

## **Morality Within the Realm of the Morally Permissible**

by Elizabeth Harman

### **0. Introduction**

It is common to distinguish among moral theories by comparing which behaviors are morally wrong according to each theory, and why these behaviors are morally wrong according to each theory. But there is another way that moral theories differ: they differ in their views of the realm of moral permissibility.

I will argue that we can divide moral theories into four categories according to their views of the realm of the morally permissible. First, consider this question:

1. Can an agent face two morally permissible options, one of which is morally better than the other?

Views that answer “no” to this question hold that morality is very demanding, such that one is always morally required to behave in the morally best way available to one; or that morality is not very demanding because one often has options that do not differ morally at all. Such views do not acknowledge a *non-trivial realm of moral permissibility*. They hold that there are sometimes behaviors that are merely morally permissible (not also morally required), but they hold that whenever one has more than one morally permissible option, the options do not differ morally. Views that answer “no” to this question fall into the first category. (Such views are false, in my view, because they deny the existence of supererogatory actions.)

Views in the other three categories answer “yes” to the first question. (It might be useful to refer to the chart on page 4 while reading my mapping out of the four categories of views.)

Now consider this question:

2. Are there any cases in which an agent has more than one morally permissible option available to her, and one of those options is such that, all things considered, she should take that option?

Views that answer “no” to this question hold that whenever we have moral permissibility, we also have what we might call all-things-considered permissibility. On these views, if an action is not morally wrong then it is never the case that one should not perform that action, all things considered. (Such views are false, in my view, because we often have options available to us that we should not take but that are not morally wrong.) Views that answer “yes” to question 1 but “no” to question 2 fall into the second category.

Views in the remaining two categories answer “yes” to both questions 1 and 2.

Now consider this question:

3. Given that there are some options one should take all things considered, but which are not morally required, are any of these such that one should perform them *for purely moral reasons*?

Views that answer “no” to this question hold that moral reasons cannot settle that one should do something without making it morally required. Views that answer “yes” to the first and second questions, but “no” to the third question, fall into the third category. Many ethicists are committed to a view in the third category, but I will argue that views in the third category are implausible for reasons not generally recognized. Once we grant that there is a robust realm of the permissible, in which one sometimes has permissible options that differ morally, and once we grant that sometimes one should take one of these options and should not take others of these options, it is strange to hold that it is only options that are partly favored by non-moral considerations that can win out to be all-things-considered required. Consider supererogatory options. Sometimes one should not perform a particular supererogatory action, all things considered (perhaps it would be too costly to oneself). Some moral views might hold in addition

the following: Sometimes a supererogatory action is such that, all things considered, one should perform it; one's reasons favor it; failing to perform that action would be doing something that one all things considered should not do; but it would not be morally wrong. I will argue that indeed this is sometimes the case. If any supererogatory action is such that one should perform it, all things considered, because of the moral considerations that favor it, then failing to perform that action is what I will call a *morally permissible moral mistake*: an option that one should not take—for moral reasons—but that is not morally wrong. I will argue that there are some morally permissible moral mistakes; but the crucial claim I need for this paper is a weaker claim: that a moral theory can hold that there are some morally permissible moral mistakes. Such moral theories fall in the fourth category; they answer “yes” to all three questions.

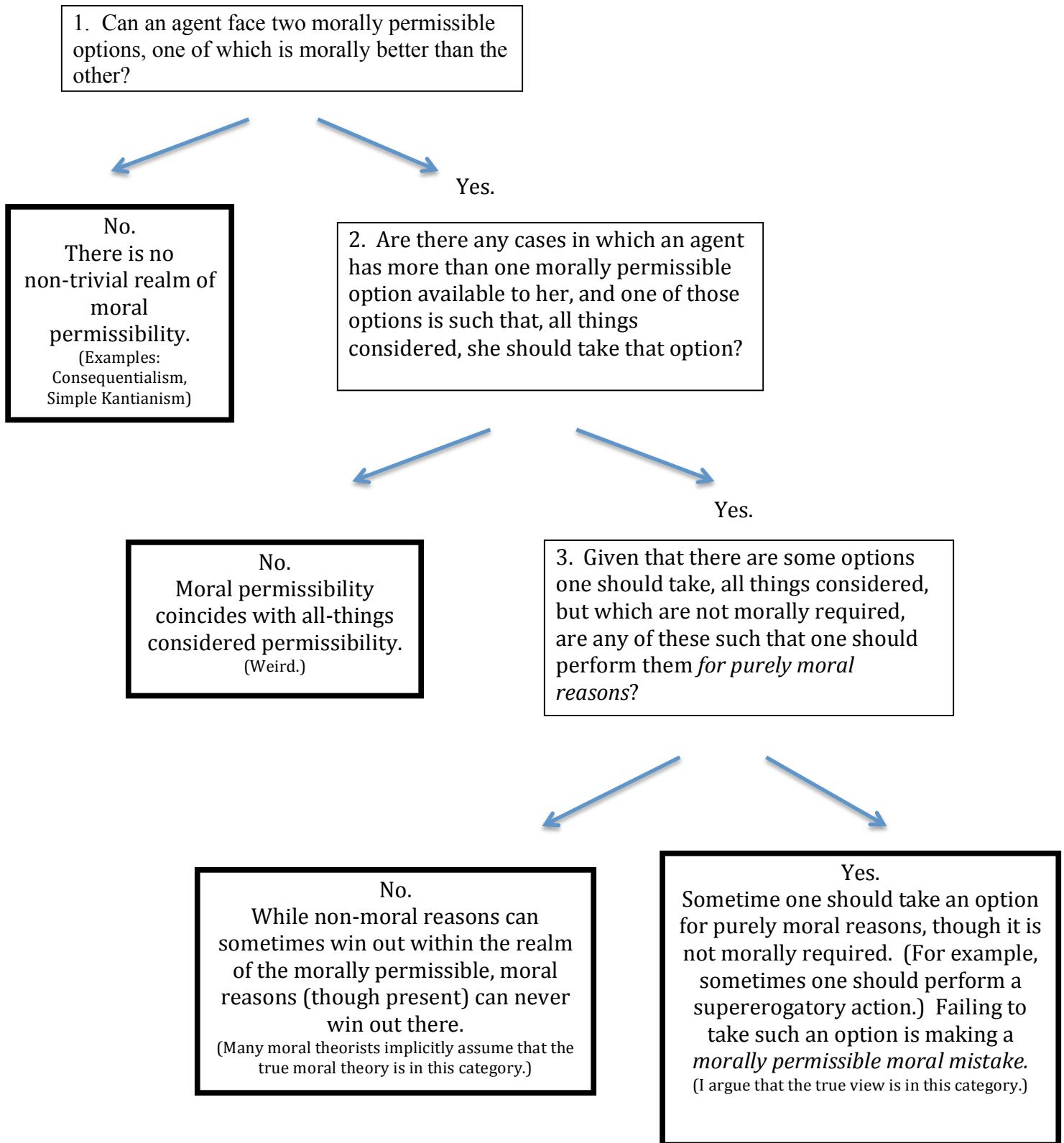
While I claim that the phenomenon of morally permissible moral mistakes is underappreciated by moral philosophers, I claim that recognition of the phenomenon is pervasive in ordinary moral thinking. Consider how natural it would be in ordinary conversation to say to a friend, “Well, you don't have to do it. But it would be a nice thing to do. You should do it.”

Moral theories that answer “yes” to all three questions hold that some options are morally permissible moral mistakes: they are actions that one should not take, all things considered—for moral reasons—but that are not morally wrong. These theories see moral reasons as relevant not just to which actions are morally required but also to which actions one should perform, all things considered. These theories see morality as playing an important role *within the realm of the morally permissible*.

My goals in this paper are the following. I will introduce and develop the categorization of moral theories into the four categories I have mentioned; I think this categorization is interesting in its own right. I will argue that the true view is in the fourth category, thus arguing for the existence of which I call *morally permissible moral mistakes*. Meanwhile, I will also be

arguing for a weaker claim: that a view in the fourth category is coherent. We should recognize that there can be moral theories in the fourth category.

## Four categories of Moral Theories



## 1. Is there a non-trivial realm of moral permissibility?

Consider the following question:

Question 1: Can an agent face two morally permissible options, one of which is morally better than the other?

Some ethical theories answer “no” to this question. Consider a standard consequentialist view:

Consequentialism: An option is morally permissible just in case there is no alternative option with morally better consequences. Consequences are morally better or worse impartially (betterness is not relative to agents or times)<sup>1</sup>. Options themselves are morally better or worse only to the extent that their consequences are morally better or worse.

On this view, if one option is morally better than another, that is because its consequences are better, but then it is morally required to choose that option over the other. So the answer to Question 1 is “no.” Consider also the following view:

Simple Kantianism: An option taken for a particular reason is morally permissible just in case its agent can reasonably wish it to be a law of human nature that agents in general take options like that for reasons like that. Morality consists of this test of moral permissibility. An option is morally better than another option just in case the first option is morally permissible and the second option is morally impermissible.

On this view, the answer to question 1 is also “no.” (I’m not claiming that any Kantian view really has this form.) If morality is simply a test of moral permissibility for actions, then there is no room for one morally permissible option to be morally better than another morally permissible option.

---

<sup>1</sup> This sentence clarifies that the consequentialism in question is, as we might say, really consequentialism, and not a deontological view “consequentialized”.

The correct answer to question 1 is “yes,” however. The existence of the supererogatory shows the answer to be “yes.” Consider the following two cases.

Burning building: A passerby, Angela, sees a building on fire, and a child at the second-story window. Angela could keep her distance, or she could attempt to enter the building and save the child. This would involve risking his life to try to save the child’s life.

Crying stranger: While staying at a hostel in Rome, Brian, an American college student, hears another guest crying in his room. When Brian sees the other man in the hallway, he has two options: say nothing, or ask if he is okay and wants to talk, thus opening himself up to being burdened by the man’s problems and to being delayed in his plans to explore the city.

I make the following claims about these cases.

It is morally permissible for each agent to try to help, but it is also morally permissible to refrain from trying to help.

Trying to help is morally better than not trying to help.

In both cases, trying to help would be supererogatory: it would be a morally good thing to do, but it is not morally required. I think that cases like this show that the answer to question 1 is “yes.” But this is controversial. Some people deny that any actions are supererogatory. (One might hold that helping in these cases is morally required, as Consequentialism might hold. Or one might hold that helping is not morally required but also is not morally better, as Simple Kantianism might hold. Alternatively, either Consequentialism or Simple Kantianism might hold that it is not morally required for Angela to help, but it is morally required for Brian to help.<sup>2,3</sup>)

---

<sup>2</sup> Simple Kantianism might hold that we can reasonably wish for a world in which people do not constantly risk their lives to save others, but we cannot reasonably wish for a world in which people do not reach out to sad strangers.

Views that answer “no” to question 1 fall into the first category: these are views on which there is no non-trivial realm of moral permissibility.

## **2. What is the relationship between moral permissibility and all-things-considered permissibility?**

Let’s restrict our attention now to views that answer “yes” to Question 1. Consider the following question.

Question 2. Are there any cases in which an agent has more than one morally permissible option available to her, and one of those options is such that, all things considered, she should take that option?

This question is less controversial than question 1. Of those people who answer “yes” to question 1, I think most will answer “yes” to question 2 as well. Consider the following cases.

Pool: Carla is playing a friendly game of pool. No money is at stake, nor is either party likely to feel particularly happy or upset with a win or a loss. But Carla is doing her best to win. She contemplates a difficult shot at the number 4 ball; if she is successful, she’ll be well positioned for another shot. Alternatively, she could take an easy shot at the number 5 ball; but then she’d have few good options for the next shot. Given how her friend is playing, Carla has little hope of winning unless she can sink several balls on this turn. Carla should go for the trickier shot.

Teeth: Ernie is tired and feels like going to bed without brushing his teeth. Ernie should brush his teeth.

Deportation: Fiona is on her way to a deportation hearing when she sees a man trip and fall down on a busy sidewalk. The man is not seriously hurt, but Fiona would

---

<sup>3</sup> What a consequentialist view would hold about the case may depend on what Angela should think about the likelihood that she can save the child and survive herself; if this is less than 50% likely, then a consequentialist view might hold that, given her information, she should not try to help.



normally stop to help him and make sure he is okay. If Fiona stops, this may result in her being deported. Fiona should not stop.

I claim that in the above three cases, each agent faces two morally permissible options, but *should* perform one of them rather than the other. I do not think my claims about these cases are controversial.

As with most philosophers' uses of example cases, my claim is that in the most natural version of the above cases, given my non-moral descriptions of them, the further claims I make about the cases (the normative claims) are true.

I would like to introduce another example case, which I will use in a slightly more unusual way. Here is the case:

Charity: George is deciding between working with two different charities. One addresses lung cancer and the other addresses famine. The work he could do to contribute to the famine charity would make more of a difference in the world: he is in more of a position to be helpful to them, and their work saves more lives. But George's mother died of lung cancer recently. Working with the lung cancer charity would enable him to cope with the loss of his mother: it would be a statement of love and he would meet others who have suffered similar losses. It is morally permissible for George to volunteer with either charity. But all things considered, he should volunteer with the lung cancer charity.

I do not think it is obvious from the non-moral details I have given about this case that the normative claims are true of it. I claim the following:

The Charity case is a possible case: it is possible for the normative claims to be true in a case in which these non-normative claims are true.

These four cases show that the answer to question 2 is "yes."

Although the view is unusual, we can consider a moral view that answers “yes” to question 1 but “no” to question 2. One might have the view that the realm of the morally permissible coincides with the realm of all-things-considered permissibility. Within this realm, various considerations tell in favor of various options. These considerations make these options *eligible* as reasonable choices to make; but they do not settle that one *should* take one option rather than another.

It is true that considerations *sometimes* make an option *eligible* as a reasonable choice to make without making it the option one *should* take. For example, suppose you are deciding whether to have Japanese food or Indian food for lunch. You are torn. You enjoy both foods, and both restaurants, for quite different reasons. The Indian food is yummy, filling, and less expensive, but the restaurant is further away. The Japanese food has a delicate sophistication which is more of a meaningful aesthetic experience for you; it is equally yummy, the restaurant is closer, but Japanese food is less filling and more expensive. You realize that even a slight improvement in one option or the other would not settle that you should take one option rather than the other. If you suddenly found a \$1 off coupon to one restaurant, for example, that would not settle that you should go there. If you are in such a case, your options are *on a par*: it is not the case that you should take either one, nor that you should not take either one, nor do *small sweetenings* (or improvements) in either option make it the case that you should take that option. (By contrast, if the two options were exactly equally supported by your reasons, then a small sweetening of either option would settle that you should take that option.)<sup>4</sup> In this case, it seems that your reasons to go to the Japanese restaurant make that option *eligible* as a reasonable option

---

<sup>4</sup> It is controversial whether it is possible for two options to be on a par. If this is possible (as I believe it is), then there are *four* possible relations that two available actions  $\phi$  and  $\psi$  can stand in: it may be that it is better to  $\phi$  than  $\psi$ ; it may be that it is better to  $\psi$  than  $\phi$ ; it may be that it is equally good to  $\phi$  as to  $\psi$ ; or it may be that  $\phi$ -ing and  $\psi$ -ing are on a par. For an argument that two options can be on a par, see Ruth Chang, “The Possibility of Parity,” *Ethics* 2002.

for you to take, without making it the option that you *should* take; and the same is true of your reasons to go to the Indian restaurant.

While this does *sometimes* happen—sometimes one faces two significantly different options that are both supported by one’s reasons without either being *avored* by one’s reasons—I think this is unusual. A view that answers “yes” to Question 1 but “no” to Question 2 holds that this kind of situation is in fact incredibly common, and that whenever one finds oneself with more than one morally permissible option, these options are also all-things-considered permissible: neither is favored by one’s reasons.

### **3. Morally Permissible Moral Mistakes**

Now let’s confine our attention to views that answer “yes” to both question 1 and question 2. On these views, an agent can sometimes face at least two morally permissible options that themselves differ morally; and sometimes, among morally permissible options, an agent should take one of these options rather than the other. Now consider:

Question 3. Given that there are some options one should take, all things considered, but which are not morally required, are any of these such that one should perform them *for purely moral reasons*?

Note that the examples we considered in section 2 do not help us to answer this question. In each of those cases, the explanation of why the agents should take the particular options involves some non-moral considerations, such as strategic considerations or prudential considerations.

Let’s consider views that answer “no” to question 3. On these views, although moral reasons are present in the realm of the morally permissible, moral reasons cannot win out. When an agent faces at least two morally permissible options that differ morally, she might be in any of the following three situations:

- (a) she should take a morally worse option over a morally better option, because of some non-moral reasons, such as prudential reasons

(b) she should take a morally better option over a morally worse option, because of some non-moral reasons, such as prudential reasons

(c) neither option is such that she should take it, either because each option is *equally supported* by her reasons, or because the options are *on a par*

We saw an example of a case in which (a) is true with the Deportation case: it is morally worse to rush to one's own deportation case rather than stop to help someone who has fallen, but rushing on is the thing that Fiona should do.<sup>5</sup> We can see an example of a case in which (b) is true if we consider someone who is choosing between volunteering at two different charities; one of them does more good, so that volunteering there is the morally better choice, but it is also much more convenient for her. She should volunteer at the morally better, more convenient charity, but not for purely moral reasons; part of the reason she should do so is that it is convenient for her. (Indeed, we can imagine a version of the case in which *the primary factor* in the explanation of why she should choose that charity is the convenience.) I do believe that there are cases in which, between two options that differ morally, the reasons to take the morally better option are *on a par* with the reasons to take the morally worse option; but I do not think it is obvious what these cases look like. (I am not sure whether a morally good option can ever be *equally supported* by the agent's reasons as a morally worse option is; this does not matter for our discussion.)

Views that answer “yes” to questions 1 and 2, but “no” to question 3, recognize that cases in which an agent faces at least two morally permissible options that differ morally may be of the form (a), (b), or (c), but they leave out an important further possibility:

(d) she should take a morally better option rather than a morally worse option *because of the moral reasons* that favor the better option

---

<sup>5</sup> Many people facing deportation have moral duties to family members to avoid getting deported. Suppose that Fiona has no family who are relying on her to stay in the country, and that her reasons for saying are self-interested. Then hers is a case of type (a).

But there are cases in this category. Consider this case:

Mechanic: Hannah is a car mechanic who owns her own small business. Irene, who is passing through town, brings her car in. Irene tells Hannah about her life. Irene has barely enough money to make it where she is heading with her three kids, starting a new life having escaped from domestic abuse. She can just cover the cost of the repairs and make her trip, but she'll have to drive much longer days than she was planning because the repair will eat up the cost of one night in a motel. Hannah knows that it's morally permissible for her to charge Irene for the repair. But she considers not charging her. Hannah concludes, "I shouldn't charge her." She's right.

My claim about this case is:

The Mechanic case is a possible case: it is possible for the normative claims to be true in a case in which these non-normative claims are true.<sup>6</sup>

I think my claim about this case is pretty uncontroversial, but that its import is generally underappreciated. We often have options available to us that are clearly supererogatory.

Sometimes one takes a supererogatory option, and in doing so one is responding appropriately to one's reasons. And sometimes it is not just that the supererogatory option was *a* reasonable option to take, but it was *the* option favored by one's reasons. Sometimes an agent thinks,

---

<sup>6</sup> My point here is that I haven't filled in all the details of the Mechanic Case to make it uncontroversial and obvious that it is *morally permissible* for Hannah to charge Irene, and yet that she *shouldn't* charge her. My claim is that there is some way of filling in those details so that those normative claims will be *true*. I think it should be fairly uncontroversial that there is *some way* of filling in the details so that those normative claims will be true, though I do not think there is a particular way of filling in the details such that it will be generally uncontroversial that *that way* of filling in the details makes the normative claims true.

Note that I am *not saying*, regarding a *complete* non-moral description of the situation, that it is merely *possible* that the normative claims are true in that case. Rather, I am saying that it is possible for the *partial* description I gave to hold of a case in which the normative claims hold. (Thus, I do not mean to suggest that the moral might not supervene on the non-moral.)

regarding a supererogatory action, “I should do this,” and she is right about that: given all of her reasons, that is the thing she should do.

Consider Hannah’s deliberative process. She recognizes that it is morally permissible to charge Irene for the repair to her car. Of course it is! But still she considers helping Irene. That she considers helping her seems utterly reasonable, even though the only reasons to help Irene are *moral* reasons: that it would be good for Irene and her children, who are escaping a bad situation. Hannah’s choice to help Irene also seems reasonable. My claims about deliberation are these:

In deliberating about what to do, an agent might come to *know* that a particular option is supererogatory, while continuing to consider taking that option. Furthermore, the agent might consider the possibility that she *should* take that option. She might conclude that she *should* take that option, and she might be right, even though throughout her deliberation, she *knows* that the option is supererogatory—it is morally good to do, but it is not morally required—and she knows that the only considerations that favor doing it are moral considerations.

This is true, but it is also reflected in the *phenomenology* of deliberation. It often seems to an agent that she knows that an option is supererogatory, even while deliberating and even while she concludes that she should take it because of the moral considerations that favor it. This seeming is often accurate.

These claims about deliberation and about the accuracy of the phenomenology of deliberating about whether to take a supererogatory option must be denied by any moral theory that answers “yes” to questions 1 and 2 but “no” to question 3.<sup>7</sup>

---

<sup>7</sup> The importance of taking the phenomenology of deliberation seriously is also stressed by Horgan and Timmons in their “Untying a Knot from the Inside Out: Reflections on the ‘Paradox’ of Supererogation” *Social Philosophy and Policy* (2010) 27: 2: 29-63.

I conclude that the answer to question 3 is “yes”, because sometimes an agent should perform a supererogatory action; and in some of these cases, the only considerations that favor that action are moral considerations.

Sometimes one should take an option, for moral reasons, though it is not morally required. I call the failures to take such options *morally permissible moral mistakes*. Here are some terminological stipulations:

S’s  $\phi$ -ing is a mistake = (def) S should not  $\phi$ , all things considered

S’s  $\phi$ -ing is a moral mistake = (def) S should not  $\phi$ , all things considered, and the reasons against  $\phi$ -ing that win out to make it the case that S should not  $\phi$  are moral reasons

S’s  $\phi$ -ing is a morally permissible moral mistake = (def) S’s  $\phi$ -ing is a moral mistake but S’s  $\phi$ -ing is not morally wrong

In the Mechanic case, Hannah’s charging Irene for the repair would be a morally permissible moral mistake. She should not charge her; the reasons against charging her that win out to make it the case that he should not charge her are *moral reasons*; and yet it is not morally wrong to charge her.

In sections 4 and 5, I will consider some objections to my argument in this section.

#### **4. Rejecting Morality Within the Realm of the Morally Permissible**

I have argued that the answer to question 3 is “yes”, and that moral reasons in favor of an option sometimes win out to make it the case that an agent *should* take that option without making that option morally required. Restricting our attention to views that answer “yes” to questions 1 and 2, let’s consider more carefully views that answer “no” to question 3. How can they resist the argument I made in section 3? What alternative picture(s) can they offer?

There is more than one way that such views might go. A view might hold:

- What looks like a *supererogatory* action that should be performed in cases like Mechanic is actually a morally required action

Or:

- What looks like a supererogatory action that *should* be performed in cases like Mechanic is actually an action made *eligible* by the agent's reasons, but it is not an action the agent should perform.

Or:

- What looks like a supererogatory action that should be performed *that is supported by purely moral reasons* is actually supported by some non-moral reasons, and that is a crucial part of why it should be performed.

Or, finally:

- What looks like a supererogatory action that *should* be performed in cases like Mechanic is actually an action that the agent should not perform, in light of all of the agent's reasons; this is part of what is morally admirable about the action.

And these responses might be combined to handle different cases differently. I will take each kind of response in turn. My reaction to each response will be to acknowledge that there are some cases with the structure described, but to deny that the Mechanic case, and others like it, *must* have the structure described by my opponent rather than the structure I have described.

Let's consider first this response to cases like the Mechanic case:

- What looks like a *supererogatory* action that should be performed in cases like Mechanic is actually a morally required action

My opponent points out that morality sometimes forbids doing what is ordinarily permissible.

Sometimes it is morally wrong to charge someone for a service or a good, although it is normally morally permissible to do so. That's certainly right. Consider the following case:



Asthma: A woman and her daughter are fleeing an abuser. Their car breaks down, and it will take all of their cash to pay for the repair. They have somewhere to stay when they get where they are going, but they have no way to pay for the daughter's asthma medicine, which they need to buy right away to prevent the possibility of an asthma attack, which could be life-threatening. Julie, the mechanic, learns all of this. Her garage has been making more money recently, and she has not already made any plans for the extra money.

It may well be morally required for Julie to give the woman the repair for free. The health risk to the daughter makes Asthma a significantly different case from Mechanic. My opponent says this:

The Asthma case shows that *sometimes* a mechanic may be morally obligated not to charge a customer, because of the burden on the customer and the mechanic's ability to get by without the payment. If we consider the original Mechanic case, in which Hannah *should* give Irene the repair for free, this is simply a case in which it is *morally required* to do so. It's not as morally wrong for Hannah to charge for the repair in Mechanic as it would be for Julie to do so in Asthma, but it is still morally wrong.

I am not moved by this. Given what is actually at stake for Irene and her children in Mechanic – driving longer hours on their trip, versus one more night in a motel and easier driving days – and even taking into account what they have recently been through, it simply could not be *morally wrong* for Hannah to charge them. Nevertheless, it is utterly reasonable for Hannah to *consider* giving them the repair for free, and she might well conclude, correctly, that she *should* do so.

This first opponent that we are considering wants to agree with me that there are versions of the Mechanic case in which Hannah *should* give them the repair for free, but disagrees with me that it is morally permissible to charge them in these versions of the case. My response is that there is simply no version of the Mechanic case in which it is morally wrong for Hannah to charge

them (although it may be wrong for Julie to charge for the repair in Asthma, which is not a version of the Mechanic case).

Note also that this first opponent must deny the claims I made about deliberation. The first opponent must hold that, while Hannah is correct that she *should* give Irene the repair for free, she is simply wrong in her belief that doing so is supererogatory. One cannot know *first* that an option is supererogatory while leaving it open whether one *should* perform that action, if it is only favored by moral considerations. Rather, if one opens the question of whether, all things considered, one *should* perform the action, then one is also opening the question of whether the action is supererogatory. These claims are implausible.

Let's turn now to an opponent who says this:

- What looks like a supererogatory action that *should* be performed in cases like Mechanic is actually an action made *eligible* by the agent's reasons, but it is not an action the agent should perform.

Recall the case of choosing whether to eat Indian food or Japanese food for lunch. Sometimes two significantly different options are both supported by one's reasons, but neither is favored by one's options. Sometimes they are not *exactly equally* supported by one's reasons, as we can see by noting that small improvements (or sweetenings) in each option do not make either option the one the agent *should* take. In such cases, the two options are *on a par*. As I've already commented, I find it implausible that a supererogatory action is ever *equally supported* by the agent's reasons as a non-supererogatory alternative, but that is not crucial for our discussion. I will restrict my attention, in discussing this opponent's view, to the idea that sometimes a supererogatory option may be *on a par* with a non-supererogatory option. I certainly agree that this *can* happen.

This second opponent might furthermore say the following thing about the phenomenology of deliberation:

An agent such as Hannah, who knows that one of her options is supererogatory, may nevertheless consider taking that option, yes. But Hannah does not ask herself the question “what should I do?” Rather, she asks herself, “what to do?” In deciding *what to do*, Hannah is asking a different question. She is asking what she chooses. This is a question that she must settle for herself, even if her reasons do not favor one option over another. Knowing the answer to “what should I do?” does not settle the answer to the question she’s really asking, “what to do?” in any cases in which her reasons do not *favor* one option over another.

It is certainly true that sometimes we must choose between options that we know to be *on a par* or *equally favored* by our reasons. I must choose one can of diet coke from the fridge, though I know it is not true that I *should* take the one on the left rather than the one on the right. We often know, when we are choosing between two very similar options, that neither option is favored by our reasons. But how often do we know this about choices that differ significantly? I think it is unusual to know this. Rather, I think the normal mode of deliberation when choosing between options that differ significantly involves implicitly taking seriously the possibility that one option is *favored* by one’s reasons and trying to figure out which one it is.

How often do our reasons settle that we should do one thing rather than another, and how often do our reasons simply make more than one significantly different option *eligible* to us? I think the latter situation is rare. We should not mistakenly think it is more common than it is by confusing that situation with the following situations:

- (i) It is difficult for the agent to know how to weigh some very different kinds of factors that tell in favor of two different options.
- (ii) It is difficult for the agent to know what will happen if she chooses each of two different options.

- (iii) It is better for the agent to choose between two options now, rather than waiting to think about it more, even if waiting would increase the likelihood that the right choice would be made.
- (iv) The agent has two options that are both wonderful and that would lead to good outcomes.
- (v) The agent is choosing between two different types of actions; it is not the case that, in general, agents choosing between these types of actions should always choose one particular type of action over the other.

Each situation, (i) through (v), is compatible with its being the case that, in light of all of her reasons, the agent *should* pick one of her options rather than any of her other options.

Sometimes an agent's reasons favor an outcome, though it is very hard to see that that is true, as in (i). Sometimes there is a lot of information that is relevant that the agent does not have, which can increase how difficult a decision *feels*, as in (ii); but it does not mean that the choice is up for grabs; there are still facts about what credences are warranted in light of whatever evidence the agent does have, and there may be a single option that the agent should take in light of those credences. Sometimes there is a reason to choose now, even if doing so would involve choosing an option that is worse in light of all of one's reasons (perhaps ignoring the reason to choose now), as in (iii); in such a case, it may be true that the agent should have chosen differently, in light of her reasons, even though it is also true that she should not have waited to figure out that that was the right choice to make. Sometimes we say to someone, "you can't go wrong," when what is really true is not that she cannot act as she *should not*, in light of all of her reasons, but that she has two options that are both wonderful, as in (iv). Sometimes an agent should make a particular choice, given the details of her situation, although a different choice would have been warranted if her situation had been a bit different; thus, it is not the case that agents in her situation should always choose in a certain way, as in (v).

Usually one's reasons support one choice over the others available to one. While there is a lot of moral permissibility in our lives, there is not very much all-things-considered permissibility in our lives. We often do things that we should not do. Often, it isn't so terrible that we do things we should not do. There is a certain amount of weakness of will in our lives that isn't terribly bad for us (eating that extra chocolate or skipping that day at the gym). Sometimes we choose to make a choice quickly rather than making the right choice; this may be the life policy that one *should* adopt, although it means that one sometimes does particular things one *should not* do. And sometimes we simply are wrong about what we should do; but the wrong choice is not seriously worse than the right one. On the other hand, sometimes it is terrible that we do things we should not do; sometimes we make big mistakes.

My third opponent says this:

- What looks like a supererogatory action that should be performed *that is supported by purely moral reasons* is actually supported by some non-moral reasons, and that is a crucial part of why it should be performed.

The third opponent might point out that people are often made to feel good by helping others, and that this always counts as a reason in favor of helping others, so it is hard to find a morally good action that should be performed *for purely moral reasons*. In fact, the claim I have made is a bit more modest: that there are cases in which an agent has two morally permissible options, and should perform one of these options, where the reasons in favor of the option that *win out* to *make it the case* that she should take that option are moral reasons. One might hold that if Hannah should give Irene the repair without charging her, the reasons that win out to make it the case that she should do so are the considerations that it would be good for Irene and her children in a difficult time. Even if it would make Hannah happy, that isn't part of what makes it what she should do. But alternatively, we can simply stipulate that she won't be made happy, that

somehow helping Irene would make Hannah feel emotionally involved in an unsettling way. So let's set aside the thought that helping someone often makes one happy.

The third opponent might instead revise a thought that the second opponent had: that often our reasons *underdetermine* what we should do. Consider, for example, Ruth Chang's view, according to which one's "given reasons" (the reasons we find ourselves with) often underdetermine what one should do, but then one makes a *choice* to *endow* a reason with extra strength, creating a new, coinciding "voluntarist reason" which settles that one *should* act in the way that reason supports.<sup>8</sup> Chang's motivating examples include cases like these:

Romance: Kyle and Louis are in a romantic and sexual relationship, but they have not defined it. So far it is casual. Kyle has no compelling reason to take care of Louis's needs, for example by cooking him dinner if he is exhausted after a long day of work, or by picking up Louis's laundry left on the floor of Louis's apartment. It is not the case that Kyle *should* do these things. But if Kyle *commits* himself to Louis – a mental act – it thereby becomes true that Kyle should do these things.

Career: Megan is trying to decide between two career directions. She will enter either a Masters of Fine Arts program in order to become an actor or a Masters of Social Work program in order to become a psychotherapist. Megan feels passion for acting, but the career path is uncertain and stressful. Megan has a real talent for both acting and therapy; her listening skills and empathy are an asset for both. Therapy is a safer path; she would feel that her work has meaning because she would be helping others; but that path involves less of an outlet for her creative energy. Megan's reasons do not settle what she should do. Megan thinks it over

---

<sup>8</sup> See Ruth Chang, "Voluntarist Reasons and the Sources of Normativity" (in *Reasons for Action*, Sobel and Wall, eds. (New York: Cambridge University Press, 2009), pp. 243-71), which is part of a bigger project, encompassing additional papers and a book manuscript, *Making it Matter*.

and makes a *commitment* – a mental act – to pursue a career as a psychotherapist.

It thereby becomes true that Megan *should* become a psychotherapist.

Chang clarifies that she does not see the relevant commitments as *the formation of intentions*, and the phenomenon she is discussing is not the phenomenon of how having formed an intention may influence how one should act. Also, the commitments are private mental acts. The phenomenon Chang is interested in is not the way that commitments we make to others impact how we should act.

Chang's view presents a challenge to my own view because if she is right, then it may be that whenever one *should* choose a supererogatory act over a non-supererogatory act, this is because one has privately *committed* to this endeavor, in thinking it over, thus creating a new “voluntarist reason” which is (perhaps) not properly thought of as a moral reason. So even in cases in which it looks like there are only moral reasons in favor of performing the supererogatory act, it turns out that there is a non-moral reason in favor of it.

Perhaps something close to Chang's description does happen *sometimes*. Perhaps sometimes an agent *should* perform a supererogatory actions simply because she *wants to*, or because she feels like doing so. But there would still be many cases, such as Mechanic, in which the agent lacks that independent inclination, and is moved by the considerations that favor the supererogatory action. And it is not clear that the mere fact of inclination or desire ever provides a reason. It may be that inclinations and desires *indicate* the presence of reasons, without themselves being reasons – one may have an inclination or desire for a reason, where that reason is typically also a reason to do the thing one wants to do.<sup>9</sup>

As I understand her views, Chang and I disagree about the extent to which a person's given reasons *underdetermine* what she should do. The choices that Chang discusses – whether

---

<sup>9</sup> There is an extensive literature on the question of whether a person's desires or inclinations themselves *provide* reasons, or should be *responsive* to independently-existing reasons. I don't want to take a stand on that issue in this paper, but I do have sympathy with the view that desires and inclinations do not provide reasons.

to be seriously involved with a particular person, what hobbies to have, what career to pursue, etc. – are difficult choices. But I think it is rare that in such cases a person faces two options that are equally favored by her reasons or are on a par. As I argued above, while situations (i) through (v) are common, none of them need be situations in which a person’s given reasons do not settle what she should do.

My fourth opponent says:

- What looks like a supererogatory action that *should* be performed in cases like Mechanic is actually an action that the agent *should not* perform, in light of all of the agent’s reasons; this is part of what is morally admirable about the action.

This opponent says that what is so admirable about supererogatory actions such as Hannah’s in Mechanic is Hannah’s willingness to act contrary to what her reasons support, in order to help another: it is precisely because she *should not* do it that the action is supererogatory, and praiseworthy.

Unfortunately I do think that the idea that supererogatory actions are ones that agents should not perform, in light of all their reasons, is common.<sup>10</sup> But the view is false. I will make three points about this.

First, we should distinguish the question of whether it is in Hannah’s interest to help from the question of whether she *should not* help. Sure, it is contrary to her own interests to help. But

---

<sup>10</sup> Some are puzzled by the existence of the supererogatory – how can an action be morally good to do but not morally required? One solution to this apparent puzzle holds that supererogatory actions are recommended by a perspective that includes only a *subset* of one’s reasons, and that these actions are not morally required because an agent has other reasons that *tell against* performing the actions, such that all things considered the agent should not perform the actions. An example of this view occurs in Jamie Dreier’s “Why Ethical Satisficing Makes Sense and Rational Satisficing Doesn’t” (in *Satisficing and Maximizing*, Michael Byron, ed., Cambridge University Press, 2004, 131-154). The view also appears in Douglas Portmore’s “Position-relative consequentialism, agent-centered options, and supererogation,” *Ethics* (2003) 113: 303-332. Portmore’s commitment to the view that all supererogatory action is a mistake is pointed out by Betsy Postow in “Supererogation Again,” *Journal of Value Inquiry* (2005) 39: 245-253. Portmore repudiates the view in “Are Moral Reasons Overriding?” *Ethical Theory and Moral Practice* (2008) 11: 369-388.



that does not mean that she should not help. Morality often *requires* self-sacrifice of various kinds; and we *should* engage in that self-sacrifice.

Second, while self-sacrifice is admirable, there is nothing in particular admirable about a willingness to do what one *shouldn't do*, for someone else's sake. Suppose that you were willing to break a promise to Joe in order to help Karen. The fact that you *shouldn't* break the promise doesn't make it admirable to do it anyway in order to help Karen.

Third, it is true that sometimes one shouldn't perform a supererogatory action, because of one's self-regarding reasons. Sometimes a supererogatory action would be so personally costly that one's reasons tell against doing it.<sup>11</sup> When an agent performs such an action, there is definitely something special about it. It is praiseworthy, as any supererogatory action is, and perhaps it is praiseworthy in a unique way. A willingness to disregard the force of one's own interests for another's sake may warrant special praise.<sup>12</sup>

But not all supererogatory action involves a failure to appreciate the force of the reasons one has. Hannah may be seeing her reasons correctly and clearly when she concludes that she *should* give the woman the repair for free.

## 5. Other moral categories

Someone might object to my discussion so far by claiming that “morally permissible moral mistake” is just a new term for an already-familiar phenomenon, so I have introduced unnecessary new terminology, which is therefore obfuscating. The objector holds not just that *ordinary morality* is committed to the existence of morally permissible moral mistakes (a claim with which I agree), but that *moral philosophy* already countenances the phenomenon. I will consider three different versions of this objection.

---

<sup>11</sup> Raz claims that no supererogatory actions are such that they should not be performed. “Permissions and Supererogation” *American Philosophical Quarterly* (1975) 12.

<sup>12</sup> Unless it involves a systematic inappropriate devaluing of oneself – too common among women in sexist cultures like our own.

The first objector holds that someone makes a morally permissible moral mistake just in case she fails to take her *morally best* option. This suggestion fails in both directions. First, it is not true that one's morally best option is in general the option that one *should* take, all things considered. For example, it may be that the morally best thing I could do right now would be to go to the local hospital and offer up one of my kidneys and some of my liver. But that is not the thing that I should do right now, all things considered. Failing to do that is not making a morally permissible moral mistake, but it is failing to do the morally best thing I could do right now. The claim fails in the other direction as well: sometimes an agent ought to perform a supererogatory action, but it is not the morally best thing she could do right now. (The morally best thing she could do may be so personally costly that, all things considered, she should not do it.)<sup>13</sup>

The second objector holds that someone makes a morally permissible moral mistake just in case she fails to take an opportunity to fulfill an *imperfect duty*; after all, each action that would fulfill an imperfect duty has moral reasons in favor of it, but is not morally required. This suggestion fails because it is not true that whenever one has an opportunity to fulfill an imperfect duty, one *should* do so, all things considered. For example, suppose that you are on your way home from work and you can feel a migraine coming on. If you get home quickly, you can take a pill that would prevent a full-blown migraine that would last several hours. If you suddenly see an opportunity to do something nice for someone – perhaps to help someone who is picking up

---

<sup>13</sup> Paul McNamara has some work on the supererogatory in which it might seem that he makes one of my central claims. Both McNamara and I say things like this: “Sometimes one ought to do something that is not morally required, and the thing one ought to do is a supererogatory action.” But McNamara and I use “ought” in quite different ways, so he is not making the same claim that I am making. He uses “ought” to pick out an agent’s morally best actions. So when he says “sometimes an agent ought to do something that is not morally required” he simply means that sometimes an agent’s morally best option is not morally required. And when he says “sometimes an agent ought to perform a supererogatory action” he simply means that sometimes an agent’s morally best option is a supererogatory action. McNamara makes these kinds of claims in his “Supererogation, Inside and Out: Toward an Adequate Scheme for Common Sense Morality” (in *Oxford Studies in Normative Ethics*, Ed. Mark Timmons, Oxford University Press (2011), 202-235) and in other papers. McNamara does not address the question of whether it is ever true that one ought to perform a supererogatory action, all things considered.

trash in a park – it is not the case that you should take that opportunity, although helping others is something that you have an imperfect duty to do. Rather, you should hurry home and take your pill. Failing to stop to help the person picking up trash is not a morally permissible moral mistake, because it is not a failure to do something that you should do.<sup>14</sup>

Finally, the third objector holds that someone makes a morally permissible moral mistake just in case she does something *suberogatory*. The category of the suberogatory is a kind of inverse of the category of the supererogatory: while supererogatory actions are morally good to do, but not morally required, suberogatory actions are morally bad to do, but not morally wrong. It is controversial whether any actions are suberogatory, but I agree with the objector that some are. The objector’s suggestion fails because it is not true that every morally permissible moral mistake is suberogatory. For example, consider Hannah’s charging Irene for the repair in Mechanic. This would not be a morally bad thing to do; but she should not do it, for moral reasons. So the category of morally permissible moral mistakes is not identical to the category of suberogatory actions. Nevertheless, the objector may be right about something: every suberogatory action may be a morally permissible moral mistake: something that one should not do, for moral reasons, but that is not morally wrong.<sup>15,16</sup>

---

<sup>14</sup> This case provides another counterexample to the claim that one should always take a morally better option over a morally worse option. It is morally better to stop and help the person picking up trash, rather than driving straight home; but you should not stop and help.

<sup>15</sup> I see it as an open question whether all suberogatory actions are morally permissible moral mistakes. I discuss the suberogatory in my “Morally Permissible Moral Mistakes” (forthcoming in *Ethics*).

For interesting discussion of the suberogatory (also called “offence”), see Roderick M. Chisholm, “Supererogation and Offence: A Conceptual Scheme for Ethics,” *Ratio* (1963) 5: 1-14; David Heyd, *Supererogation*, Cambridge: Cambridge University Press (1982); Gregory Mellema, *Beyond the Call of Duty: Supererogation, Obligation, and Offence*, State University of New York Press (1991); Julia Driver, “The suberogatory,” *Australasian Journal of Philosophy* (1992) 70: 3: 286-295; Hallie Liberto, “Denying the Suberogatory,” *Philosophia* (2011) 40: 395–402; and Paul McNamara, “Supererogation, Inside and Out: Toward an Adequate Scheme for Common Sense Morality” (2011).

<sup>16</sup> I develop and defend the notion of a morally permissible moral mistake in three other papers: “Morally Permissible Moral Mistakes” (forthcoming in *Ethics*), “Eating Meat as a Morally

## 6. Conclusion

I have argued that we can divide moral theories into four categories, according to how they view the realm of the morally permissible. The first category includes views on which it is not possible for an agent to have two morally permissible options, such that one option is morally better than the other; these views do not acknowledge a non-trivial realm of moral permissibility. The second category includes views on which there is a non-trivial realm of moral permissibility, but on which it is never the case that an agent *should* act in one morally permissible way rather than in another morally permissible way. The third category includes views on which there is a non-trivial realm of moral permissibility, and on which sometimes one should act in one morally permissible way rather than another; however, views in this category deny that sometimes one should take one morally permissible option rather than other *for moral reasons*. Views in the fourth category recognize the robust presence of morality within the realm of the morally permissible: on these views, sometimes one should take one morally permissible option rather than another, for moral reasons. These views hold that some options are *morally permissible moral mistakes*. I have argued that the true moral view acknowledges the existence of morally permissible moral mistakes. But I have also argued for the following weaker conclusions. It is coherent to hold that there are morally permissible moral mistakes. We should acknowledge that a moral theory could hold that there are morally permissible moral mistakes.<sup>17</sup>

---

Permissible Moral Mistake” (forthcoming in *Philosophy Comes to Dinner*, Chignell, Cuneo, and Halteman, eds., Routledge) and “Gamete Donation as a Laudable Moral Mistake” (manuscript).

<sup>17</sup> For helpful comments on drafts of this paper, I thank Alex Guerrero, Peter Graham, and audiences at the Massachusetts Institute of Technology and the 2014 Arizona Workshop in Normative Ethics.