

02/01/14

Comments welcome.

Feel free to cite.

Morally Permissible Moral Mistakes

by Elizabeth Harman

eharman@princeton.edu

Introduction

In this paper, I argue for a moral category of actions (and omissions) which has been ignored by moral theorists: morally permissible moral mistakes. A moral mistake is something that an agent should not do, all things considered, such that the considerations that tell against doing it – that make it something she should not do – are moral considerations. It is commonly thought that everything we should not do *for moral reasons* – that is, every moral mistake – is morally wrong. But, I argue, this is not true.

I argue that there are morally permissible moral mistakes. I argue that morally permissible moral mistakes can have some features that may be surprising: they do not just arise out of reasons of beneficence; they can be morally good things to do; and they can be praiseworthy.

I distinguish the category of morally permissible moral mistakes from several other moral categories with which it might be confused.

Finally, I argue that recognizing this category is useful in a number of ways. It helps us to see some features of supererogatory behavior that otherwise are ignored. It makes new moral views available to us. It makes new interpretations of our own and others' commitments

possible. And it enables us to make sense of combinations of views that otherwise appear inconsistent.

I. Some moral mistakes are morally permissible

It is intuitively natural to characterize morally wrong actions as follows: morally wrong actions are those actions which (a) one ought not to perform, all things considered, and (b) one ought not to perform *for moral reasons* –that is, the reasons against those actions, which explain why one should not perform them, are moral reasons.¹ More generally, the following two claims appear to be true:

Morality is overriding: If it would be morally wrong for S to ϕ , then S ought not to ϕ , all things considered.²

All moral mistakes are morally wrong: If S ought not to ϕ , all things considered, and the reasons against ϕ ing that win out to make it the case that S ought not to ϕ are *moral reasons*, then it would be morally wrong for S to ϕ .

I believe that morality is overriding, and I will not take issue with that claim.³ In this paper, I will argue that it is false that all moral mistakes are morally wrong. I will explore what follows when we recognize that there are morally permissible moral mistakes.

¹ In this paper, I use “ought to” and “should” interchangeably.

² This claim and the next one are universal generalizations, though I leave the statement of the universal generalization implicit. “S” ranges over agents, and “ ϕ ” ranges over behaviors, which include actions and omissions.

³ A number of different claims have been called the claim that “morality is overriding”. My indented statement specifies what I mean by that terminology.

First, some terminological clarification. Throughout the paper, I use “moral reason” to refer to *other-regarding* moral reasons. It is a hard question whether one is ever morally required to treat oneself well; I will not address this question. If we counted self-regarding reasons among moral reasons, the claim that all moral mistakes are morally wrong might appear obviously false. Often I ought not to do something, all things considered, because it would be bad for me, but it is not morally wrong to do it; if in this case the reasons against the action—that it would be bad for me—count as moral reasons, then we would have too easy a counterexample to the claim that all moral mistakes are morally wrong.

(I think it is an open question whether all moral reasons are other-regarding. As I understand it, a moral reason is any consideration that tells in favor of a behavior being morally required, morally good to do, morally wrong, or morally bad to do. Note that it need not be the case that all reasons that are *morally relevant* are *moral reasons*. For example, the consideration that ϕ ing would inconvenience me, or would be bad for me, may be morally relevant to whether ϕ ing is morally required, without that consideration’s being a moral reason (that consideration may not tell in favor of ϕ ing’s being morally wrong or morally bad, nor tell in favor of refraining from ϕ ing’s being morally required or morally good).⁴)

I will use the terms “mistake” and “moral mistake” as follows:

S’s ϕ ing would be a mistake iff (def)

S ought not to ϕ , all things considered

⁴ Portmore (2008) points out that once we distinguish *reasons that are morally relevant* from *moral reasons*, we will see that, if there are agent-relative moral permissions (which surely there are, I say), then it is not the case that moral reasons are morally overriding, that is, it is not the case that one’s moral reasons alone settle what is morally required.

= (equivalently)

S ought not to ϕ , in light of all of S's reasons⁵

S's ϕ ing would be a moral mistake iff (def)

S ought not to ϕ , all things considered

and

the reasons against S's ϕ ing that win out to make it the case that S ought not to ϕ
are moral reasons

Given these terminological stipulations, the claim that I called "All moral mistakes are morally wrong" does indeed assert that all moral mistakes are morally wrong.

I will now offer a counterexample to the claim that all moral mistakes are morally wrong:

Amanda is a philosophy professor who has a two-year-old daughter. It is 11pm.

Amanda receives an email from her undergraduate student Joe, with a third draft of the paper that is due tomorrow at 12noon. She has already commented on his first two drafts.

Joe is struggling in the class, but she can tell that he is on the verge of some kind of breakthrough. If Joe fails the class, he will lose his scholarship and have to drop out of

⁵ Note that I am using the word "mistake" to apply to *actions (or failures to act)* rather than *thought processes*. When we say that someone has made a mistake, very often we mean that she has reasoned poorly. But we also sometimes refer to actions themselves as mistakes. We might say "it was a mistake to call that poker hand" or "don't make the mistake of paying the carpenter before she finishes the job." A person can perform an action that is a mistake without making any mistake of reasoning, as when someone correctly realizes what she should do but then out of weakness of will fails to do it; one might truly say, "She made a mistake. How did it happen, you ask? She knew what she should do but she was weak-willed."

school. It would take half an hour to read the draft and write the comments, and Amanda is tired. Her daughter will wake her up early. Amanda realizes that she is not morally obligated to spend the thirty minutes to give Joe comments, but nevertheless she deliberates about whether to do it. Upon reflection, Amanda thinks, “I should do it!” She’s right. She gives him the comments.

Here’s my claim about the Amanda case: this could be a true story.⁶ First, Amanda is not morally obligated to give Joe comments. Some professors never give comments to undergraduates on drafts of their papers before the deadline; they are not failing to fulfill any moral obligations. Some give comments on up to one draft of each paper. It is certainly not morally obligatory to give comments on three drafts of a paper, and certainly not when the third draft comes in so close to the deadline. (It would be different if Amanda had said she would give comments on several drafts, and would even give comments at the last minute. Let’s suppose she did not.) Second, while it is not morally required for Amanda to give Joe comments, it may nevertheless be true that all things considered, she should give the comments. Finally, when we

⁶ My claim about this case is that it *could be a true story*: that the normative claims that are part of the case *could be true* of a case with these non-normative features. I don’t want to claim (about this case or the other cases in this paper) that the non-normative features of the case that I lay out make it clear that the normative claims hold in the case. Rather, I just need the weaker claim (about this case and my other cases) that the normative claims *could be true* in *some* case like this.

Thus, my use of described cases in this paper is a bit different from the common use in which a philosopher describes a case and then claims that a certain normative claim *must* be true in the case as described.

ask what reasons there are in favor of Amanda's giving Joe comments, we see that they are *moral reasons*: that Joe would benefit from the comments, and that they might enable him to stay in school. In this case, Amanda's giving Joe comments is *supererogatory*, but still it is the thing she should do. More generally, I claim:

Sometimes a morally supererogatory action is the action that an agent ought to perform, all things considered.

Given that supererogatory actions are not morally required, and given that (often) all they have going for them are moral considerations in favor of them, it follows that it is false that all moral mistakes are morally wrong. (I do need the further assumption that sometimes when a morally supererogatory action is the action an agent ought to perform, this is because of the features of the action that make it morally good and not because of some other features of the action.)

If we plug in *Amanda* for *S* and *failing to give Joe comments* for ϕ , then we get a counterexample to the claim that all moral mistakes are morally wrong: Amanda ought not to fail to give Joe comments; the reasons against failing to give Joe comments (that is, the reasons in favor of giving Joe comments) that win out to make it the case that she shouldn't fail to give him comments are moral reasons; and yet failing to give him comments is not morally wrong.⁷

⁷ The claim that all moral mistakes are morally wrong is implicitly held by most ethicists, I think, including both those who believe that some actions are supererogatory and those who deny that any actions are supererogatory. This paper does not engage the debate regarding whether any actions are supererogatory, though it does take a side in that debate. There clearly are some supererogatory actions; the cases I offer in this paper provide some examples.

While most *moral philosophy* implicitly assumes that all moral mistakes are morally wrong, *ordinary moral thinking* involves implicit commitment to the existence of morally

I also have a more general argument against that claim:

1. Sometimes a morally supererogatory action is the action that an agent ought to perform, all things considered.
2. In some of those cases, all the reasons in favor of the supererogatory action are moral reasons.

Therefore:

3. It is false that all moral mistakes are morally wrong: there are cases in which an agent ought not to ϕ (ϕ = fail to perform some supererogatory act), and the reasons against ϕ ing (that is, the reasons in favor of the supererogatory act) that win out to make it the case that she shouldn't ϕ are moral reasons, and yet ϕ ing is not morally wrong.

I find claim 1 hard to doubt, but I understand that some people might question it. They might say that when it comes to the supererogatory, we are simply in the realm of the permissible—not just the morally permissible, but also the permissible regarding what one ought to do, in light of all one's reasons. As a first pass, an objector might assert this:

Every supererogatory action is permissible, and merely permissible, in light of all one's reasons: that is, it is not the case that one should perform it, all things considered, nor is it the case that one should not perform it, all things considered.

This claim is far too strong to be plausible, however. Importantly, sometimes supererogatory actions are mistakes; their agents should not perform them, given all their reasons. Consider the following case:

permissible moral mistakes. It is common for one person to say to another, “you ought to do it, though you don't have to do it” in a context in which both know that only moral considerations favor doing it, and in which by “you don't have to do it” she means “it is not morally required.”

Betsy suffers from severe migraines, but has learned to avoid them by carefully following certain rules. For example, if she ever gets a mild headache, she lies down with her eyes closed for thirty minutes; otherwise, about 1/3 of the time, a mild headache leads into a migraine. Betsy feels a mild headache coming on when her neighbor Timmy, a six-year-old boy, comes to the door and asks if he could perform a song he's been practicing for her. Betsy knows that Timmy would really appreciate her attention, but also that he will understand if she says she's not feeling well. Listening to Timmy's song, and thus risking a migraine, would be supererogatory. The migraines are really awful, and Timmy won't be too upset if Betsy says "no." All things considered, Betsy should tell Timmy it's not a good time for his song, and go lie down right away.

The Betsy case, as described, could be a true story. Cases like this show that sometimes it is the case that one should not perform a supererogatory act.⁸

So now the objector to claim 1 is left with the following asymmetrical claim:

Some supererogatory actions are such that their agents should not perform them, all things considered, but no supererogatory actions are such that their agents should perform them, all things considered.⁹

⁸ Raz (1975) claims that no supererogatory actions are such that their agents should not perform them.

⁹ An objector to claim 1 could also hold that *all* supererogatory actions are mistakes, that whenever a person has a supererogatory option, all things considered, she should not take that option. It is not true that whenever anyone performs a supererogatory action, she is making a mistake. Nevertheless, some people have committed themselves to this view. For example,

Portmore (2003) has this implication, pointed out by Postow (2005), and repudiated in Portmore (2008).

I also see Dreier (2004) as committed to the view that all supererogatory actions are mistakes. Dreier argues that supererogation is possible because while a certain action might be recommended “from the point of view of beneficence” that point of view ignores certain reasons, such as self-interested reasons, which are nevertheless relevant when the agent considers what to do. On Dreier’s proposal, an action is supererogatory because although it is morally good to do, it would make sense not to do it. It would make sense not to do it because some considerations that are not relevant from the point of view of beneficence outweigh the considerations in favor of doing it. But while Dreier explains how it could make sense to fail to perform a supererogatory action, his account does not explain how it could make sense to perform a supererogatory action. On his account, all supererogatory actions are mistakes.

A further worry about Dreier’s account of how supererogation is possible is that it threatens to prove too much. Dreier distinguishes all three of the following: the perspective of justice, the perspective of beneficence, and the perspective from which one decides what to do, taking all one’s reasons into account. His account of how one might reasonably not perform a supererogatory action threatens to also explain how one might reasonably not perform a morally required action: although that action is required from the perspective of justice, one might have reasons not to perform it that are only relevant from the perspective from which one takes all one’s reasons into account. It is not clear that Dreier’s account of the possibility of the supererogatory can accommodate that morality (that is, moral requirement) is overriding. Dreier is aware of this worry and says that an explanation of why reasons of justice are weightier than reasons of beneficence is needed. (One problem will be that reasons do not divide neatly

It would be odd if this claim were true. I can certainly think of cases when I performed supererogatory actions and I think I was making *the right choice* – not just a justified choice, but a choice that was a better choice than my alternatives, and not morally better but all things considered better (sometimes the *morally best* choice is a *worse choice*, all things considered¹⁰); and, more distressingly, I can think of cases when I failed to perform a particular supererogatory action and I think I made *the wrong choice* – I should have performed that action, though it was not morally required.

Someone might resist my argument by adopting the asymmetrical claim and denying claim 1. Or someone might resist my argument by granting claim 1 but denying claim 2 – holding that it is sometimes true that an agent ought to perform a supererogatory action, all things considered, but only because of non-moral reasons in favor of the action. Those who adopt either of these views can take my paper as arguing for a weaker conclusion: that the view that some moral mistakes are not morally wrong is a *coherent* view that should be taken seriously. This weaker conclusion, as I'll argue, is important in its own right.

Let me summarize the disagreement between the view I am arguing for and the common view that I am opposing. Figures 1 and 2 may help to illustrate the difference. (They are on pages 13 and 14.)

between the two categories. The very same reason may be the source of a moral requirement in one situation but merely a reason to perform a supererogatory action in another.)

¹⁰ See part b of section IV for more discussion on this point.

Both views agree that morality is overriding, so both hold that if a behavior is morally required, then all things considered one should behave that way; and if a behavior is morally wrong, then all things considered one should not behave that way.

Both view agree that behaviors that are morally permissible without being morally required can fall into all of the following three categories: something one should not do, all things considered; something one should do, all things considered; and something that is neither such that one should nor should not do it. Both views agree that morally permissible behaviors can fall into all three of these categories because of *non-moral reasons* in favor of or against these behaviors. Suppose I am playing a low-stakes game of poker, in which it is morally permissible to either call the current bet or fold; nothing morally significant hangs on what I do. It may be that I *should* call the bet, all things considered, though it is not morally required to call the bet. In this case, I *should not* fold, all things considered, though it is not morally wrong to fold my hand.

The two views disagree regarding the following category of behaviors: moral mistakes, that is, those behaviors that should not be engaged in, all things considered, *for moral reasons*. The common view holds that this category is the same category as: morally wrong behaviors. I am arguing that this is incorrect. This category really encompasses two moral categories: morally wrong behaviors and morally permissible moral mistakes.

It follows that the two pictures also disagree regarding the following category of behaviors: those behaviors that should be engaged in, all things considered, *for moral reasons*. The common view holds that this category is the same category as: morally required behaviors. I am arguing that this is incorrect. This category really encompasses two moral categories:

morally required behaviors and behaviors that it would be a morally permissible moral mistake not to engage in.

My argument that there are morally permissible moral mistakes proceeds by pointing out that some supererogatory actions have the following two features: all things considered, they should be performed; and the reasons they should be performed are the very *moral reasons* that make them morally good things to do. Nevertheless, these actions are supererogatory; they are not morally required. These supererogatory actions fall into the second-to-last box in Figure 2: behavior it would be a morally permissible moral mistake not to engage in. *Failures to perform these supererogatory actions* fall into the second box in Figure 2: they are morally permissible moral mistakes.

(Note that I will also use the term “mere moral mistakes” for “morally permissible moral mistakes”, later in this paper.)

Figure 1. The Common View I am Opposing

The common view I am opposing holds that every way a person could behave falls into one of the categories in these five boxes.

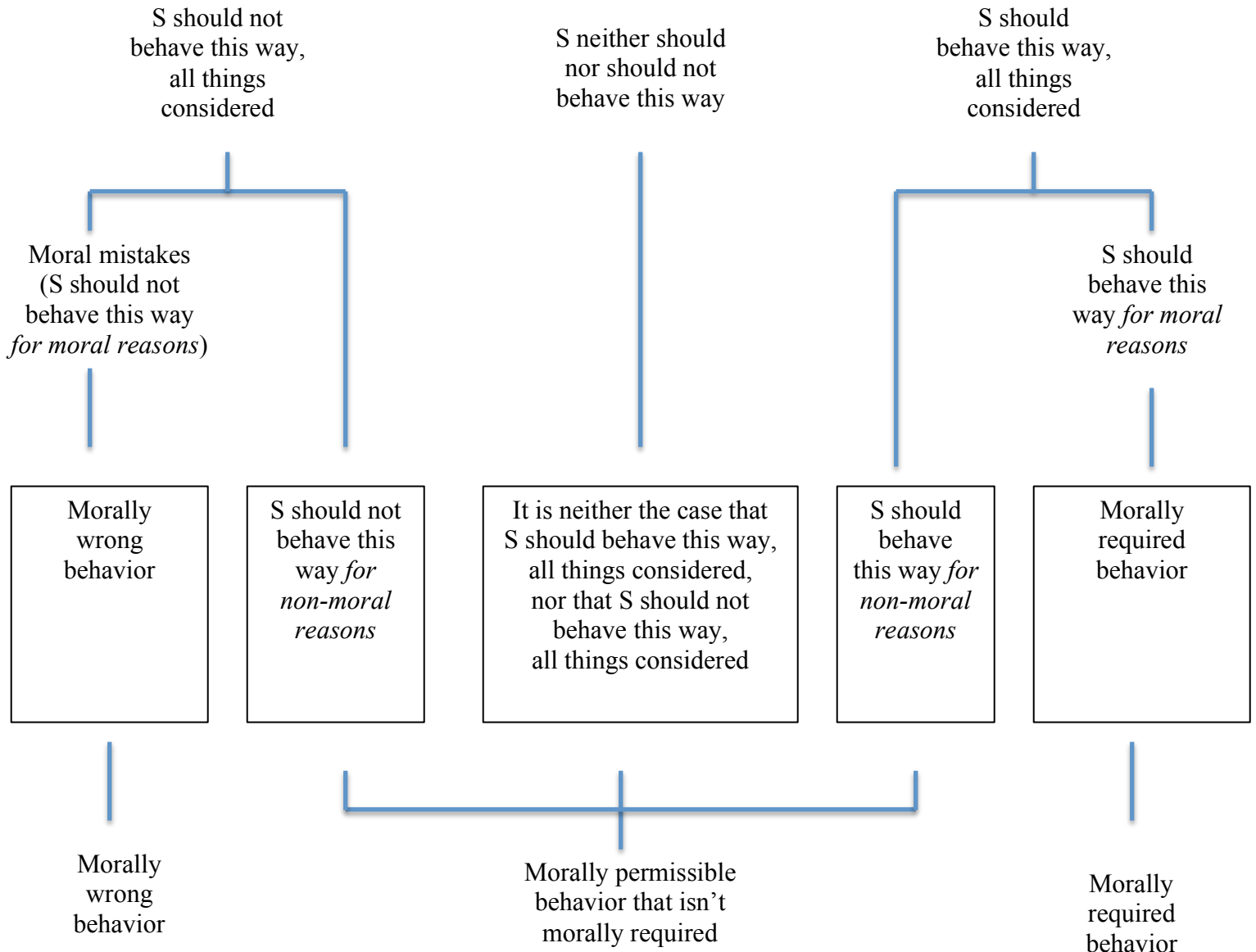
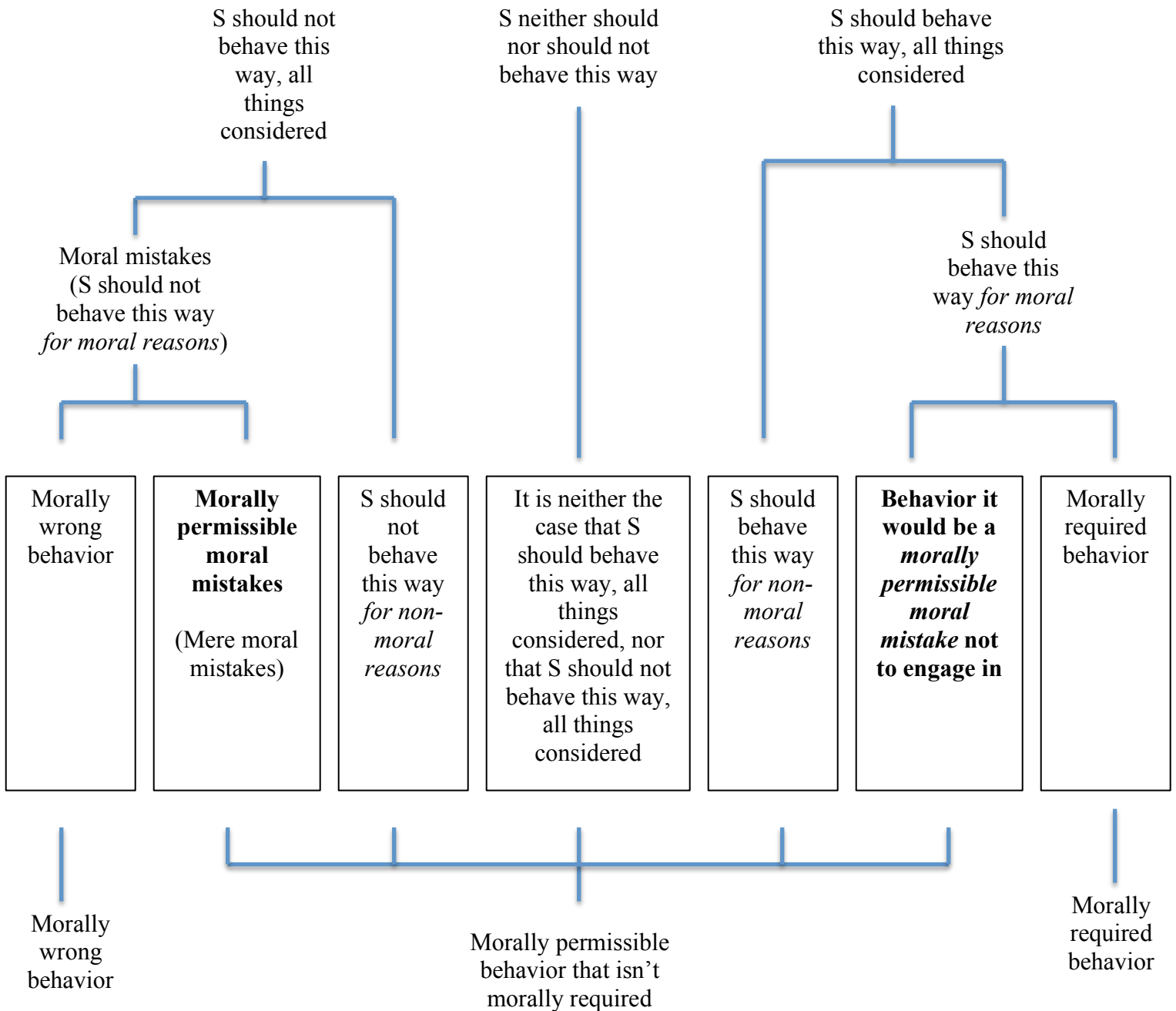


Figure 2. The Correct View

The correct view holds that every way a person could behave falls into one of the categories in these *seven* boxes. The two added categories are in bold.



II. Morally permissible moral mistakes can arise from moral constraints

I've argued against the claim that all moral mistakes are morally wrong. Someone sympathetic to that claim might try to rescue some of its spirit by saying the following:

Some of our moral reasons are reasons of beneficence; they can make an action morally good without making it morally required. But other moral reasons are in their nature morally constraining. They cannot make an action good without making it morally required, nor can they make an action bad without making it morally wrong. Our constraint-based moral reasons morally require whenever they require at all.

Such a thought motivates the following claim:

All constraint-based moral mistakes are morally wrong: If S ought not to ϕ , all things considered, and the reasons against ϕ ing that win out to make it the case that S ought not to ϕ are constraint-based moral reasons (such as reasons not to lie, not to kill, and to keep one's promises), then it would be morally wrong for S to ϕ .

This claim is not true, however. Consider the following case.

Cara has promised to pick Sam up at the airport. It was important to him to get Cara to *promise* because Sam is a nervous flier who feels very anxious and strung out after traveling, and it would make him feel a lot better to see a friendly face at the other end of his travels. Unforeseeably, Cara's car breaks down. The mechanic can fix the car in two days for \$100 or in two hours for \$500. This means it would cost an extra \$400 to be able to pick Sam up. Cara realizes that it would be morally permissible not to pick Sam up under the circumstances, and that Sam would understand that too. Nevertheless, Cara continues to deliberate. Upon reflection, she concludes, "I should still pick Sam up, because I promised." She is right. She picks him up.

Here's what I claim about the Cara case: this could be a true story. It could be that Cara ought to pick Sam up, all things considered, because she promised to pick him up, even though she is not morally required to pick him up, because there is an unexpected significant cost to picking him up. In this case, if Cara were to fail to pick Sam up, she would be making a *constraint-based* morally permissible moral mistake.

In saying that the Cara case could be a true story, I do not mean to suggest that it is always the case that we should keep our promises; of course that is not true. My claims are just these:

It is not unreasonable to continue to deliberate about whether to keep one's promise, even after realizing that circumstances have changed so that one is not morally obligated to keep one's promise.

In the course of such deliberation, one might conclude "I should still keep my promise" and one might be right about that.¹¹

In some such cases, one should keep one's promise *because one promised* and not for some other reason.¹²

¹¹ In "Morality with the Realm of the Morally Permissible" (manuscript), I argue that looking closely at deliberation about whether to perform some supererogatory actions shows that there are some morally permissible moral mistakes.

¹² The fact that some constraint-based moral mistakes are not morally wrong is related to the fact that some supererogatory actions are not acts of beneficence. This is a case in which keeping one's promise would be supererogatory, but it's not an act of beneficence.

III. Some moral mistakes are morally good and praiseworthy things to do

I've argued that not all moral mistakes are morally wrong, and I've argued that not all constraint-based moral mistakes are morally wrong. Someone might try to rescue something of the spirit of the original claim by saying the following:

*At least we can say this about moral mistakes: they are never morally good actions; and they are never praiseworthy (if the agent knows what she is doing).*¹³

Here is the claim under consideration:

No moral mistakes are morally good actions or praiseworthy actions: If S ought not to ϕ , all things considered, and the reasons against ϕ ing that win out to make it the case that S ought not to ϕ are *moral reasons*, then it is not the case that S's ϕ ing would be a morally good action, and it is not the case that S would be praiseworthy for ϕ ing (if S knew the nonmoral facts that make ϕ ing something S ought not to do).

This claim is false. We can see that it is false by considering a case in which an agent ought to perform a particular supererogatory action, all things considered, but she also has another

¹³ We can have an uninteresting case of a moral mistake that is praiseworthy if we consider a case in which someone does something objectively wrong (so it is a moral mistake) but she doesn't know what she's doing: a woman gives her husband what she has every reason to believe is the cure to his illness but it's actually poison. She poisons her husband (that's a moral mistake) but she is praiseworthy (she tries to save his life). An interesting case would be one in which an action is a moral mistake due to non-moral features of the case that the agent knows about, but nevertheless the agent is praiseworthy for acting. That is the kind of case I will claim is possible.

supererogatory action available to her, which is less morally good than the one she ought to perform. Consider this case, which is similar to the Amanda case we saw earlier:

Deborah is a professor with a two-year old child. At 11pm, two students email drafts to Deborah. Jim sends his third draft; she's already given him comments on two drafts; he is struggling but nearing a breakthrough. If Jim fails the class, he will lose his scholarship and have to drop out of school. It would take her thirty minutes to give him comments. Tom also sends a draft; it's a second draft, with one new section that she has hasn't seen; the first draft was already reasonably good. Tom would benefit from comments, and it would only take her ten minutes to give him comments, but he'd do fine without comments too. Deborah definitely shouldn't give them both comments. She needs sleep. She's not morally required to give comments to either one. Deborah deliberates about what to do. In fact, she should sacrifice thirty minutes of sleep to give Joe comments. But that's not what she decides to do; she spends ten minutes giving Tom comments instead.

(Again, my claim is that this is a *possible* case.) Deborah makes a mistake: she does something that she should not do, all things considered. (She comments on Tom's paper rather than Jim's.) Her mistake is a moral mistake; the reasons against doing what she does, that make it something she should not do, are *moral reasons*—they are the reasons in favor of commenting on Jim's paper instead. But what Deborah does is a morally good thing: she gives Tom comments and helps his paper to be better even though she's already given him comments on one draft, at some real cost to herself, the cost of some sleep. What Deborah does is praiseworthy. It's not as praiseworthy as commenting on Jim would have been, but still it is praiseworthy.¹⁴

¹⁴ The existence of these “second-best” supererogatory actions, such as Deborah's commenting

IV. Other Moral Categories

Some moral mistakes are morally wrong; I have argued that some moral mistakes are not morally wrong. Let me introduce some more terminology:

S's ϕ ing is a mere moral mistake iff (def)

S's ϕ ing is a morally permissible moral mistake

All morally wrong actions are moral mistakes. A mere moral mistake is *merely* a moral mistake in that it is a moral mistake but (unlike some moral mistakes) it is not also morally wrong.

Now that we have seen that there are morally permissible moral mistakes, we can recognize some interconnected distinctions. The following distinctions are *inclusive*; the first category is a subset of the second category:

morally wrong actions *versus* moral mistakes

morally required actions *versus* actions an agent ought to perform, all things considered,
for moral reasons

The following distinctions are *exclusive*; no behavior falls into both categories:

morally wrong actions *versus* mere moral mistakes

morally required actions *versus* actions an agent ought to perform, all things considered,
for moral reasons, but which are not morally required

on Tom's paper, is underappreciated in discussions of the supererogatory. For example. Horgan and Timmons (2010) gloss the paradox of supererogation as arising because supererogatory actions are "morally best" (29). Similarly, Dreier (2004) proposes that supererogatory actions are those that are required from the perspective of beneficence, and Portmore (2008) says "for there is a sense in which supererogatory acts are acts that agents morally ought to perform" (379).

To better understand the category of *mere moral mistakes* (that is, morally permissible moral mistakes) and the four distinctions above that the category gives rise to, I will now examine some well-recognized moral distinctions and discuss whether these distinctions enable us to characterize mere moral mistakes. If so, then we have already been talking about mere moral mistakes in talking about these other categories. If not, then the category of mere moral mistakes is not already captured by these distinctions. The discussion will also help to clarify what mere moral mistakes are.

a. What is Objectively Wrong *versus* What is Subjectively Wrong

Someone might think that we can see an instance of a mere moral mistake by considering cases in which an agent has false beliefs about her situation. A wife poisons her husband, thinking she is giving him the cure; she is blameless for her false belief. In this case, what she does is objectively wrong (she poisons him); but it is not subjectively wrong (she gives him what she has every reason to believe is a cure).

Proposal A: The wife makes a mere moral mistake: she does something she should not do, for moral reasons (poison her husband), but what she does is morally permissible (she gives him what she has every reason to believe is a cure).

The proposal fails. The wife does not make a mere moral mistake, because the normative statements above are, we might say, true only relative to different bodies of evidence – that she should not do what she does is true relative to her actual situation, ignoring her evidence; but that it is permissible to do what she does is true relative to her evidence. When an action is a mere moral mistake, that is so because *it is something the agent should not do for moral reasons* and yet *it is morally permissible* – where these normative statements are true *relative to the same*

*body of evidence.*¹⁵ There can be objective mere moral mistakes and there can be subjective mere moral mistakes. But one cannot identify something as a mere moral mistake by exploiting both objective and subjective moral claims.

If we consider the case of Amanda, who has no non-moral false beliefs about her situation, had she failed to give Joe comments, that would have been an objective moral mistake and also a subjective moral mistake. If we consider a variant of the case in which, unbeknownst to Amanda, she's going to be woken up at 4am by a prearranged test of the fire alarm system in her apartment building, that is a case in which, objectively, she ought to go to sleep right away (the impending fire alarm makes sleep much more urgent), and failing to give Joe comments is not a mistake, objectively. But if in this case, she fails to give Joe comments, she does make a subjective moral mistake.

b. What is morally best *versus* what is morally required

Consider the following proposal:

Proposal B: What is morally best *is the same as* what one ought to do, all things considered, for moral reasons. S's ϕ ing is a mere moral mistake just in case S's

¹⁵ By speaking of moral claims being true relative to bodies of evidence, I don't intend to make any controversial claims about how we should understand objective and subjective uses of normative expressions. My intention is to grant to the Proposal that all the claims it makes, about what the agent ought to do and what is morally required, have *true readings*, while explaining that the Proposal fails. (There is quite a rich literature on objective and subjective uses of normative expressions, including the following: Dowell (2013); Jackson (1991); Kolodny and MacFarlane (manuscript); MacFarlane (2014); Smith (manuscript).)

refraining from ϕ ing is morally best though S's refraining from ϕ ing is not morally required.

This proposal fails because often the morally best option available to an agent is not the one that she should take, all things considered. The morally best thing I could do right now might be this: go to the local hospital and offer up a kidney and some of my liver, to save two lives. But that's not what I should do right now, all things considered. So I am not making a mere moral mistake in failing to do so.

c. The morally right thing to do *versus* the morally required thing to do

One might make the following proposal:

Proposal C: S's ϕ ing is a mere moral mistake just in case S's refraining from ϕ ing is the morally right thing to do though S's refraining from ϕ ing is not morally required.

This proposal uses the expression "the morally right thing to do" to pick out behavior one should perform, all things considered, for moral reasons, including both morally required behavior and behavior it would be a mere moral mistake not to engage in. This strikes me as a good way to use the expression "the morally right thing to do," but I don't think that talk of "the morally right thing to do" is generally used in this way. Often people use "the morally right thing to do" to pick out only actions that are morally required. I think it would be good if talk of "the morally right thing to do" were reformed in line with Proposal C.¹⁶

¹⁶ Usage of "morally right" varies. Often people also use "morally right" to mean "morally permissible": when they say "that is the morally right thing to do" they pick out a morally required action by using the definite article ("that is the (only) morally permissible thing to do") and when they say "that isn't morally right" they mean "that isn't morally permissible." (Thanks to Jamie Dreier for discussion of this point.) Another usage allows "morally right" to cover any

People also talk simply of “the right thing to do”; this means the same as “what one should do, all things considered.” The expression “the right thing to do” doesn’t just pick out actions that are favored by moral considerations, so the following alternative proposal also fails: S’s ϕ ing is a mere moral mistake just in case S’s refraining from ϕ ing is *the right thing* to do though S’s refraining from ϕ ing is not morally required. This alternative proposal would hold that my failing to call a hand in poker (when I should have called) was a mere moral mistake; but it was a mistake without being a moral mistake.

d. What one ought to do *versus* what one must do

Proposal D: S’s ϕ ing is a mere moral mistake just in case S ought to ϕ but it’s not the case that S must ϕ .

This proposal fails because the ought/must distinction applies even to non-moral cases. If you are coaching me while I play online chess, there may be certain moves that I *must* make; it would be a serious mistake not to make them. Other moves are such that I ought to make them, but I don’t have to make them: it’s a mistake not to make them, but not a serious mistake.¹⁷

morally good action, including both morally required and supererogatory actions (excluding merely morally permissible but morally neutral actions); for example, this is Sliwa’s use in (manuscript). No one to my knowledge has used “morally right” in the way I suggest in the main text. (Sliwa’s use is different because it covers all supererogatory actions, even those that one should not perform because one should perform a different supererogatory action instead.)

¹⁷ We might consider a similar proposal regarding “morally ought” and “morally must”. My response to this proposal is similar to my response to Proposal C. It might be a good idea to use “what one morally ought to do” to pick out exactly those ways that one should behave, all things considered, for moral reasons; but that is not how that expression is currently used.

e. The suberogatory

Suberogatory actions are actions that it is bad to do, but not wrong to do. (They are a kind of inverse of the supererogatory, which are good to do, but not morally required.)¹⁸

Proposal E: All and only mere moral mistakes are suberogatory.

The proposal is false because some mere moral mistakes are not suberogatory. If Amanda fails to give Joe comments, she's failing to do something good; but she's not doing anything bad. But a weaker claim is true: all suberogatory actions are mere moral mistakes.

f. What it would be indecent to do versus what it would be morally wrong to do

Judith Jarvis Thomson, in her famous paper "A Defense of Abortion," seems to hold that certain behavior is indecent though not morally wrong.¹⁹ For example, if you are a stranger to me and in order to continue to live, all you need is the cool touch of my hand on your forehead, and I must merely walk across the room to touch you, Thomson's view appears to be that since you lack a positive right to my doing this, it is not morally required that I do this; but it would be indecent for me not to do it. Thomson is wrong about this case. If I can save a stranger's life by walking across the room to lay my hand on his forehead, and if I am unlikely to often find myself in similar situations, then I am morally required to save his life. Rather than showing that some behavior is indecent without being morally wrong, I think this shows that some behavior is morally wrong though it violates no one's rights.²⁰

But there may be actions which are indecent without being morally wrong.

¹⁸ See Driver (1992). Liberto (2012) argues that there are no suberogatory actions.

¹⁹ Thomson (1971).

²⁰ For interesting arguments that some behavior can be morally wrong – and indeed can wrong someone – without violating anyone's rights, see Cornell (dissertation).

Proposal F: All and only mere moral mistakes are indecent but not morally wrong.

If there are actions which are indecent without being morally wrong, these actions would fall under the category of the suberogatory, and my comments above apply to them. Not all moral mistakes are suberogatory, so not all moral mistakes are indecent. Yet actions that are indecent without being morally wrong are all mere moral mistakes.

g. What is justified by reasons versus what is required by reasons

Joshua Gert has argued that philosophers overlook a certain feature of reasons, which is that reasons have two dimensions of strength, not one. On the one hand, reasons have a certain requiring strength; on the other hand, they have a certain justifying strength.²¹ A reason's requiring strength is sometimes—indeed, often—weaker than its justifying strength. Consider the fact that Mary's daughter would be thrilled if Mary baked cupcakes today. This fact has no requiring strength, I claim. Yet it has some justifying strength; it could make it reasonable for Mary to bake cupcakes today, even if Mary dislikes baking.

Proposal G: S's ϕ ing is a mere moral mistake just in case S's refraining from ϕ ing is justified but not required by moral reasons.

This proposal fails because an action might be justified in Gert's sense (it might be a reasonable thing to do) without being the action that the agent ought to perform, all things considered. The notion of justification cannot be substituted for all-things-considered requirement. Gert is working with only one kind of requirement, but to understand the phenomena I have been

²¹ Gert (2012).

describing, we need two kinds of requirement: there is what is morally required; and there is what an agent ought to do, all things considered.²²

h. Perfect Duties versus Imperfect Duties

Consider the following proposal:

Proposal H: S's ϕ ing is a mere moral mistake just in case S has an imperfect duty to refrain from ϕ ing but S lacks a perfect duty to refrain from ϕ ing.

This proposal fails because there are many actions that would satisfy an imperfect duty, and so it is true that one has an imperfect duty to perform them, but that are not such that one should perform them, all things considered. Consider the case of Betsy, who should go lie down to stave off a possible migraine rather than listen to the boy at her door. Listening to him would fulfill an imperfect duty; but it is not the case that she should listen.²³

²² Horgan and Timmons (2010) argue that reasons can function in three ways: they can require, justify, or "confer merit" on an action. We might consider the proposal that failing to ϕ is a moral mistake just in case ϕ ing would have merit. This proposal fails because an action might have merit without being the thing that an agent should do.

²³ Horgan and Timmons (2010) argue that a supererogatory act involving great self-sacrifice cannot fulfill an imperfect duty, so they might deny that Betsy's listening to the boy would fulfill an imperfect duty. They say that the duty of beneficence does not "call for" great self-sacrifice, *therefore* instances of great self-sacrifice do not "fulfill" this duty. But they are wrong to think that a duty cannot be fulfilled by an action that it does not call for. If I promise to give you some food for helping me move, and if buying you some pizza would have fulfilled my promise, but instead I cook you a gourmet meal, then I do fulfill my duty to you. Doing more than is called for can be a way of fulfilling a duty.

V. Objection: Moral Reasons always Morally Require

In this section, I will consider an objection to my argument that some moral mistakes are not morally wrong. The objector hopes to vindicate this claim:

All moral mistakes are morally wrong: If S ought not to ϕ , all things considered, and the reasons against ϕ ing that win out to make it the case that S ought not to ϕ are *moral reasons*, then it would be morally wrong for S to ϕ .

The objector asserts:

Moral reasons are such that, if they win out, then they make an action morally wrong. If an other-regarding reason wins out but the action in question is not morally wrong, then the other-regarding reason is not a *moral reason*.

The objector grants much of what I said about the cases of Amanda (who is considering giving Joe comments on a paper) and Cara (who promised to pick Sam up at the airport): it is possible for there to be cases in which it is morally permissible for Amanda to go straight to sleep and for Cara to fail to pick Sam up, but in which all things considered, Amanda should give Joe comments, and all things considered, Cara should pick Sam up. The objector denies, however, that in these cases the reasons in favor of commenting on Joe's paper and picking Sam up, which win out to make it the case that the agents should do these things, are *moral reasons*.

I will consider two versions of this objection.

The first objector is using "moral reason" stipulatively. The objector asserts that he can use the term "moral reason" so that it only applies to reasons that win out against an action when the actions in question are thereby morally wrong. There are some problems with this strategy. What will this objector say about moral reasons *against* actions that we all things considered *should* perform? Consider a case in which an agent must save one person or ten, and all else is

equal. She should save the ten. But there is a moral reason to save the one; yet it does not make saving the one morally required. The proponent of this terminological proposal could either deny that losing moral reasons are moral reasons or say that moral reasons are such that *if they win out*, they make actions morally wrong. The former option greatly strains against ordinary usage of “moral reason”; the latter option raises the question: why are losing moral reasons more like winning moral reasons *in cases where the action is morally wrong* and less like winning other-regarding (apparently moral) reasons *in cases where the action is merely what the agent ought to do, but not morally required*?

But none of that really matters. My response to the first objector’s proposal is this: maybe one can stipulate a use of “moral reason” as the objector suggests. What I need for my purposes is that I have invoked a sufficiently ordinary sense of “moral reason” such that it is clear what I am saying, the claim that all moral mistakes are morally wrong has some intuitive appeal, and yet upon seeing my arguments, we can see that it is false. What I mean by “moral reason” is what we ordinarily mean by “moral reason”: an other-regarding consideration that tells in favor of an action’s being morally required, or a morally good thing to do, or morally wrong, or a morally bad thing to do.

Let me turn to the second version of the objection. This objector goes along with me in using “moral reason” in the ordinary way, but makes a substantive claim. The substantive claim is what when moral reasons against an action win out, the action is then morally wrong; when other-regarding reasons against an action win out, but the action is not morally wrong, these other-regarding reasons are not moral reasons. The objector has agreed that there are possible cases in which Amanda should give Joe comments, all things considered, and Cara should pick Sam up, all things considered, but neither action is morally required. The objector holds that in

these cases, the considerations in favor of giving Joe comments and picking Sam up *are not moral reasons*.

We can see that this claim is false by considering pairs of similar cases, in one of which behaving in a certain way is morally wrong and in the other of which it is not, where the reasons against the action are the same. Compare a case in which Cara has promised to pick Sam up at the airport, and then nothing in particular happens to make this more difficult, to the case of Cara outlined above, in which it would cost four hundred extra dollars to pick Sam up, and yet, all things considered, Cara should pick Sam up. My claim is that the very same consideration—that *Cara promised to pick Sam up*—wins out in both cases, but that it makes failing to pick Sam up morally wrong only in the first case. This consideration either is a moral reason in both cases or it isn't a moral reason in either case. It is clearly a moral reason in the first case, in which nothing happens to make it harder to pick Sam up; so it is a moral reason in the second case too.

VI. New Moral Understandings

Recognizing that there are such things as mere moral mistakes can serve at least three purposes in our moral theorizing. First, it can dissolve a “puzzle of the supererogatory”.

Consider the following “puzzle of the supererogatory”.²⁴ On the one hand, supererogatory actions are not required—by definition, they are above and beyond the call of

²⁴ What I am calling “a puzzle of supererogation” is not the same as what is sometimes called “the paradox of supererogation”, though they are related. The “paradox of supererogation” asks how there can be actions that are favored by moral reasons but not morally required. (The paradox is discussed in Raz (1975), Dreier (2004), Horgan and Timmons (2010), and others.) The puzzle I am discussing asks how someone can take her own action to be favored by moral reasons without also taking it to be morally required; this puzzle is raised in Urmson (1958).

duty. On the other hand, people sometimes do perform supererogatory actions. What can be going on when someone chooses to perform a supererogatory action? Since she is performing the action, she must believe it is *the thing to do*, and so she must believe it is required. Thus, it appears that it is impossible for anyone to perform a supererogatory action *while understanding that it is supererogatory*. That is the puzzle. We might add on: Could there really be a category of action that is above and beyond the call of duty but such that no one who performs it knows that it is above and beyond the call of duty? This may suggest that there are no supererogatory actions. Also: It's natural to think that there is a special kind of praiseworthiness that attaches to supererogatory actions. But can someone be praiseworthy in that special way if she thinks that she is just doing what is required? This may suggest that no one is ever praiseworthy in the special way that seems to attach to the supererogatory.

The statement of the puzzle gets things wrong in at least two ways. First, one *could* choose to do something without thinking it is *the thing to do*, for example when one chooses one of several identical soda cans from the refrigerator. I think there is a hidden assumption that when one chooses to perform a supererogatory action, one is surely not taking that action to be *on a par with* one's non-supererogatory options; indeed, one takes one's action to be *the thing to do*. Let's grant that assumption (for now). The more significant flaw in the reasoning in the development of the puzzle is this: there are two kinds of *requirement* that are relevant. One is *moral requirement*; the other is *requirement in light of all one's reasons*. To put this another way: an agent who performs a supererogatory action could take what she is doing to be the thing she should do, in light of all her reasons, without taking it to be morally required.

(We can see what I have said as offering a solution to the puzzle on the puzzle's own terms – granting the assumption that an agent who performs a supererogatory action always takes

her action to be *the thing to do*. But that assumption is surely too strong. Some people who perform supererogatory actions do not take those actions to be *the thing to do*. There must be cases in which a supererogatory action is on a par with one or more non-supererogatory options; an agent who performs one of these actions may well recognize this, and choose the action without taking it to be the thing to do.)

Recognizing the existence of mere moral mistakes can also provide a second, third, and fourth benefit in our moral reasoning. It makes new views available to us. It makes it possible for us to interpret our own views or others' existing views in new ways. And it may help us to make sense of combinations of view that otherwise seem inconsistent.

In fact, all four of these benefits are available even to those who are not convinced by my arguments for *the existence* of mere moral mistakes. My paper can also be seen to argue for the weaker claim that it is *conceptually coherent* to suppose that there are some mere moral mistakes, and thus to suppose that particular actions or types of action are mere moral mistakes. By seeing that this is conceptually coherent, we can recognize the availability of certain interpretations of existing views, and the possible coherence of certain combinations of views.

New views are made available:

I have recently developed the following view about the ethics of being an anonymous donor of sperm or eggs: this is a morally good thing to do, and is praiseworthy, but all things considered, one should not do it, because of moral reasons one has regarding one's children (including one's genetic children); being a gamete donor is a morally good, morally praiseworthy, morally permissible moral mistake. Our reasons regarding our own children include reasons to be available to them when they need our help, and reasons to have real relationships with them; anonymous donors set themselves up not to do the things these reasons

support.²⁵ I will not defend this view here. But the view captures a number of different things that each have some intuitive appeal. My arguments in this paper—for the existence and coherence of mere moral mistakes—shows that these views can actually be held together.

More generally, the category of mere moral mistakes includes those actions that have serious moral things to be said against them, but are not morally wrong. We must be careful to consider, when we find a behavior problematic for moral reasons, that it might be merely a moral mistake rather than morally wrong. I think that when moral objections to a practice strike us as weighty, we fail to recognize that there are two different ways moral objections to a practice can win out: they may make the practice morally wrong, or they may make it merely a moral mistake.

This brings me to the third benefit of taking seriously the category of mere moral mistakes: we can interpret our own views or others as implicitly committed to the view that certain kinds of action are mere moral mistakes. I will give two examples.

Consider the view that one should not buy or eat meat, and yet one should not criticize or impede the meat-purchasing and meat-eating practices of others. It is striking that vegetarians often seem to have this view: they keep a strict practice of avoiding meat themselves, for moral reasons; and yet they are willing to support restaurants that serve meat, and are willing to eat alongside others eating meat, saying nothing at all in criticism. It is not generally true that it is morally fine to fail to criticize or impede in any way the morally wrong behavior of others. The practices of these vegetarians makes perfect sense, however, if we attribute to them the following implicit belief: that buying and eating meat is a mere moral mistake. They might even think that buying and eating meat is *suberogatory* – that it is bad to do, though not morally wrong. Yet still

²⁵ I develop this view in “Gamete Donation as a Laudable Moral Mistake” (manuscript).

appropriate responses to suberogatory actions (which are, after all, morally permissible) vary greatly from appropriate response to morally wrong actions.²⁶

Consider next the view that one should not have an abortion, for moral reasons, and yet that abortion should not be illegal.²⁷ One way to make sense of this view is that those who hold it are implicitly committed to the view that abortion is a mere moral mistake.²⁸

VII. Further Considerations in Favor of the Existence of Mere Moral Mistakes

Our moral lives include a substantial realm of moral permissibility. Each of us has her own projects and her own interests, as well as her own loved ones, and we have agent-relative permissions to privilege our projects and our interests in substantial ways. Nevertheless, we often have reasons regarding others which tell in the other direction; these include reasons of

²⁶ I develop this view about vegetarianism and accommodation of meat eating in “Eating Meat as a Morally Permissible Moral Mistake” (manuscript).

²⁷ Thanks to Kirsten Egerstrom for asking me about this.

²⁸ My argument for the existence of mere moral mistakes proceeded by arguing that there can be *individual cases* in which a certain kind of action is a mere moral mistake. In this section, I have pointed out that once we recognize the *coherence* and *possibility* of mere moral mistakes, we can consider views according to which a certain kind of action is, in general, a mere moral mistake. Someone might claim that mere moral mistakes only arise in isolated instances, but that no kind of action is in general a mere moral mistake. I am not sure what could be said in support of that claim. In any case, my main goal has been to argue that there are some mere moral mistakes—or at least for the weaker claim that mere moral mistakes are coherent. Once we recognize the coherence of the concept of a mere moral mistake, that leaves it open for people to hold that particular kinds of action are such that each action of that kind is a mere moral mistake.

beneficence (but not only those). We regularly face questions of whether to privilege ourselves over others, even in cases in which it is permissible to privilege ourselves. It makes sense to deliberate about what to do in these cases, even though we know that we are not required to follow our other-regarding reasons: our other-regarding reasons continue to exert a force on us even though they do not in these cases render the behavior they favor morally required. It would be wrong to think that in these cases, all one's options are on a par. Sometimes doing a supererogatory thing is a mistake; the case of Betsy from section I is one example. Sometimes failing to do a particular supererogatory thing is a mistake: just as one's self-regarding reasons may win out in some of these cases, similarly one's other-regarding reasons may win out in some of these cases.

VIII. The Significance of Mere Moral Mistakes

One might ask: what is the significance, or upshot, of recognizing that a particular action (or each typical action of a particular kind) is a mere moral mistake? The answer here is subtle. On the one hand, mere moral mistakes are not morally wrong. This means, for those actions that we might have been tempted to think are morally wrong, that if they are mere moral mistakes then agents who perform those actions (knowing what they are doing²⁹) are not blameworthy in the way one is blameworthy for knowingly doing something that is in fact morally wrong. So these actions are in a sense less morally bad than morally wrong actions, in that they are not blameworthy in the way that morally wrong actions are blameworthy. But these actions are not as morally innocent as other morally permissible actions. Indeed they are not as innocent as many failures to perform supererogatory actions. Many failures to perform supererogatory actions are merely failures to do something that would have been morally good, but these failures

²⁹ That is, knowing the non-moral features of their situations.

are not *mistakes*: it is often not the case that, all things considered, one should have performed the supererogatory action in question. Mere moral mistakes are *mistakes*: they are failures to act as one all things considered should act, and they are failures to respond to one's *moral reasons* appropriately. These are moral failings.

To bring out the significance of an action's being a mere moral mistake, let's suppose that each instance of an individual's eating meat is a mere moral mistake. Some people think eating meat is morally wrong; given our supposition, these people are *overstating* the moral failing involved in eating meat. Some people think that eating meat is morally permissible, but that it would be *morally better* to refrain from eating meat; if that is the whole of their position, these people are *understating* the moral failing involved in eating meat.³⁰

Works Cited:

Cornell, Nicholas Wrongs Without Rights, dissertation in progress at Harvard University.

Dreier, James (2004). "Why Ethical Satisficing Makes Sense and Rational Satisficing Doesn't," in *Satisficing and Maximizing* (Michael Byron, ed., Cambridge University Press), 131-154.

³⁰ Thank you to Jamie Dreier, Tyler Doggett, Alex Guerrero, Peter Graham, Christopher Heathwood, Barry Maguire, Jennifer Morton, Alastair Norcross, Michael Otsuka, Tamar Schapiro, Holly Smith, and Sharon Street for comments on drafts of this paper or related papers. Thank you to audiences at the Arizona Normative Ethics Workshop, Corridor, the Eastern APA, New York University, Oxford University, the Pacific APA, Princeton University, Rutgers University, Syracuse University, University of California, Los Angeles, University of Texas at Austin, and Yale University, where this paper or related papers were discussed.

- Driver, Julia (1992). "The suberogatory," *Australasian Journal of Philosophy* 70: 3: 286-295.
- Dowell, Janice (2013). "Flexible Contextualism about Deontic Modals: A Puzzle about Information-Sensitivity," *Inquiry* 56: 2-3: 149-178.
- Gert, Joshua (2012). "Moral Worth, Supererogation, and the Justifying/Requiring Distinction," *Philosophical Review* 121: 4: 611-618.
- Jackson, Frank (1991). "Decision-Theoretic Consequentialism and the Nearest and Dearest Objection" *Ethics* 101: 461-482.
- Kolodny, Niko and MacFarlane, John (manuscript). "Ought: Between Subjective and Objective."
- Harman, Elizabeth (manuscript). "Eating Meat as a Morally Permissible Moral Mistake."
- Harman, Elizabeth (manuscript). "Gamete Donation as a Laudable Moral Mistake."
- Harman, Elizabeth (manuscript). "Morality with the Realm of the Morally Permissible"
- Horgan, Terry and Timmons, Mark (2010). "Untying a Knot from the Inside Out: Reflections on the "Paradox" of Supererogation," *Social Philosophy and Policy* 27: 2: 29-63.
- Liberto, Hallie (2011). "Denying the Suberogatory," *Philosophia* 40: 395-402.
- MacFarlane, John (2014). *Assessment Sensitivity: Relative Truth and its Applications* (Oxford: Oxford University Press).
- Portmore, Douglas (2003). "Position-relative consequentialism, agent-centered options, and supererogation," *Ethics* 113: 303-332.
- Portmore, Douglas (2008). "Are Moral Reasons Morally Overriding?" *Ethical Theory and Moral Practice* 11: 369-388.
- Postow, Betsy (2005). "Supererogation again," *Journal of Value Inquiry* 39: 245-253.
- Raz, Joseph (1975). "Permissions and Supererogation," *American Philosophical Quarterly* 12.

Sliwa, Paulina (manuscript). "Moral Worth and Acting from Moral Knowledge."

Smith, Holly (manuscript). Making Morality Work.

Thomson, Judith Jarvis (1971). "A Defense of Abortion," *Philosophy and Public Affairs* 1: 1.

Urmson, J. O. (1958). "Saints and Heroes" in Essays in Moral Philosophy (A. Melden, ed.,
Seattle: University of Washington Press).