral reasons are practically overriding: whenever morality holds that one ought to ϕ rather than ψ , it is all-things-considered irrational to ψ rather than ϕ . The anti-rationalist view cannot accept *Supremacy*. To accept the existence of supererogatory actions, it must be the case that one occasionally has permission to behave in a morally sub-optimal way, which *Supremacy* denies.

But the denial of *Supremacy* and the acceptance of *Authority* does not guarantee that supererogatory actions, like Rose’s, will be rationally *permitted*. It must be that morality generates practical reasons of at least sufficient weight to allow that actions that are morally better than those that are rationally required are rationally permitted. Hence, the anti-rationalist view must accept:

²⁸A helpful introduction to the relationship between morality and practical reasons is offered by David Brink in “Kantian Rationalism: Inescapability, Authority, and Supremacy” in *Ethics and Practical Reason*, ed. Cullity and Gaut (Oxford: Oxford University Press, 1997).

²⁹Brink, 255.

Permission: If a particular act ψ is rationally permitted, and ϕ -ing is morally decisive with respect to ψ -ing, ϕ -ing is rationally permitted.

Permission holds that all action that is morally better than is rationally required is rationally permitted. To accommodate the anti-rationalist view, one must accept *Permission* but deny *Supremacy*.

However, it is not obvious how one might coherently accept *Permission*. How could we guarantee that in any circumstance in which behaving in a morally sub-optimal way is rationally permitted, behaving in a morally optimal way is also rationally permitted? After all, it would seem that ϕ -ing and ψ -ing are *both* rationally permitted only when the practical reasons to ϕ are of equivalent strength to the practical reasons to ψ . Assume now that ϕ -ing is morally optimal and ψ -ing is morally sub-optimal, but both are rationally permitted. Now assume that we *add* prudential reason to act in a morally sub-optimal way, such that there is now greater practical reason to ψ . Wouldn't this render ϕ -ing rationally impermissible (contra *Permission*), given that there is now stronger practical reason to ψ than ϕ ?

Though there may be many additional possibilities, one novel way to accept *Permission* would be to adapt Portmore's distinction between moral requiring strength and moral justifying strength. One can accept *Permission* if one accepts that, the prudential reason to ψ can, at best, rationally *justify*, but cannot rationally *require* ψ -ing. In other words, non-moral reasons lack rational requiring strength, but maintain rational justifying strength. Like Portmore's view, one will have rational permission to perform a morally sub-optimal act if the prudential reason to do so is of sufficient comparative weight to the moral reason to avoid the morally sub-optimal act. *Supremacy*, on this view, fails: I can be rationally justified in acting in a morally sub-optimal way (depending, of course, on the weight of the non-moral reasons involved). Nevertheless, because non-moral reasons cannot require a person to behave in a morally sub-optimal way, acting in a morally better way is always permitted, satisfying *Permission*.³⁰

³⁰Some might be attracted to the claim that non-moral reasons can rationally require. If, for instance, morality is indifferent between ϕ and ψ and one has stronger prudential reason to ψ , some might hold that it is irrational (given this prudential reason) to ϕ . Such a verdict, however, would not require a radical revision of my account; one might say, for instance, that though non-moral reasons have rational requiring strength, the rational requiring strength of moral reasons lexically dominates the rational requiring strength of non-moral reasons. This entails that when morality is not indifferent between ϕ and ψ , non-moral reasons cannot require one or the other. But in a case of moral indifference, given that there is equal *moral* rational requiring reason to ϕ rather than ψ , the comparatively

Of course, this yields a complex account of the relationship between practical reasons, moral reasons, and other, non-moral domain-specific reasons. Why isn't this view *ad hoc*? There are two points to be made in this regard. First, my proposal requires no more structural complexity as an account of *practical* rationality than the traditional view requires as an account of *moral* rationality. Indeed, both views accept a roughly "Portmorian" account; mine in the domain of practical reasons, the traditional view in the moral domain. It seems to me if the traditional view is allowed a Portmorian account, so is the anti-rationalist view.

Second, this account passes the most important test of any theory of practical reason: conformity to considered judgment. Take Rose. If we accept this account of practical reasoning, we have the power to accommodate the plausible suggestion that Rose's morally required action is, in fact, supererogatory. Rose has *prudential* reason to avoid assisting this family. This prudential reason is of sufficient rational justifying strength; the prudential requirement to avoid assisting the family is sufficiently strong to justify the act of not doing so. If so, though assisting the family is morally good, indeed, *morally required*, it is not rationally required, and hence Rose is rationally permitted not to assist the family. But given that assisting the family is morally better than is rationally required, Rose has rational permission to assist the family as well, given that her prudential reason not to do so lacks rational requiring strength. Hence, in this case, assisting the family, for Rose, is supererogatory. The same applies in the case of Urmson's soldier. One might say that, given the lives at stake, there is surely greater moral reason for the soldier to throw himself on the grenade than to refrain. But given that this action, quite literally, is an instance of *self-sacrifice*, we should expect that the soldier is rationally permitted, given the rational justifying strength of prudential reasons, not to hurl himself on the grenade. These are the correct answers. Worth noting, however, is that these answers are delivered without adopting a non-standard version of *moral* rationality, and its attendant problems (as illustrated in §§3-6).

If we accept *Permission* and reject *Supremacy*, we can accept that a supererogatory action is morally better than the merely *rationaly erogatory* action (satisfying *Morally Good*). This view can also accept that supererogatory acts are rationally permitted, but not required (satisfying *Permissible not Required II*). And this view can also accept that, were it not for the required sacrifice—that is, the prudential reasons against—supererogatory actions would be rationally required; without prudential sacrifice to ratio-

insignificant non-moral rational requiring reason can tip the balance.

nally justify a morally sub-optimal act, one is rationally required to perform the action for which there is overriding moral reason (satisfying *But for Sacrifice II*). To accept *Permission* but deny *Supremacy* in the way just illustrated is necessary and sufficient to accommodate the existence of supererogatory actions on the anti-rationalist view.

One loose end should be tied. I noted in §1 that the comparative size of *S* relative to the set of all supererogatory actions is controversial. Horgan and Timmons, for instance, hold that at least some supererogatory actions are supported only by “favoring” rather than “requiring” reasons, and hence that one needn’t be required to perform them even if they involve no sacrifice (such as taking a stranger on a fun outing). But my view seems to treat moral reasons as always of rational requiring strength, and if so *S* is coextensive with the set of all supererogatory actions. However, my view is flexible on this point. To capture the possibility of a narrower *S*, one might add a clause to *Authority*, holding that any moral favoring reason (if, in fact, such things exist) is also a *pro tanto* rational favoring reason.³¹ If this is right, one might then amend *Permission* to claim that if a particular act ϕ is either morally decisive *or* morally favored with respect to a rationally permitted act ψ , ϕ -ing is rationally permitted (though merely morally favored acts would not be required, given that they possess no rational requiring strength). This would hold that actions supported by favoring reasons are permissible, not required; they are morally better than is rationally required, and hence they are supererogatory without being within the scope of *S*.

9. Objections

The anti-rationalist view faces objections. I consider three here. First, I argue that though my view accepts the claim that some supererogatory actions will also be morally required, this implication is no slight on my account. Second, I argue that though some might object to the denial of *Supremacy*, one should not consider the ability to accommodate *Supremacy* an advantage of the traditional view. Third, some might argue that my view faces problems—just as Portmore’s view did—with *Stan*. I argue that this is not so.

³¹Furthermore, given note 3, one might say that not only does prudence have rational justifying strength, but, e.g., norms of autonomy, capabilities, etc., also have rational justifying strength of a specified weight. This would expand the realm of supererogatory acts without requiring that morally sub-optimal acts be supported by prudential reasons *in particular*.

9.1. Can Supererogatory Actions be Morally Required?

On my view, it is perfectly possible that an action might fall into the class of supererogatory actions, but might also be morally required. Though this is not true for all supererogatory actions (some will be morally decisive with respect to rationally required actions, but will not be morally overriding), it is true for some, including those acts that are optimal but not rationally required. Dreier objects to this claim as follows:

[The anti-rationalist view] isn't right as it stands. If it were, then our verdict for people who fail to go beyond the call of duty would be that they act wrongly, though it is understandable and rational that they do. Failing to perform would not be morally permissible, because one would have moral reasons for performance and no moral reasons for nonperformance. Choosing not to do the supererogatory act would be like shoplifting a sweater that you really, really want and can't afford. But they are not alike: Stealing the sweater is morally wrong, and failing to [perform the supererogatory act] isn't.³²

This objection is question-begging. I have argued here for the rejection of the traditional view in favor of an alternative analysis. To be told that my view is false because it does not accommodate the traditional view is simply to pound the table against the anti-rationalist view. Let me put this point another way. The *concept* of the supererogatory for which we seek an illuminating analysis is expressed by, e.g., the thought that some actions go "beyond" one's duty. The traditional view offers a conception of this concept: that supererogatory actions are morally good, but not morally required. But the traditional view is not *analytic*; it is a theory of the supererogatory that can (and, I argue, should) be denied. Insofar as I deny it, to be told that my view fails *because* it cannot accommodate the traditional analysis is to beg the question.

However, my view *would* be in trouble under two conditions. First, if it does not capture the concept of the supererogatory, i.e., the sense in which action goes beyond duty. Second, if it offers incorrect *first-order* verdicts, i.e., that a particular action ϕ is supererogatory when, in fact, it is not.

With regard to the first, take Urmson's soldier. Urmson appears to indicate that jumping on a live grenade to save others is morally superior to refraining from so doing, but that it is not a feature of any individual's

³²Dreier, 149.

“duty”. But one has a duty to perform only those actions one is *rationally required* to perform. If conforming to a moral requirement is not rationally required, one has no duty to conform to such a requirement (any more than one has a duty to follow the commands of Anton LaVey). Hence it is perfectly acceptable to say of this case that one has no duty to conform, and also to say that, e.g., we could not “possibly say that they failed in their duty by not trying to be the one who sacrificed himself,” or that “If he had not done so” we would not have “said to him ‘You ought to have thrown yourself on that grenade’”. Furthermore, because he had sufficient rational permission not to jump on the grenade, no “superior” could “have decently ordered him to do it” (given the strength of the prudential reasons against doing it).

Looking at Urmson’s case in this way illustrates a general point. We have a tendency to accept that supererogatory actions cannot be morally required because we have a tendency to accept *Supremacy*: we have a tendency to believe that moral requirements, in all cases, determine our *duty*. Under these conditions, supererogatory acts cannot be morally required. But if we reject *Supremacy*, we are perfectly licensed to say that *sometimes moral requirements go beyond our duty*. Sometimes we have no duty to conform to moral requirements any more than we have a duty to conform to, e.g., legal requirements or requirements of etiquette. If so, there is no pressure to declare that supererogatory actions cannot be morally required.

With regard to the second condition, note that the anti-rationalist view can accept that exactly the same actions that the traditional view declares are supererogatory are, in fact, supererogatory. In fact, the set of actions that are declared supererogatory on both views are identical. For the traditional view, non-moral requirements have moral justifying strength, sufficient to render actions that are morally better than those that are morally justified supererogatory. For the anti-rationalist view, non-moral requirements have *rational* justifying strength, sufficient to render actions that are morally better than those that are rationally justified supererogatory. But there is no reason to believe that the moral justifying strength of prudential reasons is any stronger or weaker in comparison to the moral requiring strength of moral reasons than the rational justifying strength of prudential reasons is in comparison to the rational requiring strength of moral reasons.

How, on any theory, do we determine whether a given action is supererogatory? Presumably, we ask whether that action is among those actions that make up “how we should live”. If ϕ is not part of “how we should live”, but is perhaps morally *better* than we should live, any view that accepts *Supremacy* will interpret this as an instance of action that is morally

better than is morally required, and hence that ϕ is supererogatory. But a view that rejects *Supremacy* will treat “how I should live” not as capturing my moral requirements, but as capturing my *rational* requirements. Hence this view will suggest that ϕ is not necessarily morally better than is morally required, but is morally better than is *rationally* required, and is therefore supererogatory.³³ The anti-rationalist view is sensitive to precisely the same considered judgments as the traditional view, and hence needn’t deviate from the first-order verdicts of the traditional view.

I should consider one brief challenge to this argument. Not all theories of practical reason can accept the argument I have so far offered, even if they accept *Permission*, *Authority*, and the denial of *Supremacy*. Consider, e.g., Sidgwick’s classic *dualism* of practical reason.³⁴ This view holds that there are two fundamental sources of practical reasons: prudence and morality, but that one is always rationally permitted to behave in accordance with the demands of either system of norms.³⁵ This position accepts *Authority*, *Permission*, and the denial of *Supremacy*. But such a view would render my view wildly non-coextensive; because no acts of any moral quality at all are rationally required, any act has even the barest moral worth is supererogatory.³⁶

However, this objection is misguided. Concentrate, for a moment, on the traditional view. For the traditional view, the existence of supererogatory actions is not just an interesting fact about the moral point of view, but is also a reason to reject *act-consequentialism*. Because act-consequentialism delivers problematic verdicts regarding which acts are supererogatory (i.e., none of them), we are licensed to reject act-consequentialism on the basis of such verdicts. Hence, on the traditional view, considered judgments of the supererogatory are relevant to the selection of the true first-order theory of morality or moral rationality (such as the optimizing view). But if we accept the anti-rationalist view, we should treat our considered judgments about the supererogatory as important data in evaluating substantive theories of practical reason, just as—if we accept the traditional view—we should treat our considered judgments about the extent of the supererogatory as important data in evaluating substantive theories of morality or substantive theo-

³³For a more in-depth argument to this effect, see [Author’s Published Paper], op. cit.

³⁴Henry Sidgwick, *The Methods of Ethics* (Indianapolis, IN: Hackett Publishing Company, 7th ed., 1981 [1907]), 507-509. See also Peter Singer, *Practical Ethics* (Oxford: Oxford University Press, 2nd. ed., 1993), 320-2.

³⁵This reading of Sidgwick is controversial. See David Brink, “Sidgwick’s Dualism of Practical Reason” in *Australasian Journal of Philosophy* 66 (1988).

³⁶Thanks to an anonymous reviewer for posing this objection.

ries of the nature of moral rationality (such as, e.g., the optimizing view). If any theory of practical reason offers incorrect or implausible verdicts about the realm of the supererogatory, or any other verdicts about rational justification, we should reject this view. If we accept the anti-rationalist view, our considered judgments about which actions are supererogatory are fair game when it comes to evaluating theories of practical reason, as they have been all along in evaluating theories of morality (and moral rationality) under the traditional view. If so, there is no reason to believe that the true theory of practical reason will offer an implausibly inflated or deflated set of supererogatory actions. If the dualist view is committed to a range of supererogatory actions that stretches considered judgment, we should reject that view. This move is no more problematic than the traditional rejection of act-consequentialism on the basis of the traditional view.

9.2. *Why Deny Supremacy?*

My account of the supererogatory is compatible with the acceptance of *Supremacy*. But for my account to be compatible with a non-zero set of supererogatory actions, *Supremacy* must be false. But that morality is rationally overriding is a traditional approach to the nature of practical reason. Indeed, some hold that investigation of moral questions just *is* the investigation of practical reason itself.³⁷ Though I defend the rejection of *Supremacy* in more detail elsewhere, I offer three reasons here to believe that rejecting *Supremacy* is not as problematic as it may at first seem.³⁸

First, to deny only *Supremacy* is to adopt a rather mild form of anti-rationalism about the moral point of view. Recall that to reject *Supremacy* one need not reject *Permission*, or claim that the non-moral reasons will be decisive against reasons to conform to moral requirements. Rather, all one need believe is that moral agents, at least on occasion, will have rational permission to perform morally sub-optimal acts. My view can accept the claim that, at least on occasion, conforming to morally decisive actions will be rationally decisive, depending on the relative strength of the moral and non-moral reasons involved. For instance, if it is morally required of me not to harm someone else for a mild chuckle, it is perfectly compatible with

³⁷Classically, this view is held by Kant, and is developed in substantial detail by Korsgaard; cf. *The Sources of Normativity* (Cambridge: Cambridge University Press, 1996). For a contemporary restatement, see Sarah Stroud, “Moral Overridingness and Moral Theory” in *Pacific Philosophical Quarterly* 79 (1998); Robert Loudon, “Can We Be Too Moral?” in *Ethics* 98 (1988).

³⁸See [Author’s Published Paper], op. cit.

my view to say that refraining from such harm is rationally required: the prudential reason in favor of harming someone, in this case, is too weak in comparison to the moral reason against to rationally justify such harm.

Second, recall Gus. It is plausible to believe Gus's potential donation is supererogatory. But, as I argue in §6, it is extremely difficult to accept that Gus is justified in refusing to donate and Stan *isn't* justified in beating up his neighbor. But, given this, we must say something about what sort of justification Stan maintains. For my money, it is much more plausible to say that Stan is *rationally* justified rather than *morally* justified in beating up his neighbor. If so, there is independent reason to believe that *Supremacy* fails: doing so is the only plausible way to allow Gus's justification in refusing to donate without saying that Stan is *morally* justified in beating up his neighbor.

Third, and most importantly, the anti-rationalist view interrupts the coincidence between *moral reasons* and *practical reasons* to no greater extent than, and hence should be no more problematic than, the traditional view when combined with *Supremacy*. The traditional view—assuming it is combined with *Supremacy*—suggests that moral reasons are themselves not always practically overriding. After all, non-moral reasons can *morally* justify a morally sub-optimal action and thus render a morally sub-optimal act *morally*, and hence practically, permissible. Hence the coincidence between moral reasons and practical reasons, on my view, is precisely equivalent to the coincidence between moral reasons and practical reasons on the traditional view combined with *Supremacy*. This still leaves a number of paradoxes surrounding the traditional view, which I argue defeat it. But those concerned about the rational authority of morality should not be worried to any *additional* extent about the anti-rationalist view combined with the denial of *Supremacy*. Moral reasons have rational authority to precisely the same extent under the anti-rationalist view and the traditional view conjoined with *Supremacy*. And if this is the case, it seems to me that the case against the anti-rationalist view on grounds of *Supremacy* comes to very little indeed.

9.3. Reconsidering Stan

The traditional view seems to go wrong in holding that Stan could be morally justified in subjecting his neighbor to beatings especially if Stan's welfare is comparatively morally insignificant. But my view might be committed to a conclusion that is, for all intents and purposes, identical. After all, I hold that prudential reasons might be balanced against moral reasons from the

perspective of practical rationality, and it might be the case that in certain circumstances, prudential reasons are sufficient to grant rational permission to perform actions that do not conform to moral requirements. Why, then, haven't I rationally justified Stan's beating of his neighbor? And if I have, haven't I implausibly made his refraining from so doing supererogatory?

This objection is reminiscent of a comment made by Dreier. Recall that Dreier holds that the anti-rationalist proposal would be equivalent to accepting the claim that "choosing not to do the supererogatory act would be like shoplifting a sweater that you really, really want and can't afford." Generally, we think that refraining from shoplifting a sweater—or, in the case of *Stan*, refraining from beating up one's neighbor—is not at all supererogatory.

There are two questions that should be separated here, however. The first concerns whether my view is committed to the claim that Stan's beating of his neighbor for the purpose of purchasing a new car can possibly be *justified*. The second concerns whether my view is committed to the claim that Stan's beating of his neighbor for the purpose of purchasing a new car can possibly be *supererogatory*. Take the first question. As I argue above, if Gus's refusal to donate to Oxfam is justified, we cannot plausibly reject Stan's justification to beat up his neighbor. If we accept that the reasons involved are fungible, and that the moral reason to save ten is at least of equivalent moral importance to the moral reason not to subject a person to beatings, it follows that Stan is justified. But how can I reject the traditional view on the basis of the claim that the traditional view justifies Stan's skullduggery, and endorse the anti-rationalist view, which justifies the very same thing? The answer here, as noted throughout, is that the anti-rationalist view and traditional view take very different justificatory stances toward Stan's dastardly deed. The problem with the traditional view was not that Stan is *justified* in beating up his neighbor, but rather that Stan is *morally* justified in beating up his neighbor. I admit that some will think that both verdicts are implausible. But my approach is comparatively palatable: I needn't, but the traditional view must, claim that Stan's justification is *moral* justification. Assuming that the supererogatory is to be preserved, the anti-rationalist view preserves it in a much more plausible way.

Take now the second question. Must my account hold that, were Stan not to harm his neighbor, his action would be *supererogatory*? Of course, all the same responses apply: if my view is committed to this claim, the traditional view is no less committed. But leave this aside. Both views are licensed to reject the claim that Stan's action is supererogatory. One

could, in principle, restrict the epithet “supererogatory” to only a subset of morally good, but not rationally required, actions by further supplementing one’s analysis of the supererogatory. Some, including David Heyd, hold that an action is supererogatory only if it is “*altruistic* meaning that the act must be conceived as benefiting another person (or persons).”³⁹ If we accept Heyd’s view, Gus’s donation would be supererogatory (given that it is intended to benefit others), but Stan’s failure to beat up his neighbor (given that doing so does not benefit his neighbor) wouldn’t. This would also appear to solve Dreier’s case of the sweater; insofar as refraining from stealing the sweater doesn’t benefit anyone, and is not done for the sake of someone else’s welfare, neither the traditionalist nor the anti-rationalist need claim that it is supererogatory. On whether this further limiting constraint is all-things-considered plausible, I am officially neutral.

One might suggest that admitting that the traditional view can avoid marking Stan’s failure to beat up his neighbor as supererogatory saps any motivation for adopting the anti-rationalist view. But this is not so. The problem with the traditional view is not, nor was it ever, that it provides an extensionally incorrect account of the supererogatory. Rather, it relies on an extensionally incorrect theory of moral justification. The traditional view must accept the claim that Stan’s harming of his neighbor is *morally justified*, whether or not it is officially “supererogatory”. And, as I have so far been at pains to argue, it is *this verdict* that renders any accommodation of the traditional view unacceptable.

10. Conclusion: Act-consequentialism and the Supererogatory

The traditional view is puzzling. To accept the combination of *Permissible not Required*, *Morally Good*, and *But for Sacrifice*, we must reject an optimizing view and instead accept a form of moral rationality that stretches the limits of moral justification. However, we can accommodate the existence of the supererogatory without altering the structure of moral rationality if we assume that supererogatory actions just are those that are morally good, but for which one does not have decisive practical reason. This approach to the supererogatory perfectly captures the category of actions for which the supererogatory was designed: those actions that go beyond the call of one’s duty.

My view is significant for a further reason. As noted in the introduc-

³⁹Heyd, 137. For a contrary view, see Gregory Mellema, *Beyond the Call of Duty: Supererogation, Obligation, and Offence* (Albany, NY: SUNY Press, 1991), ch. 2.

tion, a traditional objection holds that act-consequentialism cannot accommodate the supererogatory. After all, act-consequentialism requires agents to maximize the value of consequences in every case, leaving no room for actions to be “better” than is morally required. Morality, according to act-consequentialism, demands the best. This objection requires the traditional view for its cogency: it must be that someone can go beyond duty from *within* the moral point of view. But if my analysis is correct, this objection fails. The supererogatory is not an *intra*-moral category. The supererogatory exists between the demands of morality and the demands of practical rationality. Hence it is illegitimate to reject act-consequentialism on the grounds that this theory cannot accommodate the supererogatory.⁴⁰ The supererogatory is not a method to evaluate first-order moral theories, but rather a method by which to evaluate theories of all-things-considered practical rationality.

But while we’re on the subject, in addition to not being an objection to act-consequentialism, the existence of the supererogatory is grist for the consequentialist’s mill. If the supererogatory is best understood as a way in which the demands of practical reason separate from the demands of morality, to accommodate the supererogatory we must adopt a first-order moral theory that can plausibly *explain* this cleavage. But, famously, the demands of act-consequentialism do not plausibly match up with that which a given person is rationally required to do.⁴¹ Hence, because act-consequentialism can plausibly explain the distinction between moral requirements and rational requirements, act-consequentialism is well-placed to explain the existence of the supererogatory: sometimes morally required action is simply supererogatory, or beyond the call of that which we ought, rationally, to do.

⁴⁰Consider, for instance, Jason Kawall, “Virtue Theory, Ideal Observers, and the Supererogatory” in *Philosophical Studies* 146 (2009).

⁴¹Paul Hurley, *Beyond Consequentialism* (Oxford University Press, 2009); [*Author’s Published Paper*], op. cit.