

*Moral Reasons, Overridingness, and Supererogation**

DOUGLAS W. PORTMORE

IN THIS PAPER, I present an argument that poses the following dilemma for moral theorists: either (a) reject at least one of three of our most firmly held moral convictions or (b) reject the view that moral reasons are morally overriding, that is, reject the view that moral reasons override non-moral reasons such that even the weakest moral reason trumps the strongest non-moral reason in the determination of an act's moral status (e.g., morally permissible or impermissible). I then argue that we should opt for the second horn of this dilemma, in part because we should be loath to reject such firmly held moral convictions, but also because doing so enables us to dissolve an apparent paradox regarding supererogation. If I am right, if non-moral reasons are relevant to determining whether or not an act is morally permissible, then it would seem that moral theorists have their work cut out for them. Not only will they need to determine what the fundamental right-making and wrong-making features of actions are (i.e., what moral reasons there are), but they will also need to determine what non-moral reasons there are and which of these are relevant to determining an act's moral status. Furthermore, they will need to account for how these two very different kinds of reasons—moral and non-moral—“come together” to determine an act's moral status. I will not attempt to do this work here, but only to argue that the work needs to be done.¹

1. Some clarifications and a quick overview of the argument

To say that one type of reason, R1, overrides another, R2, with respect to a certain kind of normative status, N, is to say that, in any situation where both types of reasons are present and an act, A, has a certain N-status, no modification of the situation that involves affecting only what R2 reasons there are will change A's N-status. That is, if R1 reasons override R2

* Working draft of 5/21/07. You may cite this work, but please do not quote without prior permission. This paper was previously entitled “Are Moral Reasons Morally Overriding?”

¹ I attempt to do some of this work in my paper “Dual-Ranking Act-Consequentialism,” *Philosophical Studies* (forthcoming).

reasons with respect to an act's N-status, then even the weakest R1 reason overrides the strongest R2 reason in the determination of that act's N-status. Note that any thesis to the effect that one type of reason overrides another must be indexed to a certain kind of normative status. And so we must distinguish the thesis that I am denying—that moral reasons are *morally* overriding—from the bolder thesis that moral reasons are *rationally* overriding. On this bolder thesis, moral reasons always override non-moral reasons in the determination of an act's rational status. If this thesis were true, then it would always be *objectively irrational* to refrain from doing what one has best moral reason to do, even if what one has best moral reason to do is only supported by the most trivial of moral reasons and opposed by the weightiest of non-moral reasons.² This thesis is much too strong and implausible to warrant further discussion. By contrast, the question of whether or not moral reasons are *morally* overriding is open to debate.³

My answer to this question is that moral reasons are not morally overriding, and that, therefore, non-moral reasons can affect an act's moral status. But let me stave off one immediate objection. Some will say that, by definition, a moral reason is any reason that is relevant to determining an act's moral status, and thus it is conceptually impossible for a non-moral reason (a reason that is not relevant to determining an act's moral status) to be relevant to determining an act's moral status. I suppose that one could define a 'moral reason' in this way, but I will adopt a different definition so as to preserve an important distinction between reasons that are relevant to determining an act's moral status and reasons that, morally speaking, count for or against performing some action. I will call any reason that is relevant to determining an act's moral status a *morally relevant reason*, and I will call any reason that, morally speaking, counts for or against performing some action a *moral reason*. It may be that not all morally

² An act is *objectively irrational* if and only if the agent has decisive reasons not to perform the act. By contrast, an act is *subjectively irrational* if and only if the agent has beliefs whose truth would give her decisive reasons not to perform the act. These are variants of the definition of 'irrational' that Derek Parfit gives in his *Climbing the Mountain* (Oxford: Oxford University Press, forthcoming)—throughout this paper, I will be referring to the July 31, 2007 version of Parfit's manuscript.

³ A separate question is whether an agent can be morally required to perform an act that she does not have most reason to perform, all things considered. Sarah Stroud argues that the answer is "no," and she calls her thesis the "Overridingness Thesis"—see her "Moral Overridingness and Moral Theory," *Pacific Philosophical Quarterly* 79 (1998): 170-189. The point of this paper is not to discuss whether *morality* is overriding, but, rather, whether *moral reasons* are morally overriding.

relevant reasons are moral reasons, for it is possible that a reason could justify performing an act that would otherwise be morally impermissible without itself counting in favor of it, morally speaking. For instance, we might think that the fact I would personally gain from breaking a promise is not a moral reason to do so in that it does not count in favor of my doing so, *morally speaking*.⁴ Nevertheless, the reason that I have to act for personal gain may be a morally relevant reason, for, perhaps, such a non-moral reason would, if sufficiently weighty, justify my breaking this promise. In any case, it would be a mistake to rule out, by definition, the very real possibility that non-moral reasons (i.e., reasons that, morally speaking, do not count for or against any action) might be relevant in determining an act's moral status. Of course, more still needs to be said on how to distinguish moral reasons from non-moral reasons, but that will come in the following section.

Before I get to the specifics of stating the various definitions and distinctions that I will need for my argument, let me lay out the general train of thought: Where an agent has what is called an *agent-centered option*, we find, typically, all the following to be true: (1) the agent has the moral option of acting self-interestedly, doing what will best promote her own self-interest—call this *self-regarding* act '*s*', (2) the agent also has the moral option of acting altruistically, sacrificing her own self-interest for the sake of doing more to promote the good of others—call this *other-regarding* act '*o*', and (3) the agent has better moral reason to perform *o* than to perform *s*, and, hence, *o* is regarded to be a supererogatory act, whereas *s* is regarded to be a permissible non-supererogatory alternative. Imagine, for instance, that *s* is acting so as to secure a considerable benefit for oneself by pursuing some core life project and that *o* is acting instead so as to secure a far more considerable net benefit for various needy, distant strangers by dedicating one's time, effort, and money to an organization such as Oxfam.

How do we make sense of claims (1)-(3) all being true in such a case? If we were to suppose that moral reasons are morally overriding and, thus, that agents are always morally required to do what they have best moral reason to do, then, contrary to (1), we must hold that the agent is required to perform *o*, for, given (3) and the fact that *o* and *s* are the only options, this is what the balance of moral reasons supports her doing. It seems, though, that given what is at stake for the agent, she is permitted to do *s* instead. Of course, were the self-interested reasons that the agent has for

⁴ Nor is it the case that the fact that I would personally gain by performing the act counts against my performing it, morally speaking.

doing *s* moral reasons, then all they could do is tip the balance of moral reasons from being in favor of performing *o* to being in favor of performing *s*. In that case, though, we could not account for (3)—that is, we could not account for why there is better moral reason to perform *o* than to perform *s*. Furthermore, if we were supposing that agents are always required to do what they have best moral reason to do, then when the alleged moral reasons that favor her performing *s* outweigh the moral reasons that favor her performing *o*, what we end up with is not, as (2) asserts, a moral *option* to perform *s*, but instead a moral *requirement* to perform *s*. Thus, to get a moral option to perform either *s* or *o*, it seems that we must suppose that the reasons that favor the agent's performing *s* are non-moral reasons that have the power to justify her refraining from doing what she has best moral reason to do (viz., *o*). In other words, we will need to suppose that moral reasons are not morally overriding. Otherwise, we will have to give up one of the moral convictions expressed by (1)-(3).

Of course, this was all too fast. To know whether this argument indeed works, we will need to have it spelled out in greater detail and we will need the relevant terms to be carefully defined, tasks to which I now turn.

2. Some analytic truths

Before presenting this dilemma-posing argument in greater detail, I will need to explain my use of certain key terms. By the term 'reasons', I will mean practical reasons, i.e., normative reasons for action. As I see it, reasons for action are considerations that can count for or against performing some action. Such considerations are not always decisive, as countervailing reasons can defeat them, but in the absence of an undefeated countervailing reason, they are decisive and thereby generate an ought of some kind.⁵ Second, I will use the phrase *undefeated reason* such that: if a person, *P*, has an undefeated reason to perform an act, *x*, it follows that *P* does not have better reason to perform any other available alternative.⁶ This allows that there are various ways in which a reason might be defeated. A reason can be defeated because it is outweighed by some other weightier reason, but also because it is trumped, silenced,

⁵ I will be using 'ought' in the weak sense, the sense in which it can be true that an agent ought to perform *x* without it being irrational or impermissible for her to refrain from doing so.

⁶ By 'other available act', I mean to include what might misleadingly be called 'inaction' or what would more accurately be called 'intending not to perform any voluntary bodily movement', which is itself an action.

undermined, excluded, or bracketed off by some other reason.⁷ And note that I have said, ‘does not have better reason’ as opposed to ‘has at least as good a reason’.⁸ The former is broader, allowing for the additional possibility that P’s reason to perform *x* is undefeated because it is incommensurate with P’s reasons to perform some other available alternative such that there is no truth as to whether P’s reason to perform *x* is better than, worse than, or just as good as P’s reason to perform this other alternative. To sum up, then, P’s reason to perform *x* can be undefeated for any of the following three reasons: (i) it defeats (by outweighing, undermining, excluding, etc.) P’s reasons for performing all other available alternatives, (ii) it is tied for best and thus equaled by (or on a par with) P’s reasons for performing whatever other available alternatives are tied for best, or (iii) it is incommensurate with P’s reason for performing some other available alternative such that there is no available alternative that P has better reason to perform.⁹

Lastly, the reader should note that when I refer to, say, the fact that P has *an* undefeated reason to perform *x*, I do not mean to suggest that there is some *single* reason that P has for performing *x* that would remain undefeated even if it stood alone as P’s only reason for performing *x*. Rather, I want to allow for the possibility that P has several reasons for performing *x* that only collectively remain undefeated by other

⁷ For a discussion of some reasons excluding others, see Joseph Raz, *Practical Reasons and Norms*, (Princeton, NJ: Princeton University Press, 1975), pp. 37-9. For a discussion of some reasons bracketing off others as irrelevant, see Thomas Scanlon, *What We Owe to Each Other*, (Cambridge: Harvard University Press, 1998), pp. 50-4.

⁸ I prefer the phrase ‘better reason’ to ‘more reason’, since, as I have already noted, I want to allow for the possibility that P can have better reason to, say, perform *x* than to perform *y*, not because there are more reasons or weightier reasons that favor performing *x*, but because the reasons that favor performing *y* are trumped, silenced, or undermined by the reasons that favor performing *x*. For similar reasons, I will use the locution ‘P has best reason to do *x*’ for the superlative. I will also use the somewhat awkward phrase ‘better moral reason’, which might be stated less awkwardly, but also less concisely, as ‘better reason, morally speaking’.

⁹ As Ruth Chang notes, many philosophers think that “if two items A and B are evaluatively comparable, then A must be better or worse than B, or A and B must be equally good. Call this the ‘Trichotomy Thesis’” — see her “The Possibility of Parity,” *Ethics* 112 (2002): 659-688. Chang rejects the Trichotomy Thesis and argues that, in addition to these three, A and B might be on a par. The difference between being on a par and being equally good is that A and B can be on a par and a small improvement in either A or B would not necessarily make the improved item better than the unimproved item—the improved item might still just be on a par with the unimproved item. By contrast, if A and B are equally good, then a small improvement in one or the other would necessarily make the improved item better than the unimproved item. I take no stand on whether or not the Trichotomy Thesis is true.

countervailing reasons. So when I say that P has *an* undefeated reason (or *a* decisive reason or *a* morally undefeated reason) to perform *x*, this is just a convenient way of saying that the collection of reasons that P has for performing *x* is undefeated (or decisive or morally undefeated).¹⁰

In what follows, I will want to distinguish an undefeated reason from a *morally* undefeated reason. Accordingly, I offer the following analytic truths:

- A1 A person, P, has a *morally undefeated reason* to perform an act, *x*, if and only if P has a moral reason to perform *x* and does not have better moral reason to perform some other available alternative.
- A2 P has an *undefeated reason* to perform *x* if and only if P has a reason to perform *x* and does not have better reason to perform some other available alternative.¹¹

A1 makes use of the notion of a moral reason, so let me define it. Moral reasons are, of course, a proper subset of reasons for action. So if reasons are considerations that count for or against the performance of an action, then moral reasons are considerations that count, morally speaking, for or against the performance of an action. Moral reasons are the only kind of reasons that can give rise to a moral ought, where ‘ought’ is understood broadly to express either obligation or advisableness. In this weak sense of ‘ought’, it can be true to say that P morally ought to perform *x* even if P is not morally required to perform *x*, as where P’s performing *x* is supererogatory. Naturally though, the question arises: “When does a moral reason give rise to a moral ought?” This much is clear: a moral reason gives rise to a moral ought when it is morally decisive, for to say that a reason is morally decisive is just to say that it gives rise to a moral ought.¹²

¹⁰ I follow Shelly Kagan in this practice. See his *The Limits of Morality* (Oxford: Oxford University Press, 1989), 66-67.

¹¹ It might seem that A1 and A2 rule out, by definition, moral and rational satisficing, respectively. That is so only if a defeated reason cannot be a good enough reason. But, as far as A1 and A2 are concerned, a reason can be defeated in that there is a better reason to do something else and still be good enough, such that acting on that reason would be neither immoral nor irrational.

¹² Likewise, to say that a reason is rationally decisive is to say that it gives rise to a rational ought. Thus, nothing I say here rules out the possibility that a moral reason could be morally decisive but not rationally decisive—that, for instance, an agent could be morally required to do one thing, but rationally permitted to do another.

Unfortunately, this, by itself, is not very informative. We need, in addition, some substantive account of when a moral reason is and is not morally decisive—that is, some substantive account of when a moral reason does and does not generate a moral ought. The problem is that there are any number of ways a moral reason might fail to be morally decisive and thus fail to generate a moral ought. The most obvious of these is where it fails to outweigh countervailing reasons. But this is far from the only way. Another way it might fail to be morally decisive is where it is silenced or excluded by other reasons. It also might fail to meet some other necessary condition for moral decisiveness that I have yet to specify. But presumably if it does fail to meet some necessary condition for moral decisiveness (whether it be one that I have or have not specified here), there will be either an undefeated reason or a morally undefeated reason to do something else. That is, unless there is either an undefeated or a morally undefeated reason to do something else, P's moral reason to perform *x* will be morally decisive.¹³

This statement takes into account all the possible ways that a moral reason might fail to be morally decisive. Consider the possibilities. If P's moral reason to perform *x* fails to outweigh opposing moral reasons, then one of those opposing moral reasons will be morally undefeated. If P's moral reason to perform *x* outweighs opposing *moral* reasons but fails to outweigh opposing *non-moral* reasons, then one of those non-moral reasons will be undefeated. Thus I have allowed for the possibility that a necessary condition for moral decisiveness is not only outweighing opposing moral reasons but also outweighing opposing reasons of all kinds. Further consider that if the moral reason in question is excluded or silenced by other reasons, then some other reason that is not excluded or silenced will be undefeated or morally undefeated. The same would hold for any yet-to-be-specified necessary condition for moral decisiveness: if a moral reason fails to meet that necessary condition, then there would have to be some other reason that is either undefeated or morally undefeated.

¹³ This is compatible with a satisficing conception of morality, where it is permissible to perform a good enough act even if there is a better and undefeated reason to perform some other alternative, for the thought here is *not* that unless there is either an undefeated or a morally undefeated reason to do something else, P will be morally required to perform *x*. Rather, the thought is that unless there is either an undefeated or a morally undefeated reason to do something else, P's performing *x* will be *either* morally obligatory *or* morally supererogatory. And satisficers accept that when there is better moral reason to perform one act than to perform another that is merely good enough, performing the morally superior alternative is supererogatory.

Someone might object that I have overlooked the fact that in order for a moral reason to generate a moral *requirement* it is not enough that it simply outweighs opposing reasons; it must do so by a certain minimum. Similarly, someone might object that even if it does outweigh opposing reasons and by the required minimum, it must itself have a certain minimum weight in order to generate a moral requirement—i.e., it must be more than just some trivial moral reason, such as the moral reason that I have to smile at some passerby who would thereby be cheered. But note that I have allowed for these possibilities in that I have allowed that a moral reason can be morally decisive and not generate a moral requirement; the act that is supported by a morally decisive reason may not be morally required, but instead merely supererogatory. So I would admit that a trivial moral reason (even one that is neither outweighed nor undermined by opposing reasons) may not be able to generate a moral requirement even in the absence of any countervailing reasons, but surely, even a trivial moral reason must provide the agent with some impetus for performing the act in question. Otherwise, in what sense is it a genuine consideration that counts morally in favor of performing the act? Thus, even if a moral reason is too trivial to generate a moral requirement when morally decisive, it must at least generate a supererogatory ought when morally decisive. Taking into account all these details, I can now offer the following formal definition:

- A3** P has a *moral reason* to perform *x* if and only if it is the case that, if P has neither an undefeated nor a morally undefeated reason to perform some other available act, P's performing *x* is either morally obligatory or morally supererogatory.¹⁴

¹⁴ Note that this allows for the possibility that a moral reason gives rise to a moral ought whenever there is no morally undefeated reason to do otherwise, for it may be that what one morally ought to do is a function of solely *moral* reasons. Nevertheless, even if this is the case, it will still be true to say that a moral reason gives rise to a moral ought if there is neither an undefeated nor a morally undefeated reason to do something else. For if a moral reason to perform *x* gives rise to *x*'s being morally obligatory/supererogatory if there isn't a morally undefeated reason to do something else, then it will also give rise to *x*'s being morally obligatory/supererogatory if there is neither an undefeated nor a morally undefeated reason to do something else.

Note also that A3 is compatible with particularism, which holds that certain facts can be relevant to how one morally ought to act on one occasion but not on another, and that certain facts can even count in favor of performing a certain type of action on one occasion but against performing that type of action on another occasion. Even on particularism, though, it will still be true to say that some fact that counts in favor of performing a particular act-token *x* on a given occasion either is or is not of the sort that is capable of

A3 is circular in that it explicates the notion of a *moral* reason in terms of its ability to give rise to an act's being *morally* obligatory or *morally* supererogatory. This, however, does not mean that A3 is not both true and illuminating; it is, after all, illuminating to understand the relationship between moral reasons and moral oughts. Of course, given its circularity, A3 cannot tell us whether, say, the reason a person has to promote her own utility is a moral reason or a non-moral reason, but that is the job of a substantive moral theory. Different moral theories hold different views about which reasons are moral and which are non-moral. According to act-utilitarianism, for instance, the reason that a person has to promote her own utility is a moral reason, for absent either an undefeated or a morally undefeated reason to perform some other available act, she would be morally required to do so. If you must decide to watch one TV show or another and which you choose makes no difference to anyone else's utility, then you are morally required to watch the one that would better promote your utility. By contrast, commonsense morality denies this. If watching the sub-optimal TV show only affects you, then your doing so may be stupid and foolish, but not wrong. On commonsense morality, the fact that you would get more utility from watching one TV show rather than another counts neither for nor against your doing so, morally speaking.¹⁵ Hence, the reason you have to watch the optimal TV show (that it will do more to promote your utility) is a non-moral reason on commonsense morality.¹⁶

As A3 implies, not only do facts that give rise to an act's being morally obligatory constitute moral reasons, but so do facts that give rise to an act's being morally supererogatory. So, for instance, a moral theory that

generating a moral ought in the absence of some undefeated or morally undefeated reason to do otherwise. That is, even on particularism, it will be true to say that a particular fact counting in favor of a particular act either does or does not meet the conditions set forth in A3.

¹⁵ For more on this, see my "Position-Relative Consequentialism, Agent-Centered Options, and Supererogation," *Ethics* 113 (2003): 303-32, section 3. See also Michael Slote, *Common-Sense Morality and Consequentialism* (London: Routledge & Kegan Paul, 1985), chapter 1.

¹⁶ A non-moral reason is any reason that it is not a moral reason. But I have to admit that A3 states only what it is to have a moral reason *for* performing an action. To complete the picture, we must say what it is to have a moral reason *against* performing an action. We need, then, to add the following corollary: P has a moral reason not to perform *x* if and only if it is the case that, if P has neither an undefeated nor a morally undefeated reason to perform *x*, then P's refraining from performing *x* is either morally obligatory or morally supererogatory.

inextricably ties moral obligations to rights will still count the reason you have to benefit someone who has no right to your beneficence as a moral reason if the theory holds that your benefiting someone who has no right to your beneficence is supererogatory. Thus there are two kinds of moral reasons, those that give rise to an act's being moral obligatory when morally decisive and those that give rise to an act's being supererogatory when morally decisive. Call the former *requiring reasons* and the latter *non-requiring moral reasons*. So, on the rights-based theory described above, the reason you have to benefit someone who has no right to your beneficence is only a non-requiring moral reason, whereas, on utilitarianism, the reason you have to benefit someone (even yourself) is always a requiring reason.

These two, of course, are not logical contradictories, but instead logical contraries. The logical contradictory of a requiring reason is a non-requiring reason. All non-moral reasons are non-requiring reasons, but some moral reasons are also non-requiring reasons: viz., non-requiring moral reasons. For our present purposes, the notion of a 'requiring reason' is most important, which I will define as follows:

- A4** P has a *requiring reason* to perform x if and only if it is the case that, if P has neither an undefeated nor a morally undefeated reason to perform some other available act, P's performing x is morally obligatory.¹⁷

Lastly, I will define a moral option as follows:

- A5** P has a *moral option* to perform either x or y if and only if it is both morally permissible for P to perform x and morally permissible for P to perform y .

3. The argument

With these analytic truths in hand, we are now in a position to consider what I take to be a very troubling argument, which I present in standard form below. Assume that 'P' stands for a person who must choose between

¹⁷ Requiring reasons are not reasons that necessarily do require; rather, they are reasons that have a requiring force to them—a force that, if left unchecked by countervailing reasons, generates a moral requirement. In calling such reasons 'requiring reasons', I don't mean to suggest that such reasons are never responsible for an act's being supererogatory. Indeed, as should become clear later on, I think that a requiring reason to perform x will often account for why doing x is morally supererogatory.

acting so as to secure a considerable benefit for herself and acting so as to secure an only *slightly* more considerable benefit for some stranger.¹⁸ Let us call the former ‘*s*’ since it is a self-regarding act and the latter ‘*o*’ since it is an other-regarding act.¹⁹ Assume that there are no other morally relevant facts. So, for instance, assume that whatever it is that P would be doing were she to perform *s*, it would not entail stealing something, causing someone harm, or anything of the sort. And assume, for the sake of simplicity, that *s* and *o* are the only available options and that they are mutually exclusive. Now consider the following argument, and for the moment, leave aside any worries you may have concerning the plausibility of P2, P4, or P6—I will discuss those premises in the following section.

- P1** If P has a requiring reason to perform *o*, then, if P has neither an undefeated nor a morally undefeated reason to perform *s*, P is morally required to perform *o*. (From A4 and the fact that *o* and *s* are the only two alternatives.)
- P2** P has a requiring reason to perform *o*.
- C1** Therefore, if P has neither an undefeated nor a morally undefeated reason to perform *s*, P is morally required to perform *o*. (From P1 and P2.)
- P3** If P has a moral option either to perform *o* or to perform *s*, then P is not morally required to perform *o*. (From A5.)
- P4** P has a moral option either to perform *o* or to perform *s*.

¹⁸ As suggested in section 1, there is an alternative version of the argument where ‘P’ stands for a person who must choose between acting so as to secure a considerable benefit for herself by pursuing some core life project (such as a vocation as a philosopher) and acting so as to secure a far more considerable net benefit for various needy, distant strangers by dedicating her time, effort, and money to an organization such as Oxfam instead. The resulting argument would be quite compelling in that P4 below would, then, express our considered moral conviction that forgoing one’s core life project in order to do more to promote the impersonal good is morally optional, not morally required. However, it seems to me that the version of the argument given in the body of this paper is even more compelling in that we have an even stronger conviction that forgoing a considerable benefit for oneself in order to provide some stranger with an only *slightly* more considerable benefit is morally optional, not morally required—see Brad Hooker, *Ideal Code, Real World* (Oxford: Clarendon Press, 2000), pp. 151-2.

¹⁹ Yet another variant would be where *s* entailed acting so as to secure a great benefit for yourself by taking advantage of some once-in-a-lifetime opportunity and *o* entailed acting so as to secure an *inconsiderable* benefit for some stickler who you’ve promised to help with her move and who refuses to release you from your promise despite the presence of other helpers who are quite capable of getting the job done without you, only it may take a bit longer.

- C2** Therefore, P is not morally required to perform *o*. (From P3 and P4.)
- C3** Therefore, P has either an undefeated or a morally undefeated reason to perform *s*. (From C1 and C2.)
- P5** If P has a morally undefeated reason to perform *s*, then it is not the case that P has better moral reason to perform *o* than to perform *s*. (From A1.)
- P6** P has better moral reason to perform *o* than to perform *s*.
- C4** Therefore, it is not the case that P has a morally undefeated reason to perform *s*. (From P5 and P6.)
- C5** Therefore, P has an undefeated reason to perform *s*. (From C3 and C4.)
- P7** If P has an undefeated reason, but not a morally undefeated reason, to perform *s* and a requiring reason to perform *o*, then P has a moral option either to perform *o* or to perform *s* only if moral reasons do not always override non-moral reasons in the determination of an act's moral status. (From A1-A5.)
- C6** Therefore, P has a moral option either to perform *o* or to perform *s* only if moral reasons do not always override non-moral reasons in the determination of an act's moral status. (From P2, C4, C5, and P7.)
- C7** Therefore, moral reasons do not always override non-moral reasons in the determination of an act's moral status. (From P4 and C6.)²⁰

P1, P3, P5, and P7 are all analytic truths given A1-A5. For all but P7, this should be obvious. Because P7 is less obvious, I will explain the reasoning behind it. Let me begin by provisionally assuming that both the antecedent of P7 and the antecedent in the consequent of P7 are true. Thus I will assume that P does have an undefeated, but not a morally undefeated,

²⁰ This argument is inspired by Shelly Kagan's argument against agent-centered options in his *The Limits of Morality*. But whereas Kagan assumes that non-moral reasons are irrelevant in the determination of an act's moral status and so argues against agent-centered options, I assume that there are agent-centered options and so argue that non-moral reasons must be relevant in determining an act's moral status. Kagan says, "since we are concerned with what is required by *morality*, the relevant reasons—whether decisive or not—must be moral ones" (p. 66). But Kagan's inference is unwarranted; we should not just assume that non-moral reasons are irrelevant with regard to what is required by morality. Also, unlike Kagan, I do not assume that ordinary morality is committed to a *pro tanto* moral reason to promote the impersonal good. And I do not assume that a morally decisive reason always generates a moral requirement; it can instead generate merely a supererogatory ought.

reason to perform *s*. This means that the undefeated reason that P has to perform *s* must be a non-moral reason, for if it were, to the contrary, a moral reason, P would have a morally undefeated reason to perform *s* as well. Now since this non-moral reason that P has to perform *s* is not defeated by the requiring (moral) reason that P has to perform *o* (P's reason to perform *s* is, after all, an undefeated reason), it follows that moral reasons (even requiring ones) do not always override non-moral reasons. Furthermore, since we are to assume that P has a moral option either to perform *o* or to perform *s*, it follows that P is not morally required to perform *o* despite the requiring reason that P has to perform *o*. And this means that when non-moral reasons defeat requiring reasons, they not only make it rationally permissible, but also morally permissible, to act contrary to what one has best requiring reason to do. For if, to the contrary, non-moral reasons were unable to prevent requiring reasons from giving rise to moral requirements when they defeat them (if, for instance, moral permissibility were a function of solely *moral* reasons), then the undefeated non-moral reason that P has to perform *s* would be powerless to prevent the requiring reason that P has to perform *o* from giving rise to a moral requirement to perform *o*. Thus if the antecedents in P7 are true, it follows that moral reasons do not always override non-moral reasons in the determination of an act's moral status, and, thus, non-moral reasons, when undefeated, can prevent moral reasons (even requiring ones) from generating moral requirements. So P7 is true.

4. Analysis and implications

Since P1, P3, P5, and P7 are all analytic truths and so unassailable, we must either (a) reject at least one of P2, P4, and P6 or (b) accept C7. I will argue that we should opt for the latter, since P2, P4, and P6 represent three of our most firmly held moral convictions. Other things being equal, a moral theory ought to comport with our considered moral convictions. This, of course, does not mean that any theory that fails to satisfy this desideratum is ultimately untenable. Theory selection is a matter of selecting that theory, from among all the alternatives, that best meets our desiderata, and since it is doubtful that any moral theory will fully satisfy all of our desiderata, a moral theory can fail to satisfy one desideratum and still be the best theory. Nevertheless, a theory that comports with our considered moral convictions is, *other things being equal*, more plausible than one that does not.²¹ So we should be reluctant to reject any of P2, P4, and P6.

²¹ For a defense of this claim, see Hooker, *Ideal Code, Real World*, pp. 9-19.

I should also note that I am not unsympathetic to the thought that a number of our commonsense moral intuitions are suspect. In particular, I think that Peter Singer and Peter Unger have successfully shown that we should be highly suspicious of many people's untutored intuition that it is permissible for those who enjoy the kind of affluence that is so common in industrialized nations to spend large portions of their surplus income on luxury items when there are so many children in developing countries who are dying easily preventable deaths.²² But the fact that some of our commonsense moral intuitions seem suspect upon careful reflection casts no doubt on the idea that our moral theories should comport with the moral convictions about which we remain confident even *after* careful reflection. Thus, it is important to point out that neither Singer's arguments nor Unger's arguments speak against our considered moral conviction that forgoing a benefit for oneself in order to provide some stranger with an only *slightly* greater benefit is morally optional, not morally required.

Up to this point, I have merely asserted, not argued, that P2, P4, and P6 represent three of our most firmly held moral convictions. Let me now rectify this, taking each in turn.

P2 says, "P has a requiring reason to perform *o*." A moral theory must countenance P2 in order to account for our conviction that in the absence of either an undefeated or a morally undefeated reason to do otherwise, P is morally required to perform *o*. Recall that P is a person who must choose between acting so as to secure a considerable benefit for herself and acting so as to secure a slightly more considerable benefit for some stranger. To illustrate, let us suppose that the specifics are as follows. A woman named Fiona is currently accessing her savings account via the Internet, and she is about to transfer the entire balance to her escrow company so as to purchase a house of her own. She can do so by clicking on button A. However, there is an alternative. By clicking on button B instead, her savings will be transferred, not to her escrow company, but to some stranger who will benefit slightly more from the money than she would. Clearly, given the tremendous sacrifice involved, our considered moral conviction is that Fiona is not morally required to perform *o*—that is, she is not morally required to click on button B. But it is equally clear that the fact that her doing so would provide the stranger with such a considerable benefit constitutes a requiring reason for her to click on button B. Indeed, but for the costs involved, she would be required to click on button B.

²² See Unger's *Living High and Letting Die* (Oxford: Oxford University Press, 1996) and Singer's "Famine, Affluence, and Morality," *Philosophy and Public Affairs* 1 (1972): 229-243.

To see this, consider the following variant of the above case. In this case, Fiona can transfer the money to her escrow company by clicking on *either* button A or button B, and, in this case, a very rich man has agreed to transfer an equivalent sum of his own money to the stranger if, and only if, Fiona clicks on button B. So, in either case, Fiona will purchase her new home, but, by clicking on button B, she will also secure a slightly more considerable benefit for the stranger. Assume that there are no other relevant facts. Surely, in this case, she is morally required to click on button B, for there is no good reason why she should not do so. By clicking on button B, she can purchase her new home while also providing a similar benefit for the stranger, and she can do so at no cost to herself, at minimal cost to the rich man (who, given the diminishing marginal utility of money, has more money than he can effectively use to benefit himself), and at absolutely no cost to anyone else. If you think that beneficence is only required when the would-be beneficiary is in great need or below a certain threshold of well-being, then assume that both Fiona and the stranger are in great need and/or below this threshold.

Given that we think that the reason Fiona has to click on button B gives rise to a moral requirement in the absence of either an undefeated or a morally undefeated reason to do otherwise, we must conclude that it is a requiring reason. If, to the contrary, Fiona's reason to click on button B were a non-requiring reason, it could not give rise to a moral requirement under such circumstances. So Fiona's reason to click on button B is a requiring reason.

The only relevant difference between this case—the case of Fiona's clicking on button B and helping the stranger at no cost to herself—and the case of P's performing *o* is how costly it is for the respective agents to provide the stranger with a considerable benefit. But surely it is implausible to suppose that an agent's requiring reason to benefit a stranger pops in and out of existence as the costs of her doing so varies.²³ Suppose, for instance, that we were to gradually increase the cost of clicking on button B, from no cost at all, to 10 cents, to 20 cents, to 30 cents, etc. It is not as if there is less and less to be said in favor of providing the stranger with a considerable benefit as the cost of clicking on button B increases for the agent. At least, that is not what the phenomenology of the case tells us, for it feels like a case of one reason being outweighed by another, not like a case of one reason being undermined by another. If it

²³ See Kagan, *The Limits of Morality*, 49. Unlike Kagan, though, I have not committed myself to a *pro tanto* requiring reason to promote the impersonal good.

were the latter, then once the cost was high enough, Fiona should cease to feel any pull toward clicking on button B. But even when the cost is extremely high, it is still clear that the fact that clicking on button B would result in a considerable benefit for the stranger speaks in favor of Fiona's doing so, morally speaking. In any case, even if I am wrong about this, even if Fiona's reason for clicking on button B is undermined by the countervailing cost-related reasons involved, P2 is still true, and, moreover, the argument still shows that non-moral reasons can defeat moral reasons; it is just that, in that case, these non-moral reasons would do so by undermining rather than outweighing the relevant moral reasons. Either way, though, we must conclude that P has a requiring reason to perform *o* given that we accept that the reason Fiona has to click button B (that it will provide the stranger with a considerable benefit) gives rise to a moral requirement in the absence of either an undefeated or a morally undefeated reason to do otherwise. To comport with this conviction, a moral theory must countenance P2.

P4 says, "P has a moral option either to perform *o* or to perform *s*." To deny that P has such a moral option, we would have to either accept, as the ethical egoist does, that P is morally required to promote her own self-interest or accept, as the act-utilitarian does, that P is morally required to do what best promotes the impersonal good.²⁴ Yet our considered moral conviction is that agents who find themselves in P's circumstances have the moral option of either furthering their own interests or sacrificing those interests for the sake of doing more to promote the impersonal good. To comport with this conviction, a moral theory must accept P4.

P6 says, "P has better moral reason to perform *o* than to perform *s*." To deny P6 is to reject our considered moral conviction that forgoing a benefit for oneself in order to secure a greater benefit for another is something that it would be morally good to do in P's circumstances. Doing so may not be obligatory, but that does not mean that it is not morally good. Indeed, the fact that doing so is supererogatory entails that P must have better moral reason to secure the greater benefit for the stranger than to secure the lesser

²⁴ It might be suggested that the ethical egoist who accepts a desire-fulfillment account of well-being could accommodate an option to perform either *o* or *s* provided P's desires would be equally fulfilled whether P performed *s* or *o*. However, I have stipulated from the start that the choice between benefiting oneself and benefiting the stranger is a mutually exclusive one. And, as I have stipulated, the choice between *s* and *o* is the choice between furthering one's own interests and *sacrificing* those interests for the sake of doing more to promote the impersonal good. Therefore, by stipulation, *o* must be a self-sacrificing act, and so the ethical egoist must hold that *o* is morally impermissible.

benefit for herself.²⁵ A moral theory, then, must accept P6 if it is going to comport with our considered moral convictions.

This brings us to the second horn of the dilemma: if we are going to accept P2, P4, and P6, then, as the argument shows, we must accept C7 as well: the view that moral reasons are not morally overriding. And it is not just P2, P4, and P6 that we would have to give up if we were to deny C7, for there are many other instances in which there are two acts open to a person P, one self-regarding and the other other-regarding (*s* and *o*), such that we have the following three moral convictions: (1) P has a requiring reason to perform *o*, (2) P has the moral option either to perform *s* or to perform *o*, and (3) P has better moral reason to perform *o* than to perform *s*. For each instantiation of *s* and *o* where we have these three moral convictions, (1)-(3) above will serve as analogues of P2, P4, and P6 in a new argument for C7. In section 1, for instance, I described a case where *s* is P's acting so as to secure a considerable benefit for herself by dedicating her resources (her time, effort, and money) to some core life project and *o* is P's acting so as to secure a far more considerable net benefit for various needy, distant strangers by instead dedicating her resources to an organization such as Oxfam. With this instantiation of *s* and *o*, we get P2', P4', and P6' and a new argument for C7. We see, therefore, that the price of denying C7 is quite high: we will have to reject not only P2, P4, and P6, but all their analogues as well.

As if this were not sufficient reason to accept C7, I provide further support for C7 in the following section, arguing that if we want to dissolve an apparent paradox regarding supererogation while accounting for many typical instances of supererogation, we must accept that non-moral reasons can defeat moral reasons in the determination of an act's moral status.

5. Moral reasons and supererogation

In this section, I argue that in order to account for many typical instances of supererogation we must accept that moral reasons do not always override non-moral reasons in the determination of an act's moral status. I will begin with a statement of two necessary conditions for an act's being supererogatory:

- NC P's performing *x* is supererogatory only if there exists some available alternative, *y*, such that (a) P is morally permitted

²⁵ See section 5 below.

both to perform x and to perform y , and (b) P has better moral reason to perform x than to perform y .²⁶

Although there is wide agreement among philosophers that a supererogatory act has to be, in some sense, morally superior to its non-supererogatory alternatives, some might deny that (b) is necessary, claiming instead that a supererogatory act need only be more morally praiseworthy than its non-supererogatory alternatives, and, on some accounts of praiseworthiness, an act can be more morally praiseworthy than another without there being better moral reason to perform it.²⁷ While I do not want to deny (or assert) that a supererogatory act must be more morally praiseworthy than its non-supererogatory alternatives, I do want to argue that (b) is necessary.

To see why (b) is necessary, consider the following case. Suppose that Jane must choose either to buy a new pair of shoes to wear around town or to donate the money that she would have spent on those shoes to Oxfam, and suppose that both options are morally permissible. Further suppose that Jane falsely believes that buying a new pair of shoes would be what is best for others, whereas donating the money to Oxfam would be what is best for her. Suppose, then, that she fails to appreciate the moral reasons that there are for her to donate the money to Oxfam, but that she thinks that she has a compelling self-interested reason to do so, for she believes that people will like her more (something she cares about) if she donates the money to Oxfam. She also falsely believes that she has good moral reason to buy the shoes, for she believes (falsely) that people will really enjoy seeing her wear them around town. So Jane decides, with altruism in her heart, to buy the shoes. Given her actual beliefs, motives, and intentions, one might plausibly claim that her buying the shoes is more morally praiseworthy than her donating the money to Oxfam. Yet, given that there is in fact no moral reason for her to buy the shoes and

²⁶ I leave open the question of whether there are any further necessary conditions, such as "P's performing x is more morally praiseworthy than P's performing y ."

²⁷ Someone else might deny (b), suggesting that a supererogatory act is one that involves a greater self-sacrifice for the sake of others than is required, whether or not there is necessarily any moral reason for agents to make such self-sacrifices—see, for instance, Sterling Hardwood, "Eleven Objections to Utilitarianism," in L. Pojman (ed.), *Moral Philosophy: A Reader* (Indianapolis: Hackett Publishing, 1998). One problem with such an account is that it rules out the possibility of supererogation with respect to self-regarding duties. Yet it certainly seems possible to go above and beyond what such duties require. For more on this and for a discussion of other problems with this account, see my "Consequentializing Moral Theories," *Pacific Philosophical Quarterly* 88 (2007): 39-73.

considerable moral reason for her to donate the money to Oxfam, it is hard to see how we could call her choice to buy the shoes supererogatory. Of course, the easy reply here is to insist that an act cannot be morally praiseworthy if the agent's choice of action is the result of her failure to appreciate the force of the relevant moral reasons, and I am quite sympathetic to this reply. But if one concedes that in order for an act to be morally praiseworthy the agent must properly appreciate the relevant moral reasons, then one must also concede that in order for an act to be more morally praiseworthy than another there will have to be better moral reason to perform it. And this is just to concede my point: that (b) is a necessary condition for an act's being supererogatory.

A further reason to accept (b) is that it nicely accounts for the normative force that supererogatory acts have, for there is a sense in which supererogatory acts are acts that agents morally ought to perform. It would be better, morally speaking, were they to do so. Indeed, the facts that make an act supererogatory are supposedly considerations that, morally speaking, count in favor of performing that act as opposed to some non-supererogatory alternative. But if, contrary to (b), agents do not have any better moral reason to perform a supererogatory act than to perform its non-supererogatory alternatives, then it is hard to see why it would be morally advisable for them to do so and why there is supposedly something that, morally speaking, counts in favor of their doing so.

Once we accept NC, however, supererogation can seem almost paradoxical since (b) can appear to be in tension with (a), as James Dreier has explained:

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 have no reason *from the moral point of view* [that is, no moral reason] to refrain from it (if you do have 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 your...[performing the supererogatory act], and nothing at all speaks against it. [In] what sense is it "all right," "permissible," "not wrong" to fail [to do so]? There seems to be no sense at all. Supererogation, according to this way of seeing things, turns out to be impossible.²⁸

²⁸ "Why Ethical Satisficing Makes Sense and Rational Satisficing Does not," in *Satisficing and Maximizing*, ed. by Michael Byron (Cambridge: Cambridge University Press, 2004), 148.

To dissolve the apparent paradox, the supererogationist must explain why the morally undefeated reason that favors performing the supererogatory act fails to generate a moral requirement to so act. There are only two possible explanations, for either moral reasons are morally overriding or they are not. If, on the one hand, they are, then the only possible explanation for why the morally undefeated reason that favors performing the supererogatory act fails to generate a moral requirement is that it is a *non-requiring* moral reason, the kind of moral reason that is incapable of giving rise to a moral requirement even when morally decisive.²⁹ Call this *the non-requiring reason explanation*. If, on the other hand, moral reasons are not morally overriding, then another possible explanation for why the morally undefeated reason that favors performing the supererogatory act fails to generate a moral requirement is that the agent has an undefeated non-moral reason to perform some non-supererogatory alternative instead. Call this *the non-moral reason explanation*. On the non-moral reason explanation, the thought is that moral reasons (even requiring ones) are not morally overriding, and thus the undefeated non-moral reason for performing some non-supererogatory alternative can prevent the requiring reason for performing the supererogatory act from being morally decisive and thereby generating a moral requirement. On this view, then, what explains the fact that it is morally permissible for an agent to fail to perform what is supererogatory is that, although she may have most requiring reason to perform the supererogatory act, she has no worse reason, all things considered, to perform some non-supererogatory alternative.

Philosophers such as James Dreier and Michael J. Zimmerman reject the non-moral reason explanation, only because they assume (explicitly in Dreier's case and implicitly in Zimmerman's case) that moral reasons are

²⁹ One might rightly point out that even a morally undefeated *requiring* reason will fail to generate a moral requirement when it is opposed by some moral reason of equal or incommensurate weight. But this cannot explain why the morally undefeated reason that favors performing the supererogatory act fails to generate a moral requirement in those instances where it defeats whatever moral reasons there are for performing some permissible non-supererogatory alternative. Moreover, the morally undefeated reason that favors performing the supererogatory act must defeat (not just equal) these moral reasons for performing the permissible non-supererogatory alternative if it is to meet condition (b) of NC. Thus the explanation for why the morally undefeated reason for performing the supererogatory alternative fails to generate a moral requirement has to be that it is a non-requiring moral reason.

morally overriding.³⁰ On their view, it can never be morally permissible to do something one has worse moral reason to do, even if the agent has an undefeated non-moral reason to do so. Hence, they both take the non-requiring reason explanation to be the only tenable explanation.

In an effort to spell out how the non-requiring explanation might go exactly, Dreier speculates that there might be two moral points of view, one the point of view of justice and the other the point of view of beneficence. Dreier further speculates that the reasons stemming from justice are requiring reasons and that the reasons stemming from beneficence are non-requiring moral reasons. Thus supererogatory acts are, according to Dreier, more beneficent, but not more just, than their non-supererogatory alternatives.³¹ So although agents have better moral reason to perform a supererogatory act than to perform any of its non-supererogatory alternatives, they are not morally required to do so, for the reasons involved are merely non-requiring moral reasons (i.e., reasons of beneficence), whereas what the agent is morally required to do is a matter of only requiring reasons (i.e., reasons of justice), and there simply is not better requiring reason to perform the supererogatory alternative.

The problem with Dreier's proposed account and with the non-requiring reason explanation in general is that it rests on the mistaken assumption that the reasons that make a supererogatory alternative morally superior to its non-supererogatory alternatives are always non-requiring moral reasons. To the contrary, it seems that in many typical instances of supererogation the moral reason that favors performing the supererogatory alternative over its non-supererogatory alternatives is a requiring reason. To illustrate, recall the case from section 3 where the agent must choose between acting so as to secure a considerable benefit for herself (i.e., performing *s*) and acting so as to secure a slightly more considerable benefit for some stranger (i.e., performing *o*). In this case, her performing *o* is clearly supererogatory. But given that Dreier and Zimmerman insist on the non-requiring reason explanation, they must

³⁰ That Dreier thinks that an act's moral status is a function of solely moral reasons is clear from the first sentence in the above quote. Zimmerman, by contrast, is less explicit, but he does say that if there being better moral reason to perform the supererogatory alternative is essential to supererogation, then any theory wishing to accommodate supererogation will have to declare that there are two sets of moral reasons, deontic and non-deontic reasons (or what I am calling requiring and non-requiring moral reasons)—see his "Supererogation and Doing the Best One Can," *American Philosophical Quarterly* 30 (1993): 373-80, pp. 375-6.

³¹ Similarly, in note 11, Zimmerman offers an example where he supposes that reasons of fidelity are requiring reasons and reasons of beneficence are non-requiring moral reasons.

presume that, in this case, the moral reason that favors her performing *o* as opposed to *s* is a non-requiring moral reason. If the reason were instead a requiring reason, then her performing *o* would, on their view, be obligatory, not supererogatory, because, on their view, whatever non-moral reason she has to perform *s* is powerless to prevent the better requiring reason she has to perform *o* from giving rise to a moral requirement to perform *o*. But, as we have already seen in section 4, the moral reason that favors her performing *o* is a requiring reason; absent either an undefeated or a morally undefeated reason to do otherwise, the agent is *required* to perform *o*. And there are many other cases like this where the moral reason one has to perform some beneficent and supererogatory act is a requiring reason. So if we are to account for many typical cases of supererogation, we are going to have to accept the non-moral reason explanation, thereby accepting, contrary to Dreier and Zimmerman, that non-moral reasons, and not just moral reasons, are relevant to the determination of an act's moral status.

However, if we accept the non-moral reason explanation, as I have argued that we must, the performance of a supererogatory act could, then, turn out to be *objectively irrational*—that is, it could turn out to be contrary to what the agent has best/decisive reason to do, all things considered.³² Some philosophers will find this implication implausible and will, thus, take this to be a reason to reject the non-moral reason explanation.³³ But we should not be so quick to reject the non-moral reason explanation. For one, it is not clear that, on this explanation, there will be many (if any) instances where a supererogatory act turns out to be objectively irrational. For another, it is not clear that we should find it implausible to suppose that some supererogatory acts are objectively irrational.

With regard to the first point, it is important to keep in mind that taking the non-moral reason explanation to be the best explanation in certain instances does not preclude taking the non-requiring reason explanation to be the best explanation in other instances—specifically, those where the supererogatory alternative is favored by a non-requiring moral reason as

³² Recall that an act is objectively irrational just when the agent has decisive reasons not to perform that act.

³³ The account of supererogation that I gave in my “Position-Relative Consequentialism, Agent-Centered Options, and Supererogation” did have the implication that all supererogatory acts are objectively irrational, and Michael Byron and Betsy Postow rightfully objected to it for this reason. See Michael Byron’s “Alternative Consequentialisms” (manuscript) and Betsy Postow’s “Supererogation Again,” *Journal of Value Inquiry* 39 (2005): 245-253.

opposed to a requiring reason.³⁴ So how often supererogatory acts turn out to be objectively irrational will depend in part on how often they are favored by requiring as opposed to non-requiring reasons. Moreover, whether and how often supererogatory acts turn out to be objectively irrational in those cases where the supererogatory act is favored by requiring reasons (cases where we must adopt the non-moral reason explanation) depends on what, in each case, accounts for the fact that the non-moral reason that favors performing some non-supererogatory alternative is undefeated. Recall from section 2 that there are only three possibilities as to why P's reason for performing some non-supererogatory alternative is undefeated even in the face of a morally undefeated reason to perform the supererogatory alternative: (i) it defeats (by outweighing and/or undermining) P's reasons for performing the supererogatory alternative, (ii) it is equaled by (or on a par with) P's reasons for performing the supererogatory alternative, or (iii) it is incommensurate with P's reasons for performing the supererogatory alternative, such that there is no truth as to whether the reasons for performing the one are better than, worse than, or just as good as the reasons for performing the other.³⁵ The only one of these three that would result in a supererogatory act's being objectively irrational is (i). So how often supererogatory acts turn out to be objectively irrational on the non-moral reason explanation depends on how often (i) is the case. If, for instance, it turns out that the sorts of reasons that favor performing supererogatory acts are not comparable with the sorts of reasons that favor performing their permissible non-supererogatory alternatives, then the non-moral reason explanation will never imply that a supererogatory act is objectively irrational. Alternatively, if it turns out that the sorts of reasons that favor performing supererogatory acts are only very roughly comparable with the sorts of reasons that favor performing their permissible non-supererogatory alternatives, or if it turns out that the two are very often equivalent or on a

³⁴ From the fact that the non-moral explanation is needed to account for many typical instances of supererogation, instances in which the relevant reasons are requiring reasons, it does not follow that the non-requiring explanation is not the correct explanation in cases where the relevant reasons are non-requiring moral reasons.

³⁵ In my "Imperfect Reasons and Rational Options" (manuscript), I argue that (ii) is often the case, for the relevant reasons involved are typically what I call *imperfect reasons*, reasons that do not count in favor of performing some particular act-token, but instead either count in favor of performing one of some set of act-tokens that would each produce the same result or count in favor of performing a certain type of act (e.g., the altruistic type) with a certain frequency over a certain time frame, where the desired result will be achieved so long as that act-type is performed with that frequency over the given time frame.

par, then the non-moral reason explanation will only rarely imply that a supererogatory act is objectively irrational. In sum, if, in a given case, the reasons favoring the supererogatory alternative are themselves undefeated because they are either tied for best or incommensurate with the non-moral reasons that are preventing them from generating a moral requirement, then the non-moral reason explanation will not imply that the supererogatory alternative is objectively irrational.

Now the sorts of reasons that support supererogatory acts are typically altruistic (or impartial) reasons, and the sorts of reasons that support their permissible non-supererogatory alternatives are typically self-interested (or partial) reasons. Although I will not argue that these two types of reasons are either wholly incomparable or only roughly comparable, I should point out that these positions are not without their supporters.³⁶ Henry Sidgwick, for instance, held that impartial reasons and self-interested reasons are wholly incomparable, and this led him to accept a kind of *dualism of practical reason*, where it is always objectively rational to do either what would be impartially best or what would be self-interestedly best. Recently, Derek Parfit has argued that some form of dualism about practical reason is correct, but that Sidgwick was wrong to think that impartial and self-interested reasons are wholly incomparable. Parfit argues that they are instead only roughly comparable.³⁷

If Sidgwick is right, then (iii) will always be the case and so the non-moral reason explanation will never imply that supererogatory acts are objectively irrational. And if Parfit is right, then (ii) will often be the case and thus the non-moral reason will rarely, if ever, imply that supererogatory acts are objective irrational. Again, I will not defend either view. Nor is it important that I do so, for the plausibility of the non-moral reason explanation does not depend on my establishing any particular claim concerning to what extent, if at all, such reasons are comparable. To see this, we need only take note of what the various possibilities are in any

³⁶ Two types of reasons, R1 and R2, are *wholly incomparable* if and only if there is, for *no pair* of token instances of R1 and R2, some truth as to whether the one is better than, worse than, or just as good as (or on a par with) the other. And two types of reasons, R1 and R2, are *only roughly comparable* if and only if there is, for *only a few pairs* of token instances of R1 and R2, some truth as to whether the one is better than, worse than, or just as good as (or on a par with) the other. Just how rough the comparability of the two types of reasons is in proportion to how few pairs there are for which there is some truth as to whether the one is better than, worse than, or just as good as (or on a par with) the other. The notions of ‘wholly incomparable’ and ‘roughly comparable’ come from Parfit’s *Climbing the Mountain*. However, the definitions are mine.

³⁷ See his *Climbing the Mountain*.

given case of supererogation. So take any other-regarding and supererogatory act, *o*, and take any self-regarding and permissible non-supererogatory alternative, *s*. Either there is some truth as to whether the reasons supporting *o* are better than, worse than, or just as good as the reasons supporting *s* or there is not. If there is not, then the non-moral reason explanation will not imply that *o* is objectively irrational. If there is, then either the reasons supporting *o* are just as good as the reasons supporting *s* or they are not. If they are, then the non-moral reason explanation will not imply that *o* is objectively irrational. If they are not, then either the reasons supporting *o* are better than or worse than the reasons supporting *s*. Now assume that the reasons supporting *o* are requiring reasons. (If they are not requiring reasons, then we can employ the non-requiring reason explanation.) If they are requiring reasons, then, on the non-moral reason explanation, they must be worse than the reasons supporting *s* so as to account for the fact that they do not generate a moral requirement. And so, on this scenario, *o* turns out to be objectively irrational on the non-moral reason explanation. Should we, therefore, reject the non-moral reason explanation? Well, that depends on what the alternative to saying that *o* is objectively irrational might be. Clearly, the only alternative to saying that the requiring reasons supporting *o* are worse than the non-moral reasons supporting *s* (given that we have already excluded those situations in which the opposing reasons are just as good as or incommensurate with one another) is to say that they are better than the reasons supporting *s*. And, in that case, *s* would be objectively irrational. But if we are willing to say that *s* could be objectively irrational, then we should also be willing to say that *o* could be objectively irrational; both are, after all, equally permissible alternatives. Thus, in the only sort of case where the non-moral reason explanation implies that *o* is objectively irrational, we find that the only other alternative, which is to say that *s* is objectively irrational, is no more plausible. Therefore, we should not hesitate to accept the non-moral reason explanation despite the fact that it may imply that some supererogatory acts are objectively irrational.

Besides, as I will now show, there will be some supererogatory acts that are objectively irrational no matter what account of supererogation we give—at least, there will be if the relevant reasons are commensurate. To see why, consider that there is often a spectrum of supererogatory acts available to an agent, some of which go further beyond what is required than others. For instance, I could volunteer for a worthy non-profit organization on Saturday mornings, but I could also volunteer all day on Saturdays or even all day on both Saturdays and Sundays. And where

there is such a spectrum of supererogatory alternatives available to me, it will often turn out that most of them are objectively irrational if the relevant reasons are commensurate, as it is unlikely that in all such cases I will have just as good a reason, all things considered, to perform any one of them as to perform any other. The thought, then, is that if the relevant reasons are commensurate, then I will have best reason to strike a certain balance between time spent volunteering and time spent working on self-interested projects. Given this equilibrium point, it would be objectively irrational for me to spend either any less time on my self-interested projects or any more time volunteering. Thus spending more time volunteering will be both supererogatory and objectively irrational. Spending more time volunteering will be supported by moral reasons but will be contrary to what I have best reason to do, all things considered. Of course, one could deny that, in this case, there is any truth as to whether I have better, worse, or equal reason to spend more time volunteering. But that is just to deny that the relevant reasons are commensurate, and if they are not commensurate, then the non-moral reason explanation does not imply that doing so is objectively irrational.

To sum up, I have shown that the non-moral reason explanation will imply that supererogatory acts are irrational only in those instances where such acts are supported by requiring reasons that are commensurate with, but not equaled by, the non-moral reasons there are to perform some non-supererogatory alternative. And I have argued that were we to deny that the supererogatory alternative is objectively irrational in such instances, we would instead have to claim that the permissible non-supererogatory alternative is objectively irrational, which is no more plausible. Furthermore, I have argued that given that there is oftentimes a spectrum of supererogatory acts open to an agent, any account of supererogation will imply that some supererogatory acts are objectively irrational if the relevant reasons are commensurate.

I will now end my discussion of supererogation by explaining what we should say about three very different sorts of cases. First, there are those cases where the reason favoring the supererogatory alternative is a non-requiring reason. To illustrate, suppose that it is obvious that an old man who is shopping nearby would like some help with getting a heavy item off the shelf. If I do not help him (and he has not asked for my assistance), he will need to wait a minute or two until the store clerk comes by. But I could very easily help him right now, so I ask if he would like some help, he affirms that he would, and I help him. My willingness to offer and provide assistance in such instances is, on most accounts, supererogatory.

Yet this does not seem to be a case where the moral reason I have to offer and provide assistance is a requiring reason; it is just too trivial of a moral reason to be one that would generate a moral requirement even in the absence of any countervailing reasons. So, in these sorts of cases, the non-requiring reason explanation is what best explains why I do not have a moral obligation to do what I have best moral reason to do.

Second, there are those cases where the reason favoring the supererogatory alternative is a requiring reason that is defeated by the non-moral reason there is to perform some permissible non-supererogatory alternative. Such seems to be the case where I face the choice between either saving myself from years of agony or saving some stranger from a brief moment of mild discomfort. If we assume, as seems to be the case on commonsense morality, that there is no moral duty to minimize one's own suffering and that failing to save myself from such agony would not in any way hamper my ability to fulfill my moral obligations, then we should consider saving the stranger from discomfort to be supererogatory. And if we also assume that, given the great discrepancy between what is at stake for myself and what is at stake for the stranger, we can compare my conflicting self-interested and altruistic reasons and rightly conclude that the former defeats the latter, then we should also conclude that my performing this supererogatory act is objectively irrational. And I see nothing counter-intuitive about such a verdict.

Third, there are those cases where the reason favoring the supererogatory alternative is a requiring reason that neither defeats nor is defeated by the non-moral reason there is to perform some permissible non-supererogatory alternative. This could be so either because the reasons supporting each course of action are exactly equal or because there just is no truth as to whether those in support of the one are better than, worse than, or just as good as those in support of the other. Such cases will, I think, account for the vast majority of cases of supererogation, and, in these cases, it will be objectively rational both to do what would be best for others (the supererogatory alternative) and to do what would be better for ourselves (some permissible non-supererogatory alternative). In such cases, doing what would be best for others is supererogatory, but not irrational.

In conclusion, we have found that we can plausibly dissolve the apparent paradoxical nature of supererogatory acts by accepting that non-moral reasons can affect an act's moral status. In addition, then, to the reasons offered in the previous sections (i.e., wanting to accept all of P2, P4,

and P6 and their analogues), we have yet another reason to reject the view that moral reasons are morally overriding.

6. Conclusion

I have argued that moral reasons are not morally overriding, that non-moral reasons can defeat moral reasons (even requiring ones) and thereby prevent them from giving rise to a moral requirement. At the very least, I have shown that we must accept this conclusion if we want to accommodate many of our moral convictions, such as those regarding both supererogation and agent-centered options. This means that those who wish to accommodate our moral convictions have their work cut out for them. It will not be enough to provide a criterion of rightness that yields deontic verdicts that comport with our moral convictions, for, as I have shown, this compartment commits them to the view that moral reasons are not morally overriding, and so they owe us an account of how both moral reasons and non-moral reasons come together to determine an act's moral status. This is no small task. First, it requires providing both a substantive account of what moral reasons there are and a substantive count of what non-moral reasons there are. Second, it requires providing a substantive account of how moral and non-moral reasons function together to determine an act's moral status. With regard to the second, we will need answers to the following sorts of questions: (1) "Is it ever wrong to do what one has best reason to do, all things considered?"; (2) "Is it ever wrong to do what one has best moral reason to do?"; and (3) "Is it ever permissible to do something that one has neither best moral reason to do nor best reason, all things considered, to do, and, if so, when?" I do not mean to suggest that philosophers have not been working on these issues; they have.³⁸ But what is not generally appreciated is the importance of resolving these issues if we are going to have any hope of defending a moral theory that comports with our most firmly held moral convictions.³⁹

³⁸ See, for instance, Ruth Chang, "Putting Together Morality and Well-Being," in *Practical Conflicts*, edited by Peter Baumann and Monika Betzler, (Cambridge: Cambridge University Press, 2004); Owen McLeod, "Just Plain 'Ought'," *The Journal of Ethics* 5 (2001): 269-291; Sarah Stroud, "Moral Overridingness and Moral Theory," *Pacific Philosophical Quarterly* 79 (1998): 170-189; and both my "Dual-Ranking Act-Consequentialism" and my "Position-Relative Consequentialism, Agent-Centered Options, and Supererogation."

³⁹ I thank Noell Birondo, Dan Boisvert, Campbell Brown, Michael Byron, Nir Eyal, Joshua Glasgow, Betsy Postow, Mark Schroeder, David Shoemaker, Mark van Roojen, Michael J. Zimmerman, students in my Spring 2006 seminar entitled "The Limits of Morality," and the audience at my 2006 Pacific APA colloquium for helpful comments and discussions.

*Department of Philosophy
Arizona State University
P.O. Box 874102
Tempe, AZ 85287-4102*

douglas.portmore@asu.edu