

Moral Worth and Supererogation

Amy Massoud

Introduction

In this paper I introduce a novel account of moral worth that is able to explain numerous moral phenomena, most notably moral supererogation. My account is inspired by a theory of moral worth recently defended by Julia Markovits, her Coincident Reasons Thesis (CRT).¹ Although CRT is initially appealing, it generates implausible results as it stands.² I make use of a distinctively moral version of Joshua Gert's distinction between the justifying and requiring strengths of reasons, paving the way for a more plausible, multidimensional account of the moral praiseworthiness of actions. My account also captures other important, often neglected features of the moral landscape, including the fact that some morally required actions warrant more moral praise than some morally supererogatory actions.

1. Moral Supererogation and Praiseworthiness

J.O. Urmson is credited with initiating contemporary philosophical discussion of supererogatory actions in his 1958 article "Saints and Heroes," and much has been written on the topic since. Morally supererogatory actions are traditionally conceived as actions that are non-obligatory, but morally praiseworthy. As pointed out by Urmson, the paradigmatic examples of such actions include actions that are saintly or heroic, e.g., risking one's life to save a stranger. In this section, I will present four sorts of morally praiseworthy actions that indicate features of praiseworthiness that any plausible theory of moral worth ought to capture. I'll present one example from each of the following categories: (1) morally required actions that warrant considerable moral praise; (2) morally required actions that warrant an insignificant amount of moral praise; (3) morally supererogatory

¹ Markovits (2010).

² Markovits (2012). She concludes that supererogatory actions are genuinely obligatory, in order to avoid relativism about moral requirements.

actions that warrant considerable praise; and (4) morally supererogatory actions that warrant at least some praise.³

1. *A morally required action warranting considerable moral praise*

Sam is rushing to the airport to catch an international flight. Because his ticket is nonrefundable, missing it will result in a substantial financial loss. Sam is not by any means wealthy. Moreover, because his destination is a very small island country, he will have to wait several days for the next flight out. On the way to the airport, Sam witnesses a woman suffering some sort of heart condition. Sam considers the personal sacrifice involved in helping this woman, but determines that he must help her, as his assistance could save her life. Sam helps the woman, thus missing his flight.

2. *A morally required action warranting little to no moral praise*

Eva is driving home from work and she refrains from swerving into oncoming traffic. (It would be odd to praise her simply for staying on the right side of the road, but it is a morally required action, given that she would be putting many lives in danger if she were to swerve over.)

3. *A morally supererogatory action warranting considerable moral praise*

Kate learns that her co-worker, Josh, is in desperate need of a kidney transplant. Josh has a partner and two small children. Kate manages, without Josh's knowledge, to get tested, and finds that she is indeed a match. In spite of potential complications and long-term health risks, Kate donates one of her healthy kidneys to Josh. (In order to maintain the moral permissibility of this case, we may want to suppose further that Kate has no dependents of her own.)

³ Note that there's no sharp line dividing actions of the first and second categories, nor between actions of the third and fourth categories. As will become clear below, I take this to be an intuitively plausible feature of any account of moral worth.

4. A morally supererogatory action warranting at least some moral praise

Bo is shopping at a flea market and he happens upon a scarf that he thinks his friend, Dot, will enjoy. He purchases it and gives to her the next day when they meet for lunch.

Now let us see what these examples reveal about moral worth. Regarding the first two examples, both of which are, intuitively, morally obligatory, it is evident that an acceptable theory of moral worth ought to recognize *degrees* of praiseworthiness regarding morally obligatory actions. Intuitively, Sam's action of missing his flight in order to help a stranger warrants more praise than Eva's action of staying on the right side of the road. Note that in each case, the details could be manipulated in order to adjust the amount of praise that each agent deserves. If Eva is a thrill-seeker with a burning desire to swerve, yet refrains out of concern for potential crash victims, it is (at least) plausible that her action of remaining on the correct side is one that warrants more moral praise.⁴

In considering the second two examples, both of which are, intuitively, morally praiseworthy, but not obligatory, it is evident that an acceptable theory of moral worth also ought to recognize degrees of moral worth regarding supererogatory actions. Kate's action of donating one of her healthy kidneys to Josh, incurring the risks that go along with it, is more morally praiseworthy than Bo's action of purchasing the scarf for Dot. Yet each action is supererogatory because, in addition to being morally worthy, each is optional.

Another important feature of moral worth revealed by these cases is that some morally required actions are more morally worthy, or warrant more moral praise, than some morally supererogatory actions. For example, Sam's action warrants more praise than Bo's, even though Sam's action is required, while Bo's is not.

⁴ To clarify, I'm comparing the moral worth of the original Eva case (where she has no desire to swerve) with the revised Eva case (where she is a thrill-seeker with a strong desire to swerve).

In sum, I take it that a complete theory of moral worth will acknowledge: (a) the existence of genuinely non-obligatory morally supererogatory actions; (b) degrees of moral worth, regarding both morally obligatory and supererogatory actions; and (c) that some morally required actions are more morally worthy, or warrant more moral praise, than some supererogatory actions.

2. The Justifying/Requiring Distinction

Joshua Gert has argued convincingly that normative practical reasons can play (at least) two different roles in determining the rational status of token actions: they may justify, and they may require.⁵ According to Gert, when a normative reason plays the requiring role regarding a certain action, it renders an otherwise rationally permissible action rationally impermissible. For example, although eating a banana walnut muffin for breakfast is otherwise rationally permissible, it is rationally impermissible for an individual with a tree nut allergy to do so (assuming that she is aware of her allergy and of the muffin's ingredients⁶). The fact that she has a tree nut allergy constitutes a normative reason to avoid the muffin, and in this case the normative reason is playing the requiring role by rendering the otherwise rationally permissible action of eating the muffin rationally impermissible. The fact that the agent's tree nut allergy is life-threatening has greater requiring *strength* than the fact that the agent's allergy will cause her to break out in a mild rash.⁷

When a normative practical reason plays the justifying role, it renders an otherwise rationally impermissible action rationally permissible. For example,

⁵ Gert (2007).

⁶ We may need to assume further that the agent has a *desire* to avoid an allergic reaction, but I leave this complication aside, as nothing I argue for here hinges on this issue.

⁷ Here is Gert's (2007) Counterfactual Criterion of requiring strength (p. 538): "One reason, A, has more requiring strength than another reason, B, if, in playing the requiring role in actual and counterfactual circumstances, A can overcome any reason or set of reasons that B can overcome, and there are some reasons or sets of reasons that A can overcome but B cannot." For the purpose of this paper, I do not wish to commit myself to such a strong criterion. I take it that types of reasons are generally comparable in this way, but I also take it that there may be exceptional cases that run counter to Gert's Counterfactual Criterion.

though it is otherwise irrational to drink a concoction that will make one very sick, it is (at least) rationally permissible for an agent to do so if it will save her life. The fact that the concoction will save her life constitutes a normative reason to consume it, and in this case the normative reason is playing the justifying role by rendering the otherwise rationally impermissible action of drinking the concoction rationally permissible. Furthermore, the fact that drinking the concoction will save the agent's life has greater justifying *strength* than the fact that drinking it will improve her complexion.⁸

Up to this point, I've been following Gert in characterizing these roles in terms of practical reasons more generally. However, I will utilize a distinctively moral characterization of Gert's distinction below.

3. Markovits's Coincident Reasons Thesis

Julia Markovits has recently offered an account of the conditions an agent's action must meet in order to qualify as morally worthy. Here is Markovits's Coincident Reasons Thesis of moral worth:

The Coincident Reasons Thesis: An action is morally worthy if and only if—and to the degree that—the noninstrumental reasons motivating the action coincide with the noninstrumental reasons that morally justify its performance.⁹

According to Markovits, normative reasons that morally justify actions are facts which determine that an agent *ought* to perform a given action. So Markovits's sense of "justifying" is more akin to Gert's requiring role played by a reason, as opposed to his justifying role.¹⁰ In order for a normative reason to justify an

⁸ Here is Gert's (2007) Counterfactual Criterion of justifying strength (p. 539): "One reason, A, has more justifying strength than another reason, B, if, in playing the justifying role in actual and counterfactual circumstances, A can overcome any reason or set of reasons that B can overcome, and there are some reasons or sets of reasons that A can overcome but B cannot." Also, see my comment in the previous footnote.

⁹ Markovits (2012), 290.

¹⁰ In footnote 2 (2012), she acknowledges Gert's distinction, but maintains the "stronger" notion of justifying, the sort that refers to an action one morally *ought* to perform.

action in Markovits's stronger, ought-making sense, it must meet the following characteristics: (a) the agent must have *epistemic access* to it; (b) an agent must have sufficient reason to *believe* it would be best to do the act; and (c) the agent must have adequate evidence that it would be best to do the act.¹¹ Motivating reasons are facts that provide an explanation as to why the agent acted the way she did. In order for a normative reason to motivate an action, it must be a fact on the basis of which an agent *chooses*. However, an agent need not recognize her own motivating reasons, nor need such reasons be justified in order to qualify as motivating. Finally, Markovits explains, the reasons relevant to CRT (both those that justify and those that motivate) are noninstrumental, in the sense that they are reasons to act that are provided by ends we pursue for their own sakes.

As Markovits points out, her CRT has a “weird” implication. Her thesis implies that all actions that are entirely motivated by all of the right-making reasons are of equal moral worth. But this is highly counterintuitive; some actions motivated by all of the right-making reasons warrant more moral praise than some other actions that also meet this criterion. To illustrate this point, consider the examples of morally required actions I gave in the first section: Eva continues to drive on the right side of the road, and, let us suppose, she is motivated by all of the right-making reasons (e.g. she's committed to safe driving habits for the benefit of other drivers and pedestrians), each one to the right degree. Sam stops to assist the stranger in spite of the personal sacrifice his action requires, and, let us suppose, he is also motivated by all of the right-making reasons (e.g. he recognizes that the victim needs his help), each one to the right degree. A satisfactory account of moral worth will not imply that these two actions are equally morally worthy. On the contrary, the latter action warrants more moral praise. A similar point can be made regarding my two examples of supererogation. Even if Kate and Bo each perform their actions for all of the right reasons, Kate's action of donating her kidney warrants more moral praise than

¹¹ Markovits (2012), 290-1.

Bo's action of purchasing the scarf does. In other words, Kate's action is supererogatory to a greater degree. Markovits spends the rest of her (2012) paper responding to the following question: In virtue of what might one completely morally worthy action, given CRT, be worthier than another? One of her chief concerns is to be able to account for saintly and heroic (i.e. supererogatory) actions.¹²

Markovits proposes the following solution: In accordance with CRT, a saintly or heroic action performed for the right reasons merely *seems* extra morally worthy to us when it is unlikely that other agents would have taken the same action in similar circumstances. An action is (e.g.) heroic, then, just in case it is an action that is judged by us as atypical; it is atypical in the sense that most of us would not be willing to do the same in relevantly similar circumstances. One implication of this view, which Markovits acknowledges and accepts, is that whether a given action is supererogatory or not is relative to the appraiser of the action and her peers. This is so because what is judged as atypical or extraordinary relative to a given situation will depend upon what is judged as typical or ordinary. And what is judged as typical or ordinary differs among, for example, cultures. And because Markovits wants to avoid the result that obligatoriness is appraiser-relative she is left with no option but to deny that supererogatory actions are genuinely non-obligatory. She writes:

Indeed, because, on my view, whether an action counts as heroic (as opposed to merely right, let's say) is relative to the appraiser, the view sits uncomfortably with any view of supererogation according to which actions qualify as supererogatory in virtue of being heroic *and* supererogatory actions are by definition non-obligatory: this combination of views would entail appraiser-relativism, not just about which actions are heroic but about which are obligatory. The very same action could be appropriately described as required and not required (because heroic),

¹² This is not to say that an action must be saintly or heroic in order to qualify as supererogatory. For example, my account explains why Bo's act of purchasing the scarf for Dot is supererogatory, even though it is not saintly or heroic.

depending on the speaker. I think we have good reason to resist such a strong form of moral relativism.¹³

What makes supererogatory actions *seem* like they are beyond duty, according to Markovits, is that the appraiser feels it would be hypocritical to expect or demand such actions from others when the appraiser, along with those with whom she is familiar, lack the moral strength to perform similar actions. Thus, supererogatory actions are some subset of obligatory actions; actions we have a *duty* to perform. Their “special moral worth” is just a matter of appraisers categorizing them as such in order to avoid hypocrisy.

I agree with Markovits that we have good reason to avoid the strong form of relativism that would result from a conjunction of the views that supererogation is appraiser-relative and that supererogatory actions are non-obligatory. However, I also think we have good reason to avoid the result that supererogation is appraiser-relative, as well as the result that supererogatory actions are obligatory. First, it seems wrong to insist that it is only appropriate for an individual to categorize an action as supererogatory if she doubts that she (or her peers) would have the moral strength to perform such an action. For example, consider Mary. Mary was raised in a tight-knit community, and among the members of her community it’s been commonplace for each individual to go out of his or her way to improve the lives of other members. Now Mary has moved to a new neighborhood. The very evening she moves into her new home, a neighbor, Noah, shows up at her door with a fresh baked pie and a sack of local coffee. It seems totally plausible that Mary may correctly deem such an action supererogatory, even if she has no doubt that both she and her peers would act similarly toward a new neighbor.¹⁴ That is, there’s no trouble in supposing that Mary would feel that her neighbor’s action lies beyond the call of neighborly

¹³ Markovits (2012), 300.

¹⁴ It could be objected that her new neighbors are now her peers, but I would reply that she need not know how her new neighbors typically behave in order to be justified in deeming their actions supererogatory.

duty, even absent any inkling to avoid hypocrisy. Moreover, the appraiser-relative characterization forces Markovits to categorize saintly and heroic actions as morally obligatory.

Furthermore, we can imagine a morally lazy society in which it is very rare for one individual to so much as lift a finger in order to benefit another. According to Markovits's purported solution, actions like effortlessly tossing a life preserver to a drowning individual could justifiably qualify as supererogatory within this society. This result is counterintuitive.

In spite of the falsity of CRT, given my arguments thus far, it does have some very appealing features. According to CRT, actions cannot be accidentally morally worthy. To illustrate, suppose an evil doctor intends to give an innocent patient a medication that will prolong his illness, but, unbeknownst to the doctor, the medication will actually cure the patient. CRT generates the desirable result that the doctor's action has no moral worth, since, we are supposing, the doctor's (non-instrumental) motivating reasons do not coincide at all with the existing normative (non-instrumental) reasons. Another appealing feature of CRT is that, unlike strict Kantian views of moral worth, an agent can perform a maximally worthy action, even if she is not solely motivated (or perhaps even motivated at all) by a sense of duty; arguably, there are other sorts of normative reasons that ought to motivate us. Thus, CRT includes a very appealing feature of strict Kantian, or motive-based moral worth, while overcoming a highly counterintuitive aspect of such views.

In what follows, I will offer an alternative account of moral worth that retains these appealing features, while avoiding the counterintuitive implications generated by Markovits's view. I will utilize a distinctively moral version of Gert's requiring/justifying distinction to develop a positive, coincident reasons-based account of supererogation that is not appraiser-relative.

4. A Moral Justifying/Requiring Distinction

Moral reasons, I'm stipulating, are distinctively *other*-regarding. Thus, according to the view I'll defend, agents have prudential, but not moral, duties to themselves.¹⁵ I follow Douglas Portmore in supposing that nonmoral reasons can affect an act's moral status.¹⁶ Such reasons are not (by definition) *moral* reasons, but they are *morally relevant* reasons. For example, the fact that taking a particular other-regarding action will cause me severe pain is a morally relevant reason (e.g. it may make an otherwise morally impermissible action morally permissible); though it is not, from my own perspective, a moral, or other-regarding, reason. Portmore accepts Gert's distinction between the requiring and justifying roles reasons can play, and he offers a distinctively moral characterization of these roles: A morally relevant reason plays the requiring role (regarding action *a*) when, absent this reason, it would be morally permissible for an agent to refrain from performing *a*. A morally relevant reason plays the justifying role (regarding action *a*) when, absent this reason, it would be morally impermissible for an agent to perform *a*.¹⁷

The view of moral worth I will ultimately defend takes Markovits's CRT as an initial step in determining the moral worth of an action. To illustrate how the moral version¹⁸ of the justifying/requiring distinction works in conjunction with CRT, consider the following case:

Holly: Holly has promised to drive Ira to the train station on Saturday.

I assume we can agree that, absent some acceptable justification, Holly is morally required to follow through. When Holly wakes up Saturday morning and manages to get Ira to the station on time, she does what she morally ought to do.

¹⁵ Although this is my preferred view, there's an interesting question as to whether one needs to accept this stipulation in order to accept my account of moral worth. For theorists who are partial to the notion of moral duties to self, I will make some suggestions later in the paper.

¹⁶ Portmore (2008).

¹⁷ Ibid. 372.

¹⁸ By this I mean the version of the justifying/requiring distinction that is relevant from the perspective of moral deliberation.

But is her action morally worthy? If so, how morally worthy is it? According to CRT, we must determine the answers to these questions by considering the extent to which Holly's motivating reasons overlap with the moral reasons in favor of her action. Here, for the sake of simplicity, are four exhaustive normative reasons confronting Holly:

- N1) She has promised to drive Ira to the train station.
- N2) Ira is counting on her for a ride.
- N3) She wants Ira to drive her to the airport next month.
- N4) Ira is a talented baker, and she does not want to risk damaging the relationship for fear of missing out on his delicious treats in the future.

Given CRT, N1 and N2 are reasons that will give Holly's action moral worth, should she be motivated by them. N3 and N4, on the other hand, will not, according to CRT, give Holly's action moral worth. If Holly is motivated only by N1 and N2, each to the right degree, then her action of driving Ira to the train station is maximally morally worthy, because there is perfect overlap between the moral reasons that justify (in Markovits's ought-making sense) her action, and the reasons that motivate her to perform the action. If, on the other hand, Holly is motivated solely by N3 and N4, then her action has no moral worth.¹⁹ These are the extreme ends of the spectrum of moral worthiness; in between are various combinations of motivating reasons that render Holly's action morally worthy to various degrees. For example, if Holly is motivated by N1, N2, and N3, then her action is not maximally morally worthy, though it may have considerable moral worth, depending on the extent to which she's motivated by each reason.

Now we'll focus on a morally-charged situation in which the deliberating agent is not confronted by a reason that plays the moral requiring role. Consider an instance of supererogation:

¹⁹ One might object that Holly's action has at least some (consequentialist, perhaps) moral worth, in that she has come through for Ira, and has delivered him to the station as needed. However, I follow Markovits in restricting talk of moral worth to our evaluations of agents, insofar as we take their actions to be morally praiseworthy.

Holly*: Holly has promised to drive Ira to the airport, but she wakes up Saturday morning feeling sick – sick enough that she’s justified in breaking her promise to Ira.

For the sake of simplicity, let’s suppose that the following normative reasons exhaust the relevant class:

N5) She has promised to drive Ira to the train station.

N6) She is feeling sick.

Since we are supposing that Holly is feeling sick enough that she is justified in refraining from taking Ira to the station, N6 has enough justifying strength to render Holly’s decision to stay home morally permissible. However, again, it’s important to realize that although, given N6, N5 no longer plays the requiring *role* (meaning that it does not generate an all-things-considered moral obligation), N5 retains its requiring strength, because, absent N6, Holly would be required to drive Ira.

Holly is justified in taking Ira to the station in spite of her sickness, but she is also justified in staying home. But since she is justified in doing either, neither is required. If she takes Ira to the station, because of N5, in spite of N6, then her action is supererogatory. It is supererogatory because N5 is an other-regarding reason that Holly is justified in *refraining* from acting upon, given N6. If she stays home because of N6, in spite of N5, then her action is permissible, but not supererogatory, because N6 is not an other-regarding reason.²⁰ Holly’s action of driving Ira, in spite of her sickness, is supererogatory because: (a) it is not morally

²⁰ Of course, this understanding of the case seems to assume that all supererogatory actions involve altruism (or, acting for an “other-regarding” reason). As Jason Kawall (2003) has argued, there may be self-regarding supererogatory actions based on duties to self. I think my description of this case can be applied to cases of self-regarding supererogatory actions, since we might conceive as such actions as sacrifices of relatively immediate (or present self) well-being, for the sake of relatively distant (or future self) well-being. One may be inclined to categorize the duties one has to one’s own future self as a moral, not prudential duty. This strikes me as plausible, but the theorist would need to give a principled account of which self-regarding duties are moral and which are prudential. This need for a principled account will become more apparent later in the paper, when I explain the second step of the moral worth determining process.

required (because the agent is *justified* in not performing it); (b) there is at least one permissible option available to the agent that would involve comparatively less personal sacrifice; and (c) an other-regarding reason rationally justifies her in performing it, in spite of the personal sacrifice involved.

So if Holly chooses to stay home, her action of breaking her promise is morally permissible only because she is sick enough that she is justified in not keeping it. She would not be justified in breaking her promise because she'd prefer to chat with a neighbor, for example. If she breaks her promise because she is sick, then her action is morally permissible. Similarly, if she chooses to take Ira to the station in spite of her sickness, it must be *because* she has promised to do so (i.e. because she is motivated by the relevant moral reason).

5. My Two-Step View of Moral Worth

According to the version of Gert's distinction I'm utilizing, a reason plays the moral requiring role with regard to a specific action when it explains why a conflicting, otherwise morally permissible action is morally forbidden.²¹ For example, going back to the original version of the Holly case, though it is otherwise morally permissible for Holly to sleep in or spend time chatting with a neighbor, it is morally impermissible for her to perform either of these actions when they conflict with the keeping of her promise to Ira. This is because neither of the reasons for these alternative actions can justify the breaking of her promise. In other words, neither reason has *sufficient justifying strength* to render the action of breaking the promise morally permissible. When a morally relevant reason plays the justifying role with regard to a specific moral action, it explains why an action that would otherwise be morally impermissible is morally permissible. For example, though it is otherwise not morally permissible for Holly to break her

²¹ I say "my version" because Gert gives reason to think he would take issue with my explanation of the moral cases. On p. 612 of Gert (2012), he claims that "altruistic reasons typically cannot require very much at all (if anything) though they can often justify a great deal." And in footnote 4, he acknowledges that he takes his worry about Markovits's view to be a worry for at least some versions of moral rationalism.

promise, it is morally permissible for her to do so if she wakes up feeling very sick.

There are two main categories of actions with distinctively moral worth: those that are obligatory and those that are supererogatory. My proposal holds that when a particular action is morally required, because at least one moral reason in favor of it has requiring strength, and there are no existing reasons that have sufficient justifying strength to render the action it speaks in favor of non-obligatory, the action is morally worthy to the extent that the agent is motivated to perform that action for the very reason (or reasons) that plays (or play) the requiring role. Regarding the supererogatory: often, a moral, or other-regarding, action is not required, because, although there is at least one reason with moral requiring strength in favor of the action, there is at least one other reason, e.g. a self-regarding reason, that morally justifies the agent in not performing it.²² It is due to the requiring strength of the existing moral reason that the agent is *rationaly* justified making the relevant sacrifice. Such an action is more morally worthy to the extent that the agent is motivated to take that action for the other-regarding reason that provides requiring strength. Such actions are supererogatory.

Recall, however, that one of the problems with CRT is that it entails that any completely morally worthy action is just as morally worthy as another. Although I've acknowledged a moral version of the justifying/requiring distinction, the account I've offered so far has yet to overcome the worry that one action is as morally worthy as another, so long as each of these actions is

²² For example, the fact that an action will benefit someone has at least some degree of requiring strength, even if such requiring strength is insufficient to override other competing reasons. Suppose S can push one of two buttons (A) and (B), each of which will bring S some equal benefit. Suppose further that if S pushes (A), a group of innocent people in another country will receive malaria immunizations. If S pushes (B), only S will be benefitted. All else being equal, it would be morally impermissible for S to push (B). This shows that absent some justification, the fact that innocent people will be benefitted seems to carry some requiring strength. This example is adapted from Portmore's (2008).

motivated entirely for the other-regarding reasons that have requiring strength.²³ Again, this result is counterintuitive, since it seems that one supererogatory action (e.g. risking one's life to save a stranger) can be supererogatory to a greater degree, and, thus, of greater moral worth, than another supererogatory action (e.g. purchasing a small gift to bring someone joy).

Fortunately, the notion of justifying strength can help us here. Going back to Holly*, suppose that there is some point at which Holly's sickness is severe enough that she is justified in breaking her promise. Any amount of sickness below this point is too weak to justify her in breaking it (e.g. she has a mild headache). But also suppose that there is a point on the sickness scale at which Holly would no longer be justified in keeping her promise. If she is so sick that she is unable to drive safely, she is actually morally required to break her promise. Finally, we can suppose that there is some permissibility space between these two points. The more sickness Holly is experiencing (within the realm of permissibility) the greater the justifying strength of her self-regarding reason in terms of the breaking of her promise. And the greater the justifying strength, the more supererogatory her action of keeping her promise is.

An intuitive implication of this picture is that it can be impermissible to sacrifice too much for the good of another. For example, it would be foolish to risk one's life to procure a child's doll. However, given at least one reason with sufficient justifying strength from the *prudential* deliberative perspective, it can be especially morally worthy to risk one's life or well-being for the good of another.²⁴

²³ Actually, this worry pertains to the moral worth of morally required actions as well, as I will demonstrate below.

²⁴ Here it seems appropriate to point out that the view of moral worth I'm offering here is meant to be localized to a specific deliberative perspective; in this case, the moral deliberative perspective. I take it that the general structure of the view I'm offering can be extended to other localized deliberative perspectives in order to account for other types of supererogation (e.g. epistemic, legal, aesthetic). I intend to pursue such extensions in future papers.

Thus, the degree of justifying strength of competing reasons can account for our intuition that some supererogatory actions performed for all the right reasons are more morally worthy than others. Since risking one's life involves a greater sacrifice than spending a few dollars for a gift, the former action confers even more merit on the agent, in spite of the fact that each action is performed for all of the right reasons. Thus, the former action is supererogatory to a greater degree.

Moreover, the justifying role can do important work for us in determining the degree of moral worth morally required actions can have. Intuitively, although Eva is morally required to refrain from steering her vehicle across the double-yellow into oncoming traffic, and, indeed, the situation is a morally-charged one given the significant harm that will occur if she does, it would be strange to maintain that, in ordinary circumstances, she deserves moral praise for refraining from doing so. However, in life, it is pretty rare for there to be any relevant competing reasons with justifying strength. Now consider that Eva is a thrill-seeker with a burning desire to swerve, and is really struggling to stay on her side of the road. Perhaps there *is* some moral worth in refraining from swerving into oncoming traffic in spite of her urge, so long as she is motivated to refrain for the right reasons. Justifying strength, then, seems to be a *requirement* for moral worth, whether the action in question is morally required or morally supererogatory. A reason with justifying strength is just what is missing in the case of the non-thrill-seeking version of the case. Absent at least one reason with at least some justifying strength, Eva warrants no (or minimal at best) praise for staying on her side.

So I've argued that there are actually two steps involved in determining moral worth: one step involves determining the degree to which the agent's motivating reasons overlap with the requiring strength of the moral reasons that either recommend (if supererogatory), or require the action. In this step, we evaluate the action on its own. However, there's another relevant aspect of moral

worth, and this step opens the door to degrees of supererogation, as well as degrees of moral worth regarding the comparison of distinct actions that are morally required. To clarify this picture, suppose that, following CRT, we determine that a particular agent performs an action, and her motivating reasons coincide with 75% of the moral reasons either recommending or requiring her action. So far, so good, but the work of determining moral worth is not yet complete. We must also consider the justifying strength of her reasons to refrain from performing the action she has chosen.

We can see how this plays out by thinking back to my previous example of two agents who perform distinct supererogatory actions for all of the right reasons. According to the first, Bo thoughtfully purchases a small gift for Dot in order to bring her joy. In the second, Kate donates a kidney in order to prolong the life of her co-worker, Josh. Given the second step I've offered in determining moral worth, we are able to account for our intuition that the second agent's action is supererogatory to a greater degree (i.e. even more morally worthy). The second agent made her decision in the face of reasons containing considerable justifying strength, for example "I may not survive the operation." In the gift-buying case, there is some justifying strength in refraining from purchasing the gift, for example, "It costs fifteen dollars, which I could otherwise spend on myself," but the strength of this reason pales in comparison to the strength of the reason "I may not survive," and this is why, in terms of supererogation, the action itself also pales in comparison.²⁵ As it turns out, an action's moral status (e.g. whether it is obligatory or supererogatory) is independent, at least to some extent, of the degree of moral worth it has.

²⁵ This account sits well with the notion that we learn more about an individual's moral praiseworthiness or character when we evaluate actions performed in scenarios in which it may be difficult to be moral, even though the agent is not actually justified in failing to follow through. Staying on the right side of the road doesn't provide much insight into who one is as a person, but one's ability to withstand significant self-regarding justifying strength in morally-charged situations does.

Naturally, an interesting, but difficult, question arises as to how to go about integrating these two figures: a) the percentage of overlap; and (b) the amount or weight of justifying strength that pulls the agent in the direction of refraining. The view I lean towards has it that we *multiply* the two figures. Here, quickly, is why I tend toward the multiplication method. If an agent is not motivated by any of the right/recommending features of the situation, then the overlap figure is zero, resulting in zero moral worth no matter how weighty the justifying strength of the second figure. Conversely, a required action performed for all the right reasons to the right degree (100% overlap) cannot generate moral worth if the justifying weight figure is zero. These two results, I take it, are intuitively plausible. The former result captures the intuition that an agent who accidentally performs the right action is not morally praiseworthy, as in the example of the evil doctor. The latter result captures our intuition about the Eva case; that the action of driving on the right side of the road, even if performed for all the right reasons, does not warrant moral praise.

According to my Two-Step account, we need not resort to appraiser-relativism in order to account for supererogation, because an action is supererogatory if (a) it is morally permissible (because the agent is justified in performing it, but also justified in not performing it, due to existing reasons with justifying strength); (b) there is at least one permissible option available to the agent that would involve comparatively less personal sacrifice; and (c) an other-regarding reason (rationally) justifies one in performing it, in spite of the sacrifice. According to this characterization, one need not rely on an appraiser who wants to avoid hypocrisy in order to determine whether an action is supererogatory; rather, supererogation is an objective matter. Since there is no need to resort to appraiser-relativism, one is not forced to conclude that supererogatory actions are obligatory (in order to avoid the threat of relativism about obligation). Thus, qualified supererogation is also avoided. Moreover, the plausible aspects of CRT

are retained. Indeed, an agent's action will not end up having moral worth at all unless it is acquired due to such overlap, which is determined in the first step.²⁶

Conclusion

I take myself to have shown that distinguishing the requiring and justifying strength of reasons opens the door a more plausible account of moral worth that avoids appraiser-relativism about supererogation. When a moral action is not required (i.e. an agent is morally justified in not performing it), but an agent performs the action for an other-regarding reason, that action is more morally worthy to the extent that the agent is motivated to take that action for the other-regarding reason that supplies requiring strength, and the degree of overlap is then multiplied by the extent to which the agent would have been justified in refraining from performing it. Justificatory strength enables us differentiate the moral worth of one supererogatory action from another, but it also enables us to differentiate the moral worth of one morally required action from another. According to this new picture, we can also explain the sense in which some morally required actions warrant more moral praise than some supererogatory actions.

²⁶ Another benefit of my account of moral worth is that it is not susceptible to the powerful "gratuitous prevention of goodness" objection to satisficing moral views. Briefly, the objection is that satisficing moral views (the very sort of views that open the door to a non-negligible class of supererogatory actions) seem to warrant an agent in performing an action that confers less benefit on others, even when this agent could confer greater benefit at no (or nearly no) cost to herself. According to my account, an agent must be justified (due to existing morally relevant reasons) in refraining from performing a morally recommended action. Absent competing reasons with sufficient justifying strength, the action would be required. See Bradley's (2006).

References

- Bradley, B. (2006). "Against Satisficing Consequentialism," *Utilitas* 18: 97-108.
- Gert, J. (2007). "Normative Strength and the Balance of Reasons," *Philosophical Review*, 116 (4): 533-62.
- Gert, J. (2012). "Moral Worth, Supererogation, and the Justifying/Requiring Distinction," *Philosophical Review*, 121 (4): 611-618.
- Kawall, J. (2003). "Self-Regarding Supererogatory Actions," *Journal of Social Philosophy*, 34: 487-498.
- Markovits, J. (2010). "Acting for the Right Reasons," *Philosophical Review* 119 (2): 201-242.
- Markovits, J. (2012). "Saints, Heroes, Sages, and Villains," *Philosophical Studies* 158:289-311, p. 290.
- Portmore, D. (2008). "Are Moral Reasons Morally Overriding?" *Ethical Theory and Moral Practice*, 11:369-388.
- Urmson, J.O. (1958). "Saints and Heroes," *Essays in Moral Philosophy*, edited by A.I. Melden, University of Washington Press, Seattle, pp. 198-216.