THREE PARADOXES OF SUPEREROGATION

Daniel Muñoz
Monash University
To appear in *Noûs*

Supererogatory acts—good deeds "beyond the call of duty"—are a part of moral common sense, but conceptually puzzling. I propose a unified solution to three of the most infamous puzzles: the classic Paradox of Supererogation (if it's so good, why isn't it just obligatory?), Horton's All or Nothing Problem, and Kamm's Intransitivity Paradox. I conclude that supererogation makes sense if, and only if, the grounds of rightness are multi-dimensional and comparative.

1. Introduction

We don't always have to do what's best. Some good deeds are *supererogatory:* they are optional and yet better than other permissible options. Examples include enduring injury to save a life, doing a favor, and granting forgiveness.[1] Lovely stuff; hardly required.

But as innocent as the examples may sound, the theory of supererogation is fraught with paradoxes. Three in particular have been influential. The classic Paradox of Supererogation asks how supererogation could be possible even in a simple pairwise choice, like the choice between saving someone from a crumbling building or safely loitering outside (Raz, 1975; Dreier, 2004: 148; Heyd, 2016: §3; Archer, 2018: §4). If being the hero is really better, why isn't it just obligatory?

The other puzzles add a third option. In Horton's (2017) All or Nothing Problem, the agent has two ways to be the hero: either rescue the one from the building, or rescue *two*, at the same level of cost and effort—say, two crushed arms either way. It seems wrong to save just one person; letting the other die is pointless badness. And yet, saving one seems no worse than permissibly saving zero. But surely it's always worse to do wrong. How could it not be?

In Kamm's (1985, 1996, 2007: 30–1) Intransitivity Paradox, meanwhile, the agent has two

---

[1] When I say that supererogation is "better," I just mean that there is more reason to do it, not that it has nice effects or merits praise. Later I will discuss views on which supererogation isn't better all things considered, but only morally better. None of my arguments depend on how we define 'better' or 'supererogatory'.
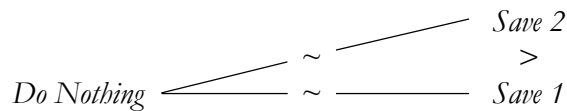
ways *not* to be heroic. Suppose you come upon the building while on your way to the cafe, where you have promised to meet a friend. You may break your promise to save a life, or you may keep the promise, but it would be wrong to just sit there. This leads to an infamous intransitivity: you may do nothing instead of saving the one (in a pairwise choice), and may save the one instead of keeping your promise, but you may *not* do nothing instead of keeping the promise. How is this possible?

These puzzles are usually discussed in isolation. It is not standard to check whether a solution to Kamm's conundrum could work for Horton's, or vice versa, nor do people discuss these two in light of the classic Paradox. But we need a unified solution. My goal is to find one.
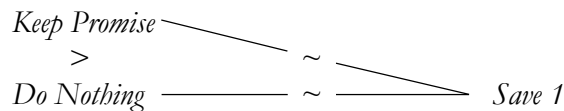
Why take on the puzzles together? For one thing, the choices involved are undeniably linked. In the classic Paradox, we have:
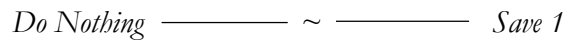
$$\textit{Do Nothing} \, \text{————} \, \sim \, \text{————} \, \textit{Save 1}$$

Where '*Do Nothing ~ Save 1*' means that *Do Nothing* "ties" *Save 1*: both options are permissible in a pairwise choice. In Horton's case, we seem to have:

$$
\begin{array}{ccc}
 & & \textit{Save 2} \\
 & \sim \, \diagup & \\
 & & > \\
\textit{Do Nothing} \, \text{————} & \sim \, \text{————} & \textit{Save 1}
\end{array}
$$
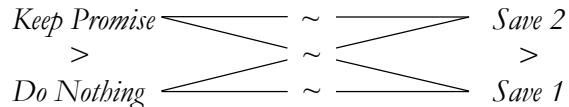
Where '*Save 2 > Save 1*' means that *Save 2* "defeats" *Save 1*: only *Save 2* is permissible in a pairwise choice (cf. Dietrich and List 2017). Kamm's case is isomorphic:

$$
\begin{array}{ccc}
\textit{Keep Promise} & & \\
> & \diagdown & \sim \\
\textit{Do Nothing} \, \text{————} & \sim \, \text{————} & \textit{Save 1}
\end{array}
$$

Indeed, the cases are built in the same way. We add an option that defeats *only one* of the classic pair. Kamm's addition beats *Do Nothing*; Horton's beats *Save 1*.

Intriguing links aside, the deeper reason to treat the puzzles together is that we need our solutions to cohere. To capture the intuitions in Kamm's or Horton's case, some philosophers propose revisionary principles of reasons and wrongness. But these proposals, I argue, can't be

extended to the combined case, where you have two ways to be heroic and two ways not to:

$$
\begin{array}{ccc}
\textit{Keep Promise} & \sim & \textit{Save 2} \\
> & \sim & > \\
\textit{Do Nothing} & \sim & \textit{Save 1}
\end{array}
$$

Other philosophers reject the intuitions in Kamm's and Horton's cases, citing a tension with certain familiar principles of reasons and wrongness. But these principles clash with our best solutions to the classic Paradox, which turn out to fit our intuitions rather nicely. If we just think through the best views about how supererogation is possible, we can debunk the principles that lead to paradox.

What makes supererogation possible is that wrongness doesn't just depend on worseness. Wrongness depends on multiple factors that are independent—for example, moral reasons and non-moral reasons. This independence can lead to some wrong acts (like *Save 1*) being no worse than permissible ones (like *Do Nothing*). It also leads to intransitivities of the kind described by Kamm, though in her case, I argue that the intransitivity arises from more than a multiplicity of factors; one factor must also be "comparative." I will later explain what this amounts to, and why I think comparativity is implicit in, or at least winked at by, our everyday views of supererogation.

We begin with the classic Paradox (§§2–3), then Horton's (§§4–5) and Kamm's (§§6–7). If all goes well, the reward is a coherent picture of how morality works beyond and beneath the call of duty, and the happy conclusion (§§8–9) is that supererogation isn't really paradoxical—just more interesting than we expected.

2. The Classic Paradox: Against 1D Solutions

The classic Paradox of Supererogation asks: if these fabulous deeds are really so *good*, why aren't they *obligatory?* The question has bite if we are drawn to:

> *The More Reason Principle*
> An option is obligatory if there is more reason to choose it than any alternative.

Which sounds hard to resist. Why permit worse choices? Or as Raz (1975: 165) asks: "How can one

3

be permitted to refrain from action that is required by reason?" And yet the More Reason Principle entails that anything less than best is wrong. Supererogation is impossible.[2]

To save the supererogatory, we need to replace this principle with something less stringent. Now, one strategy is to stick with plain reasons, and pick a fancier, more forgiving way of moving from reasons to rightness. Why not just do that?

It's worth a shot, but the fancy principles have their own problems. The main option is to say that we are only required to be "good enough," clearing a baseline of decency:

The Baseline Principle
An option is permissible just if it is better than the baseline. (Slote, 1985; Hurka, 1990)

Where "the baseline" is either an absolute standard (one's choice isn't *too* awful) or, like an average, relative to the set of one's options (one's choice isn't *too much worse* than the alternatives). Either way, this view allows for supererogation in simple two-option cases; *Do Nothing* might be permissible, even though *Save 1* is better, if both are above the line. The problem is that the Baseline Principle permits too much, since it permits acts that are above the line yet gratuitously bad. Although this problem predates Horton's puzzle (Bradley, 2006), his building case is a perfect example. It's wrong to save only one, given that saving two was no costlier, and yet *Save 1* must be above the line, since *Do Nothing* is. The Baseline Principle's permissions are too indiscriminate.

The two principles we have seen invite us to rank options from worst to best like numbers on a line. The More Reason Principle says anything below the top option is wrong; the Baseline Principle permits anything "high enough." But neither principle can rule out gratuitously worse options (like *Save 1*) while leaving in those that are justifiably worse (like *Do Nothing*).[3] To properly

---

    [2]   Some discussions of the paradox also link reasons to some other notion, like praise (Raz, 1975: 164) or goodness (Heyd, 2016: §3.1). By contrast, I am not making substantive claims about how value links up to reasons; I use 'better' only as a snappy way to say 'more reason' (see n.1).
    [3]   Another proposal is to assign each option an *interval* on the real line, not a point, and use this rule: an option is wrong just if its interval lies wholly below that of another option (cf. Gert, 2004: 505). But this proposal can't handle the combined case (see §1, above), where the options are {*Keep*

solve the classic Paradox, we need the resources to say *what it is* about the permissible suboptimal

acts that *makes* them permissible, something beyond raw betterness.

Solving the Paradox of Supererogation thus requires us to think of an action's moral status

as more than a number on a line. We need a second factor, something inherently permissive—and

this is exactly what we find in the two leading solutions to the Paradox of Supererogation.

3. The Classic Paradox: 2D Solutions

If supererogating is better, why permit anything less?

The leading answer is that the supererogatory act is only better in a respect. Supererogation

is favored by one kind of reason, but is opposed by weighty reasons of another kind, which are said

to be "justifying" rather than requiring (Archer, 2016; Lazar, 2019; cf. Gert, 2007), "partial" rather

than impartial (Parfit, 2011), or "non-moral" rather than moral (Slote, 1991; Portmore, 2011)—

whatever they are, they counterbalance the reasons to supererogate, stopping them from grounding

an obligation.[4] So we get something like:

> *The Non-Moral Reasons Principle*
> An option x is obligatory just if, for any alternative y, there is more moral reason to choose x
> than there is *combined* moral and non-moral reason to choose y.

Promise, Do Nothing, Save 2, Save 1}. This case intuitively violates:

> *Interval Order Property*
> If A > B and C > D, then either A > D or C > B.

And no relation that violates this property can be represented with intervals and the proposed rule
for determining wrongness (see Fishburn, 1970: 20–23; Rabinowicz, 2008: 33n.23).

[4]   For other 2D reasons-based views of supererogation, see Raz (1975), Dancy (1993a)
(critiqued in Postow, 2005), Bedke (2011) (critiqued in Snedegar, 2016), and Portmore (2019) (whose
views have changed since his *Commonsense Consequentialism*, in 2011). Some argue that supererogatory
beneficence is favored by "merely commendatory" reasons, which make options choiceworthy
without requiring them (Dreier, 2004; Horgan & Timmons, 2010; Little & McNamara, 2017). The
stock objection: this view permits us to let people die even when the cost is negligible (Dorsey, 2013:
361; Archer, 2016: 460). Horton's case is a special instance of this: *Save 1* seems wrong, but it would
be permissible if the reasons to *Save 2* were merely commendatory.

Where moral and non-moral reasons are two independent factors, and only moral reasons favor supererogating.

Already, this is enough to solve the paradox. Heroic sacrifices are optional, even though they are morally better, because we have mighty non-moral reasons not to harm ourselves. (This view does, however, require us to gently revise our concept of supererogation, since it's no longer better overall—just morally better. See §5, below.)

My own preferred view admits just one kind of reason—which tends to favor, justify, and require—and posits something further: *prerogatives*, which justify without at all requiring or favoring.[5] Prerogatives are purely permissive. They have weights, like reasons, but their sole function is to justify acts that would otherwise be wrong, putting a targeted gap between "most reason" and "must." They are like non-moral reasons, except they don't affect what's best.

This view also lets us replace the More Reason Principle. We say:

The Prerogatives Principle
An option x is obligatory just if, for any alternative y, there is more reason to choose x than there is *combined* reason and prerogative to choose y.

And this gives us an easy way to explain the optionality of self-sacrifice. On any view of prerogatives, I have a prerogative not to give up my limbs, and so even though the reasons favor self-sacrifice, it lies well beyond the call of duty.[6]

These are the leading solutions in the literature, and I can't stress enough that they both involve the idea that right and wrong depend on two dimensions. There isn't just a cutoff along a single scale from naughty to nice. Moreover, on either view, the two dimensions are independent, in

---

[5]   Hurka and Shubert (2012) prefer "prima facie permissions." ('Prerogative' is from Scheffler, 1982; Slote, 1984; and Kamm, 1996). But just as nothing turns on how we define 'supererogation', nothing turns on how we define 'reason' and 'prerogative'; the substantive ideas here are justifying, favoring, requiring, and permissibility.

[6]   Some say we have a basic prerogative to do what's prudent (Scheffler, 1982; Slote, 1984; Hurka & Shubert, 2012); others derive prerogatives from rights against harm (e.g. Benn, 2017; I am also happy to derive prerogatives from other rights, like body and property rights; see Muñoz ms.).

the sense that an option's ranking along one doesn't tell us how it ranks with respect to the other. The weights of the moral reasons don't tell us how the non-moral reasons lie; the balance of reasons doesn't settle which prerogatives we have; and so on (cf. Lazar, 2019: 246).

The crucial upshot of this independence is that an option can outrank its rival on one dimension while being outranked on the other. Indeed, this sort of moral mismatch is just what we would expect to find in our simple case of supererogation. Self-sacrifice is a wonderful thing, but a harm to oneself can have special significance. *Do Nothing* thus gets protected by the weightier prerogative (or non-moral reasons), even though *Save 1* is (morally) better.

We need two dimensions to make sense of supererogation, and having two dimensions can lead to moral mismatches. That is the main lesson of the Paradox of Supererogation. It's a simple point, but even on its own, it is powerful enough to solve the All or Nothing Problem.

4. All or Nothing: Reasons and Prerogatives

We begin by stating the problem more carefully.

In Horton's case, I am faced with a collapsing building and three options. I can do nothing, which costs me nothing. Or I could sacrifice my arms to hold open a small escape route, saving one of the people inside. I could also, for the same blood price, open a bigger route and save both.

Intuitively, I do not have to be the hero.

May Save None
It is permissible for me to save no one.

After all, two arms would be a massive sacrifice. But it would be wrong to put that sacrifice to waste, saving only one person and abandoning the other:

Wrong to Save One
It is wrong for me to save just one.

There is no way to justify such gratuitous badness. Still, saving one is no worse than saving zero.

No Worse to Save One
I don't have more reason, all things considered, to save no one than I have to save just one.

Morality wouldn't recommend saving fewer lives.

But as nice as these intuitions are, they leave us with a wrong act (saving one) that is no

worse than a permissible act (saving zero). And surely:

Worse to Do Wrong
If A is wrong and B is permissible, then I have more reason to do B.

Indeed, it is a "common assumption" in the literature that any permissible act is better than any

wrong alternative (Pummer, 2019: 284n.19).[7]

That is the All or Nothing Problem: if the only permissible options are saving all of the

people and saving none, then we must revise our intuitions, or else deny the seemingly undeniable—

Worse to Do Wrong.

Most solutions give up an intuition. Horton (2017) denies May Save None (he says it is false

when the agent is willing to make the sacrifice);[8] Portmore (2019: §5.4) denies that there is more

moral reason to save one than none; and McMahan (2018: 99–100) suggests that saving one might

be wrong and yet somehow not impermissible. Although these concessions are costly, they would

allow us to preserve Worse to Do Wrong.

But that principle is seriously undermined—and the intuitions are strongly predicted—by

the idea of independent dimensions, which we got from working through the Paradox of

Supererogation. Wrong acts *would* always be worse, if 'wrong' meant 'below the cutoff on the one

---

[7] Worse to Do Wrong is endorsed in some form by Darwall (2013), Ferguson (2016: 959n.23),
Horton (2017: 96) (who says we "ought rather" to do any permissible act over any wrong one),
Ferguson and Köhler (2019) (as "reasons monotonicity of permissibility"), and Portmore (2019: 22)
(who thinks the moral reasons must always favor a morally permissible act over a wrong alternative,
or else permissions couldn't depend on reasons). Two important exceptions are Bader (2019) and
Pummer (2019). For discussion of cases like Horton's, see Parfit, 1982; Kagan, 1989: 16; Tadros,
2011: 161–62; Portmore, 2011: 147; and Snedegar, 2015: 379. Barnett (ms.) discusses some
fascinating examples that make trouble for the analogue of Worse to Do Wrong in decision theory.
[8] Sinclair (2018: 49) appears to share Horton's view of the building case.

true moral scale from bad to good'. Any permissible act would be above the line and therefore better than anything below. But what if there are two scales? Then we get counterexamples. Let's see why, starting with a view of reasons and prerogatives.

Recall the basic pair: *Do Nothing ~ Save 1*. (Where '*x ~ y*' means that *x* and *y* are both permissible in a pairwise choice.) Given that *Save 1* is better than *Do Nothing*, we might expect *Save 1* to be obligatory. But it isn't. We have prerogatives not to harm ourselves, so *Do Nothing* is justifiable despite being suboptimal. The reason to save a life doesn't outweigh the combined reasons and prerogatives to self-preserve.

Now the key question. What should we expect to happen when we add the option to *Save 2?* Consider how it fares in pairwise choices. Intuitively, *Save 2* is obligatory in a pairwise choice with *Save 1: Save 2 > Save 1*. The cost is the same either way, so in effect, the agent is just choosing whether to pointlessly allow the second person to die. I assume that pointlessly letting someone die is wrong. Whatever prerogatives we have, they don't make a decisive difference in this kind of choice, where no additional harm to self (or weighty rights) are in play.

But our prerogatives do permit *Do Nothing* in a choice with *Save 2. Do Nothing ~ Save 2.* A massive harm to oneself can allow us to forgo a whole range of sacrifices; it is not as if anything beyond one life has got to tip the scales. I assume that two lives won't be the tipping point.

So in the choice from {*Do Nothing*, *Save 1*, *Save 2*}, we should expect only one wrong option: *Save 1*. And there is a very natural way to get this from reasons and prerogatives:

	Do Nothing	Save 1	Save 2
Reasons	1	5	10
Prerogatives	10	0	0

The result is that wrongly saving one is better than permissibly doing nothing. And this is possible because of reasons and prerogatives are independent. *Do Nothing* is worse than *Save 1*, but I have a mighty prerogative not to sacrifice, and only a negligible prerogative to sacrifice to save the one; I can't justify *Save 1* over purely better rescues. That is why the addition of *Save 2* is able to rule out only the better of two options. "Tied" options don't have to be morally equivalent, and they don't have to be defeated by the same things. Even the lesser option in a tie can be more resiliently permissible.[9]

So here is my main claim about the All or Nothing Problem. If we start with reasons, prerogatives, and pairwise judgments, we should expect that wrongly saving one will be better than permissibly saving no one. It is a prediction, not a paradox, that the three intuitions are true and that Worse to Do Wrong is false. We can solve the All or Nothing Problem without compromising on our intuitive judgments, all thanks to the insights we got from working through the Paradox of Supererogation.

5. All or Nothing: Parity of Reasons

But what about reasons-based views of supererogation? Do they make the same predictions about Horton's case? It's a bit complicated, but mostly: yes.

For simplicity, let's stick with the Non-Moral Reasons Principle, which says that supererogation emerges from a clash of moral and non-moral reasons: an act is obligatory *iff* the moral reasons to do it outweigh the combined moral and non-moral reasons against. Given this principle, we would expect *Save 2* to be morally best, followed by *Save 1*, and we would expect both of these sacrifices to be non-morally worse than *Do Nothing*. So the prediction is that *Save 1* will be

---

[9] This case also shows that we can't recover the facts about what's permissible from the *sum* of reasons and prerogatives in favor of each option. (*Do Nothing* has a higher sum than *Save 2*, but only *Save 2* makes *Save 1* wrong.) Prerogatives and reasons play different roles; only reasons can *require*.

wrong and yet morally better than permissibly saving no one. This is violation of what we might call "Morally Worse to Do Wrong," which is the version of Worse to Do Wrong that is in fact endorsed by fans of non-moral reasons (Portmore, 2019: 22).

But can this view accommodate our three intuitions? On the face of it, no. The trouble is with No Worse to Save One. Since *Do Nothing* is permissible, we know that it is at least as good overall as *Save 2*, which is strictly better than *Save 1*. Don't we have to conclude that *Do Nothing* is also better than *Save 1?*

This is a well-known problem.[10] But there is a solution, due to Derek Parfit: we say that *Do Nothing* is on a par with *Save 1* and *Save 2* (2011: 137–141). Two options are on a par when they are comparable, though neither is better, and they aren't exactly equal in goodness (Chang, 2002). Using '~' to denote parity and '>' plain betterness, we get a familiar structure:

$$
\begin{array}{ccc}
 & & Save\ 2 \\
 & \sim & > \\
Do\ Nothing & \sim & Save\ 1
\end{array}
$$
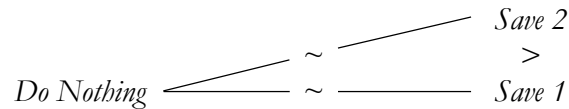
So *Save 1* is wrong because it's strictly worse than *Save 2*, but *Save 1* is still on a par with *Do Nothing*, which remains permissible because it is the best non-morally. With parity, I conclude, we can capture all three intuitions in the All or Nothing Problem and debunk Worse to Do Wrong.[11]

Indeed, we can do all this even if we don't use moral and non-moral reasons in particular. We might instead say that the reasons to save lives are "other-regarding" while the reasons to do nothing are "prudential." Or we could call it a conflict between "impartial" reasons (like *it would save two lives*) and "partial" reasons (like *it would save* my *arms*). All good. So long as the reasons to sacrifice

---

[10]   See Kagan, 1991: 927–28; Portmore, 2011: Chapter 5; Hurka & Shubert 2012: n.7.

[11]   Indeed, even non-moral parity makes trouble for Worse to Do Wrong. For example, start with *Tea ~ Coffee*, then add a "sweetened" third option: *Coffee+ > Coffee*. Plausibly, if *Coffee+* is just a slight improvement, we have *Tea ~ Coffee+*. But that means a wrong option, *Coffee*, is on a par with (and therefore no worse than) a permissible option—*Tea*. In such cases, parity is "insensitive to mild sweetening" (Hare, 2010). Sweetening is a recipe for counterexamples to Worse to Do Wrong.

are on a par with the reasons to self-preserve, we can solve the All or Nothing Problem.[12]

6. The Intransitivity Paradox

Now the twist: we have *already* solved Kamm's Intransitivity Paradox. (And the bucket of cold water: it's not entirely obvious what "Kamm's Paradox" amounts to. We'll clear this up as we go.)

The core of the puzzle, as originally presented, is that some cases appear to feature a spooky intransitivity. In Kamm's case, I have three options: keep a promise to meet someone for lunch, save a life at great personal cost, or do nothing.[13] Because life-saving is costly, I am free to do nothing instead: *Do Nothing ~ Save 1*. (Where this "tie" means that I may do either option in a pairwise choice.) But since the sacrifice is for a splendid end—saving a life—it is permissible even when it entails breaking lunch obligations: *Save 1 ~ Keep Promise*. And yet I may not skip lunch to lounge around: *Keep Promise > Do Nothing*. (This "defeat" means that I would have to keep the promise in a pairwise choice.) All together:

$$\begin{array}{ccc}
\textit{Keep Promise} & & \\
> & \diagdown & \sim \\
\textit{Do Nothing} & \underset{\sim}{\rule{2cm}{0.4pt}} & \textit{Save 1}
\end{array}$$
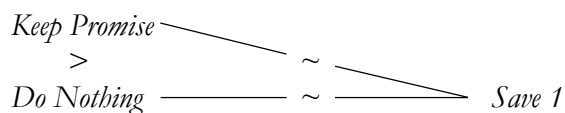
Here, we get an intransitivity in the relation of *weak defeat*, or '$\gtrsim$'.[14] Let '$x \gtrsim y$' mean that x is permissible in the pairwise choice. ('$x \gtrsim y$' is equivalent to '$x > y$ or $x \sim y$'.) Kamm's intuitions about her case entail that '$\gtrsim$' is intransitive: *Do Nothing $\gtrsim$ Save 1 $\gtrsim$ Keep Promise > Do Nothing*. This means that we have *Do Nothing $\gtrsim$ Save 1 $\gtrsim$ Keep Promise*, but not *Do Nothing $\gtrsim$ Keep Promise*. That's Kamm's intransitivity.
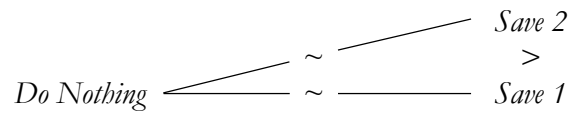
---

[12] One limitation of the parity view, however, is that, since *Save 1* and *Do Nothing* are on a par, we can't say that *Save 1* is better. This is one reason why I prefer prerogatives to non-moral reasons.

[13] Kamm's (1985) hero gives a kidney rather than losing arms. I ignore this cosmetic difference.

[14] Tying is also intransitive. Indeed, if weak defeat is intransitive, complete, and reflexive, then either trying or defeat must be intransitive, though not necessarily both (see n.23). Kamm (1985) herself doesn't use 'tie', and her term for weak defeat is 'may permissibly take precedence over'.

The air of paradox creeps in when we ask: how is this possible? If x is tall enough to at least tie with y in a height contest, and y is tall enough to at least tie with z, then x will at least tie with z. Why shouldn't the same hold for contests of permissibility? If x is at least good enough to choose over y, and y over z, doesn't x also have to be good enough to choose over z?[15]

But we have *already* shown that this kind of question is confused, because being permissible, unlike being tallest, is a matter of two independent dimensions—reasons and prerogatives. Indeed, we have just seen how reasons and prerogatives lead to intransitivity in Horton's case:

$$\begin{array}{ccc} & & \textit{Save 2} \\ & \sim \diagup & \vee \\ \textit{Do Nothing} \diagdown \quad \sim \quad & \text{———} & \textit{Save 1} \end{array}$$

Save 1 ≳ *Do Nothing* ≳ *Save 2*, but not *Save 1* ≳ *Save 2*. The key is a moral mismatch. I have reasons to help others and prerogatives not to self-harm. So, I have a greater prerogative to *Do Nothing* despite having more reason to *Save 1*. (*Mutatis mutandis* for views with multi-flavored reasons.)

The fear of transitivity, like the pull of Worse to Do Wrong, depends on the assumption that right and wrong depend on only one dimension—a dogma that would rule out supererogation from the start. We should be glad to get rid of Worse to Do Wrong. For the same reason, we should enjoy our polite farewell to the transitivity of weak defeat.

7. Comparative Prerogatives

Now things get tricky. Kamm's Paradox is kaput; there is nothing fishy about intransitivity per se. But we are still in a pickle insofar as we haven't figured out Kamm's *case*, which is trickier than Horton's. Maybe this isn't part of the paradox proper, since it's not strictly about transitivity (this is

---

[15]   Archer (2016: 445) and Portmore (2017: 294) bring out the oddness in terms of reasons. If the reasons in favor of x can justify acting against the reasons for y, and same for y's reasons against z's, then why can't the reasons in favor of x justify acting against the reasons in favor of z?

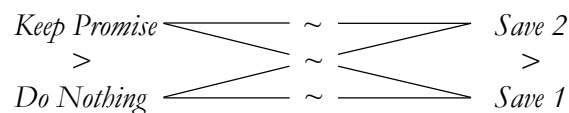the unclarity alluded to earlier). Still, we need a solution.

Here's the problem. Because *Do Nothing* ⪆ *Save 1*, we have to say that there is a mighty prerogative to *Do Nothing*. But clearly, in the choice between *Do Nothing* and *Keep Promise*, there can't be such a prerogative. If there were, promise-breaking would be permissible—which it isn't. It's a datum that *Keep Promise* > *Do Nothing*.

An essential part of this problem is that we can't just say that *Keep Promise* rules out *Do Nothing* by virtue of its being even better than *Save 1*. Keeping the promise isn't better than giving up one's arms to save a life, as shown by the fact that saving the life is supererogatory. The core mystery is that the prerogative to *Do Nothing* seems to kick in against *Save 1*—the best option—but disappear up against *Keep Promise*.

This is really puzzling. What's going on?

Kamm (1996: 313) proposes an answer: we are dealing with two independent dimensions. *Save 1* scores highest in one respect—it may "take precedence" over the biggest swatch of options, weakly defeating them—whereas *Keep Promise* ranks highest in another: it requires the greatest efforts. The idea seems to be that *Keep Promise* is urged by the bossier moral reasons, but *Save 1* is protected by the bigger prerogative (for a similar interpretation, see Lazar, 2019: 246). Clearly, Kamm's view will have trouble capturing the betterness of *Save 1*. But a deeper problem is that it makes the wrong predictions about Horton's option to *Save 2*.

Recall the combined choice:

$$
\begin{array}{ccc}
\textit{Keep Promise} & \sim & \textit{Save 2} \\
> & \sim & > \\
\textit{Do Nothing} & \sim & \textit{Save 1}
\end{array}
$$

Now, if Kamm is right that we have a mighty prerogative to save lives, and only relatively weak moral reasons to do so, then the justification to *Save 1* should be enough to outweigh the reasons to *Save 2*, and *Save 1* will be permissible.

Here's why. We know that there is some justification for *Keep Promise*, since *Keep Promise* ≳ *Save 1*. We also know that there is even *more* justification for *Save 1*; that is Kamm's main idea. But then adding *Save 2* can't make *Save 1* alone wrong. Since *Save 1* is even more justifiable than *Keep Promise*, anything that makes *Save 1* wrong will also rule out *Keep Promise*. But *Save 2* doesn't rule out *Keep Promise*. So it can't rule out *Save 1*, either. Kamm's view thus conflicts with the idea that it's wrong to save one in Horton's case. The view lets us be gratuitously nasty.[16]

The other leading take on Kamm's case, due to Alfred Archer (2016: 459), is that it involves *three* dimensions. Basically, Archer's view is just Kamm's plus the idea that an option can be "morally favored" independently of how strongly it is justified or required. This allows Archer to say that *Save 1* is more favored than *Keep Promise*. Now, like Portmore (2016: 293n.11), I think it's odd that "favoring" should be independent of other factors. But even granting Archer this, he hasn't solved the fundamental problem for Kamm's view: his view also permits me to *Save 1* even given the option to *Save 2*. Neither Kamm's view nor Archer's can be plausibly extended to Horton-style cases, where there are better and worse ways to save.

Now what? If we can't treat Kamm's case even with three dimensions, what on earth could do the trick? A leap to four dimensions? A formal *deus ex?*

I don't think we need anything so fancy. Just ask yourself: why am I free to stay out of the building, but not free to break the promise? The answer is obvious: keeping the promise wouldn't crush my arms! Promise-keeping isn't more harmful to me than doing nothing, so I don't have any harm-based prerogative to *Do Nothing* rather than keep my promise.

---

[16] Kamm (1996: 336) is sensitive to the worry that we may have to save lives, noting that the proposed view does not explain "why it is obligatory to save a life at small cost to ourselves rather than keep an important business appointment." She then suggests, roughly, that costless lifesaving is supported by stronger moral reasons than promise-keeping. I'm not sure that this will help in the four-option case, since *Save 2* isn't costless, and yet it is favored by strong moral reasons.

The big mistake was to think that we could always represent reasons and prerogatives with a single weight. Sometimes, as in Kamm's case, it matters what we are comparing to what. I have a prerogative to *Do Nothing* rather than *Save 1*, because saving the life is more harmful to me; but I don't have any such prerogative to *Do Nothing* rather than *Keep Promise*, since the harm to me is the same either way. What matters is comparative harmfulness. That is why we can't just say that *Do Nothing* gets a prerogative of weight 5, or weight 10, or what have you. The weight doesn't just depend on how harmful the act is considered in itself; we also need to know the alternative. Prerogative are in this sense "contrastive" (Snedegar, 2015), or as I like to put it, *comparative:* we have a prerogative to do x rather than y if x harms us *less* than y does.

The key to Kamm's case is that prerogatives (or non-moral reasons) are comparative.[17] Stated in the abstract, this might sound like an ad hoc addition. But it's not. It is just common sense that our prerogatives depend on comparative harms, and there is nothing natural about the idea that the costs of heroism should justify one non-heroic option over another. This idea is just confused, as Kamm's case shows, and the solution is that harms matter in a comparative way.

Let me close with one more case, introduced by Dale Dorsey (2013), which has been treated as a variant on Kamm's.[18] Suppose I am able to beat up a stranger, Roy, and take his money, which I can use either to buy a car for myself, or to save ten lives via an amazing NGO. I have three options: *Get a Car*, *Respect Roy's Rights*, or *Save 10*. Intuitively, Dorsey thinks, I don't have to give up my own luxury for charity: *Get a Car* ~ *Save 10*. (A bit heartless, but let's grant it.) I may, however, beat up Roy when lives are on the line: *Respect Roy's Rights* ~ *Save 10*. And yet I would not be justified in

---

[17] Let me be clear: I am not saying that prerogatives are "essentially comparative" in Temkin's sense (2012: 371). That would mean that the weight of the prerogative to do x over y is not a function of the options' intrinsic properties even taken together. But for all I have said so far, the prerogative's weight might be a function of how much x and y harm the agent.

[18] See Archer, 2016; Portmore, 2017. Dorsey doesn't himself mention transitivity or Kamm's paper. I have cleaned up his cases so that they fit into a single three-option choice.

assaulting someone just for the sake of new wheels: *Respect Roy's Rights > Get a Car.*

We get the same old structure:

$$Respect\ Roy's\ Rights$$
$$>\qquad\qquad \sim$$
$$Get\ a\ Car\ \qquad\qquad \sim\ \qquad\qquad Save\ 10$$
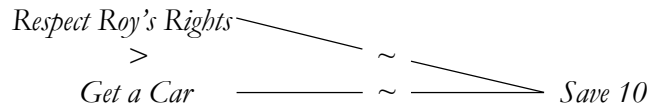
This time, however, we can't say that the two non-lifesaving acts involve the same null cost to self. Respecting Roy costs me a whole car! So why is my prerogative to *Get a Car* comparative? Why does it count for so much against the greater good and for so little against Roy's rights?

Again, I don't think the explanation here has to be ad hoc. On the contrary, the verdicts follow from two familiar views of prerogatives. First, on Hurka and Shubert's view, we have a prerogative to promote our own interests *except* if doing so would violate a right. They argue:

> Surely if it's wrong to kill one innocent person to save two others, as deontological moralities hold, it's also and equally wrong when one of the two is you: your agent-favoring permission to pursue your own good must have more weight against the duty to promote the good than it does against deontological constraints. (2012: 10; cf. Kamm, 2007: 30–1)

On a rights-based view, meanwhile, it seems clear that a right to spend on oneself doesn't entail a right to steal from others, much less beat them up!
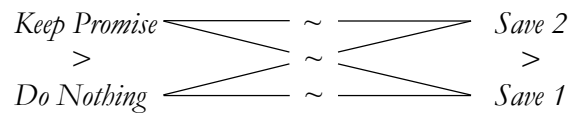
Both views entail that we have no prerogative to violate rights, even self-interestedly. We don't have to agree with this idea. But if we do, then we can explain the presence of a comparative prerogative in Dorsey's case, which turns out to be more complex than Kamm's. Instead of a difference in marginal costs, we find a difference in the significance of the same marginal cost of one shiny new car. The prospect of new wheels does little to justify stealing, but seems to do a fair bit to justify spending on oneself over strangers. This is the same comparativity as in Kamm's case—only with a different source.

8. Menu Relativity

We have finally hacked our way through the paradoxes of supererogation. Let's recap.

The fundamental idea is that we build our way up to the outré many-option choices from what we already know about simple choices between pairs. We start with the classic Paradox: why may we choose a morally subpar option? The answer is that it is justified by a second factor—prerogatives (or non-moral reasons)—which is independent of moral reasons and inherently permissive. That gives us the Prerogatives Principle: we may choose an option when the reasons and prerogatives in favor aren't outweighed by the reasons to do any alternative.

We then use this this principle to generate pairwise judgments about our full set of options.

$$\begin{array}{ccccc} \textit{Keep Promise} & & \sim & & \textit{Save 2} \\ > & & \sim & & > \\ \textit{Do Nothing} & & \sim & & \textit{Save 1} \end{array}$$

This instantly yields Kamm's intransitivity: *Do Nothing ~ Save 1 ~ Keep Promise > Do Nothing.* We even find the same intransitivity with Horton's three options: *Save 1 ~ Do Nothing ~ Save 2 > Save 1.*

Moreover, the Prerogatives Principle helps us debunk Worse to Do Wrong. The reason why *Save 1* is wrong, even though it is better than permissibly doing nothing, is that I have a prerogative to *Do Nothing* rather than save lives, whereas I have no prerogative to *Save 1*. A prerogative to keep my arms doesn't entail a right to sacrifice them however I like. By that same token, a prerogative to keep my arms doesn't entail a prerogative to keep them however I like. I have no prerogative to *Do Nothing* rather than keep my promise. Prerogatives depend on comparative harms: if two options are equally harmful to me, I have no harm-based prerogative to pick one over the other. That's why I may not *Do Nothing* rather than *Keep Promise*, just as I may not *Save 1* rather than *Save 2*.

The result is a unified and minimal solution to the paradoxes. It uses the basic ideas that make supererogation possible, along with some easy pairwise judgments, to debunk transitivity and Worse to Do Wrong—no exotic extras, no controversial apparatus.

Now I want to clarify something. My solution *sounds* like something very controversial. It sounds like I am saying: when we add new options, we change the balance of reasons and

prerogatives (cf. Kamm, 2007: Chapter 1, n.55). This idea has been defended by Lazar and Barry

(ms), who think that adding *Save 2* to {*Do Nothing, Save 1*} generates a new moral reason against *Save*

1, which makes it morally worse than *Do Nothing*. There's definitely something to this idea. (Isn't it

bad to gratuitously let someone die? Doesn't the presence of the second child make *Save 1* into a

gratuitous letting-die?) But I think changing reasons are hard to defend, and not essential to solving

our paradoxes.

On Lazar and Barry's view, reasons "change" in the sense that they depend on the menu:

Menu Relativity of Reasons and Prerogatives
The weight of the reasons and prerogatives to do *x* rather than *y* may differ depending on
the availability of a further option *z*.

In particular, there isn't more reason to *Save 1* rather than *Do Nothing* in a pairwise choice, but there

is when *Save 2* is on the menu. The presence of *Save 2* makes *Save 1* better than *Do Nothing*. That is

why *Save 1* is permissible in the pairwise choice but wrong with *Save 2* on the menu.

Again, this *sounds* like my view. I have argued that *Do Nothing* compares more favorably to

Save 1 in the presence of *Save 2* (and *less* favorably, in the presence of *Keep Promise*). Pummer agrees:

it is a familiar feature of nonconsequentialist ethics that the moral status of an act can
depend on which alternative acts are available. In this case, the presence of [*Save 2*] alters the
moral status of [*Save 1*], thereby altering the way that [*Save 1*] and [*Do Nothing*] compare
morally. (2016: 86–7)

Isn't this just Menu Relativity? Aren't we saying that *Save 2* makes *Save 1* worse, and that this

explains why *Save 1* becomes wrong?

No. We already have a simpler explanation. *Save 2* makes *Save 1* wrong, but leaves *Do Nothing*

permissible, because only *Do Nothing* is protected by a prerogative. We don't need to posit changing

reasons to capture this; it's implicit in defeat relations:
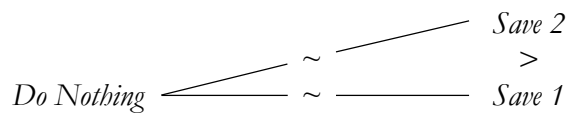
Because *Save 2* > *Save 1*, we know that the reasons to *Save 2* outweigh the reasons and prerogatives

to *Save 1* in a pairwise choice. Assuming that this doesn't change when we add the option to *Do Nothing*, *Save 1* will be wrong whenever *Save 2* is on the menu, though *Do Nothing* needn't be.

Adding options, however, clearly does have *some* effect. Even in my view. the presence of *Save 2* alters the way in which *Save 1* "compares morally" to *Do Nothing*. But rather than a change in reasons and prerogatives, we have a divergence in deontic status. Only one option is made wrong by the new addition. This seems to be how Pummer sees the case, too:

> …with the full choice situation in view, it is clear that there is something to be said against [*Save 1*] that cannot be said against [*Do Nothing*] or [*Save 2*]: the performance of [*Save 1*] constitutes a deliberate refusal to do something much better at no extra cost. This is a serious moral failing. (2016: 86–7)

The failing consists in how *Save 1* compares to *Save 2*, not in how it directly compares to *Do Nothing*. Only *Save 1* is gratuitously worse than an alternative; it alone is *defeated*. Similarly, in Kamm's case, only *Do Nothing* becomes wrong because it alone is defeated by *Keep Promise*. This explanation doesn't require any Menu Relativity.[19]

So we haven't used, and don't need, Menu Relativity to handle Kamm and Horton's cases. All we need are commonsense reasons and prerogatives—which we already needed to solve the classic Paradox. By contrast, Menu Relativity can play no role in that Paradox, since it concerns a single fixed menu, {*Do Nothing, Save 1*}.

That's why it's important to distinguish my view from the Menu Relativist's. Not only are the views different: mine is more unified. I am trying to solve a whole cluster of puzzles using only the familiar idea of a prerogative not to harm oneself, and I don't appeal to menu-relative reasons, which makes the view much less controversial than it might have sounded.

---

[19] Interestingly, in our cases, it *does* seem menu-relative whether two options have the same deontic status. For example, *Do Nothing* and *Save 1* are both permissible in a pairwise choice, but only *Do Nothing* may be chosen from {*Do Nothing, Save 1, Save 2*}. I don't deny this. When I say I reject "Menu Relativity," I am only denying that *z*'s presence can change the *reasons and prerogatives* to do *x* over *y*. Our two factors are independent and comparative, but not themselves menu-relative.

But why is Menu Relativity so controversial? And why isn't my view, with its comparativity and intransitivity, just as fishy? As it turns out, the Menu Relative view has much stronger—and stranger—implications. To see why, we need to slow down and talk about what are sometimes called principles of "consistency" (Sen, 1993) or "rationality" (Sen, 2017: 63). A more neutral name for them would be "principles of menu-insensitivity." They require options to keep their deontic properties even as we alter the rest of the menu.

Our cases, as I've described them, respect two key principles of menu-insensitivity. First:

Property α (Basic Contraction Consistency)
If x is permissible to choose from a set of options S, then x is permissible to choose from any subset S* of S to which x belongs. (Sen, 2017: Chapter 1*6)

Property *a* says that a permissible option stays permissible even if we prune other options from the menu.[20] This is how things work in Kamm's and Horton's cases, as well as the combined case. Nothing is ever made *wrong* by the removal of an alternative; *Save 2* doesn't become wrong when we take away the option to *Keep Promise*. The cases also obey:

Property γ (Basic Expansion Consistency)
Given some sets of options S_i, if x is permissible to choose from any S_i, then x is permissible to choose from the union of all S_i. (Sen, 1993: 500)

For instance: *Save 1* is permissible in pairwise choices with either *Do Nothing* or *Keep Promise*, and it remains permissible in the three-way choice.

Now here is the kicker. Properties *a* and *γ*, though consistent with my view, are deeply in tension with Menu Relativity. The upshot of the properties is that we may derive the facts about what's permissible from facts about what defeats what. More precisely: the properties hold *just if* the

---

    [20]   Property *a*, also known as "The Independence of Irrelevant Alternatives" (IIA), is often mixed up with nearby ideas. Kamm (1996: 344) confuses IIA with its homophone in social choice theory (which says: society's preference over a pair depends only on how individuals rank that pair; see Arrow, 1951, Sen, 2017: Chapter 3). Rulli and Worsnip (2016) define IIA in a way that entails Property *β*.

permissible options in any subset S* of S are exactly those that weakly defeat everything in S*.[21] If either property is violated, then the permissibility facts are so menu-sensitive that we would need Menu Relativity to explain it. Consider an a-violation: x is permissible to pick from $\{x, y, z\}$ but wrong from $\{x, y\}$. Do the reasons to do y outweigh the reasons and prerogatives to do x? The answer has to be: it depends. The reasons for y don't rule out x in $\{x, y, z\}$, but they do in $\{x, y\}$, so z's presence must make x more justifiable over y. Similarly, we need Menu Relativity to understand cases where Property γ doesn't hold. In fact, it is *only* in these two kinds of cases that we need Menu-Relativity. So long as Property a and Property γ are obeyed, we can find some way to derive the judgments about what's permissible from a possible array of non-relative reasons and prerogatives.

This is bad news for the friend of Menu Relativity. The best kind of evidence for their view is missing in our examples, where there is no violation of α or γ in sight. Moreover, given their view, we should *expect* to find violations of both of these principles, somewhere or other. That is a significant cost, especially in the case of a, since it is quite plausible that pruning the menu shouldn't make an option wrong. As Vallentyne and Tungodden (2005: 143) ask: "If [x] is a winner against [y] and z, why would it not also be a winner against y alone?"[22]

Our examples do exhibit one kind of menu-sensitivity. Rather than offending against α or γ, they infringe on their less legendary cousin:

---

[21]   I assume that, for any subset S* of our (finite!) set of options S, there is an option that would be permissible in a choice from S*; in such conditions, we say that there is a choice function f defined over S; α and γ are properties of f necessary and sufficient for f to be generated from a non-menu-relative relation '$\gtrsim$', such that $x \in f(S^*)$ *iff* $x \in S^*$ & $\forall y \in S^*$, $x \gtrsim y$ (Sen 1993: 499–500).

[22]   There is also Morgenbesser's joke (adapted from Kamm, 1996: n.31):

BARTENDER: Would you prefer red or white wine?
PATRON: White, please.
BARTENDER: I forgot: we also have beer.
PATRON: In that case, I'll take the red.

For some potential counterexamples to α, see Pettit, 1991; Sen, 1993; Kamm, 1996: 343.

Property β
If x and y are both permissible to choose from a set of options S, then if x is permissible to choose from a superset S* of S, so is y. (Sen 2017: Chapter 1*6)

$β$ says: if two options are both permissible, then adding an option can't make only one of them wrong. Equally permissible options are beaten by the same things. Clearly this isn't how things work in our paradoxes, where *Save 2* makes only *Save 1* wrong, as *Keep Promise* does to *Do Nothing*.

Is this evidence for Menu Relativity? Interestingly, no. When $β$ fails, the upshot isn't any kind of relativity—it's *intransitivity!* Property $β$ is equivalent to what I call:

Transmission Over Ties
If $x > y \sim z$, then $x > z$.[23]

And this is exactly the principle to which Kamm's and Horton's cases are counterexamples. Recall:
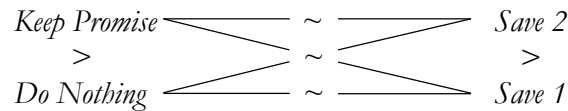
Here we have a smorgasbord of Transmission failures. (For example: *Save 2 > Save 1 ~ Do Nothing ~ Save 2*.) And it is easy to see why defeat doesn't transmit across these tied options. Since we are dealing with two dimensions, the tied options are relevantly different; there is more reason to *Save 1*, but a much greater prerogative to *Do Nothing*. When reasons and prerogatives are mismatched, tied options can be defeated by different things, which means that $β$ fails and intransitivities abound.

This is a rather nice result. My approach to the paradoxes, based on the Prerogatives Principle, fits with the failure of $β$, and unlike Menu Relativity, it doesn't conflict at all with the more plausible a and $γ$. Why no conflict? Because the approach is all about pairs. To figure out if x is permissible, just ask, for each alternative y: do the reasons to do y outweigh the combined reasons

---

[23] This equivalence holds given that a and $γ$ obtain and a choice function is defined (see n.21). Sen (2017: 66) elegantly explains the relation between Transmission Over Ties and transitivity. Transmission is strictly weaker than the transitivity of '$\gtrsim$', which is equivalent to the conjunction of Transmission with the transitivity of '$>$'. (Assuming that '$\gtrsim$' is complete and reflexive.)

and prerogatives to do *x* instead? If the answer is always "no," then *x* is permissible. If we run this test in Kamm's and Horton's cases, we get the exact answers we wanted simply by considering the reasons and prerogatives at work in pairwise choices; this leaves us with no need to posit Menu Relativity. All we need to give up is Property *β*—and its equivalent, Transmission Over Ties—which is no great loss. Any view that vindicates our intuitions will have to lose *β* and Transmission. The goal is to give up *only* these principles, and keep their more plausible relatives.

In that same vein, just as my approach lets us keep *a* and *γ*, it avoids the more scandalous kinds of intransitivity. In particular, we may hang on to:

Acyclicity
If $x > y > \ldots > z$, then it's not true that $z > x$.

What a relief. The possibility of cycles is legendarily controversial (cf. Temkin 2012), since they leave the agent with no undefeated option. Transmission failures entail no such moral doom.

Menu-Relative reasons are unnecessary for treating Horton's case; they are costly to posit; and they aren't motivated by any more basic ideas about supererogation. That is why I prefer a solution that uses non-relative (but still comparative!) reasons and prerogatives.

Let me note one last advantage. My view gives a unified account of both Horton's and Kamm's cases. Menu Relativity can at best explain Horton's. On Lazar and Barry's view, adding the option to *Save 2* makes *Save 1* morally worse vis-à-vis *Do Nothing*, as *Save 1* now involves letting the second child die for no reason. That strikes me as a reasonable proposal. But we can't extend it to Kamm's case. Throwing *Keep Promise* into the mix does make *Do Nothing* worse, but it also makes *Save 1* worse. Since both options become ways of breaking the promise, they are worsened equally. So what explains why *Keep Promise* makes only *Do Nothing* wrong? If we really wanted to stick with Menu Relativity, we could try to cook up a new answer. But I think we should instead go back to our Supererogation ABCs: take each option pair by pair, weigh up the reasons and prerogatives, and watch as the paradoxes unravel.

9. Conclusion

There are two kinds of objection to the possibility of supererogation (Dancy, 1993b). The first is that the very concept of supererogation is confused or problematic. The second is that the concept, however coherent, is empty, because morality wouldn't allow anyone to do less than best.

My concern here has been the conceptual objections—the paradoxes. I have argued that these shouldn't scare us, and that we can construct an elegant account of supererogation using nothing more than two dimensions—reasons and prerogatives, flavors of reasons—so long as we allow for comparativity. We don't need baselines, yet more dimensions, Menu-Relative reasons, or cycles of defeat. The solution is simple: you must do best unless you have a prerogative not to. You have to save lives, for example, it harms you much more than the alternative.

Solving paradoxes can be fun. But in this case it has another purpose. We are clearing the ground for a more fruitful moral debate about which acts (if any) are supererogatory. Does a non-moral reason to keep one's arms outweigh the moral reasons to save a life? Do we have any prerogative to eat the meat of non-human animals? To turn away refugees? Pollute the oceans? Or are we forbidden from doing these things? These questions are urgent and difficult. But we might see them in a new light, or at least with more clarity, if we can just get them out of the shadow of paradox.[24]

REFERENCES

Archer, A. (2016). Moral obligation, self-interest, and the transitivity problem. *Utilitas*, 28, 441-464.

  doi: 10.1017/S0953820816000091

-------(2017). Supererogation. *Philosophy Compass* 13. doi: 10.1111/phc3.12476

Arrow, K. (1951) [1963]. *Social Choice and Individual Values*. New York: Wiley. 2nd Edition 1963.

Bader, R. (2019). Agent-relative prerogatives and suboptimal beneficence. In M. Timmons (Ed.),

  Oxford Studies in Normative Ethics, Volume 9. Oxford: Oxford University Press (pp. 223-250).

Barnett, David (ms.). Graded ratifiability.

Bedke, M. (2011). Passing the deontic buck. In R. Shafer-Landau (Ed.), *Oxford Studies in Metaethics,*

  Volume 6 (pp. 128-152). Oxford: Oxford University Press.

Benn, C. (2017). Supererogatory spandrels. *Ethics & Politics*, 19, 269-290.

Bradley, B. (2016). Against satisficing consequentialism. *Utilitas*, 18, 97-108. doi:

  10.1017/S0953820806001877

Chang, R. (2002). The possibility of parity. *Ethics*, 112, 659-688. doi: 10.1086/339673

Dancy, J. (1993a). *Moral Reasons*. Oxford: Blackwell.

-------(1993b). *Beyond the Call of Duty: Supererogation, Obligation, and Offence* by Gregory Mellema.

  Philosophical Books, 34, 48-49.

Darwall, S. (2013). "But it would be wrong." In *Morality, Authority, and Law: Essays in Second-Personal*

  Ethics I. Oxford: Oxford University Press (pp. 52-71).

Dietrich, F. and List, C. (2017). What matters and how it matters: a choice-theoretic interpretation

  of moral theories. *Philosophical Review*, 126, 421-479. doi: 10.1215/00318108-4173412

Dorsey, D. (2013). The supererogatory, and how to accommodate it. *Utilitas*, 25, 355-382.

  doi: 10.1017/S095382081200060X

Dreier, J. (2004). Why ethical satisficing makes sense and rational satisficing doesn't. In M. Byron

(Ed.), *Satisficing and Maximizing*. Cambridge: Cambridge University Press (pp. 131-154).

Ferguson, B. (2016). The paradox of exploitation. *Erkenntnis*, 81, 951-972. doi: 10.1007/s10670-015-9776-4

Ferguson, B. and Köhler, S. (2019). Betterness of permissibility. *Philosophical Studies*. Print version forthcoming. doi: 10.1007/s11098-019-01319-8

Fishburn, P. (1970). *Utility Theory for Decision Making*. New York: John Wiley & Sons.

Gert, J. (2004). Value and parity. *Ethics*, 114, 492-510. doi: 10.1086/381697

-------(2007). Normative strength and the balance of reasons. *Philosophical Review*, 116, 533-562. doi: 10.1215/00318108-2007-013

Hare, C. (2010). Take the sugar. *Analysis* 70, 237-247. doi: 10.1093/analys/anp174

Heyd, D. (2016). Supererogation. In E. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Spring 2016 Edition), URL = <https://plato.stanford.edu/archives/spr2016/entries/supererogation/>.

Horgan, T. and Timmons, M. (2010). Untying a knot from the inside out: reflections on the 'paradox' of supererogation. *Social Philosophy and Policy*, 27, 29-63. doi: 10.1017/S026505250999015X

Horton, J. (2017). The all or nothing problem. *Journal of Philosophy*, 114, 94-104. doi: 10.5840/jphil201711427

Hurka, T. (1990). Two kinds of satisficing. *Philosophical Studies*, 59, 107-111.

Hurka, T. and Shubert, E. (2012). Permissions to do less than best: a moving band. In M. Timmons (Ed.), *Oxford Studies in Normative Ethics, Volume 2*. Oxford: Oxford University Press (pp. 1-27).

Kamm, F. (1985). Supererogation and obligation. *Journal of Philosophy*, 82, 118-138.

-------(1996). *Morality, Mortality, Volume II: Rights, Duties, and Status*. New York: Oxford University Press.

-------(2007). *Intricate Ethics: Rights, Responsibilities, and Permissible Harm*. Oxford: Oxford University Press.

Lazar, S. (2019). Accommodating options. *Pacific Philosophical Quarterly*, 100, 233-255. doi: 10.1111/papq.12252

Lazar, S. and Barry, C. (ms.). Acting beyond the call of duty: supererogation and optimization.

Little, M. and McNamara, C. (2017). For better or worse: commendatory reasons and latitude. In Mark Timmons (ed.), *Oxford Studies in Normative Ethics, Volume 7*. Oxford: Oxford University Press (pp. 138-160).

McMahan, J. (2018). Doing good and doing the best. In P. Woodruff (Ed.), *The Ethics of Giving: Philosophers' Perspectives on Philanthropy*. New York: Oxford University Press (pp. 78-102).

Muñoz, D. (ms.). From rights to prerogatives.

Parfit, D. (1982). Future generations: further problems. *Philosophy and Public Affairs*, 11, 113-172.

-------(2011). *On What Matters, Volume One*. Oxford: Oxford University Press.

Pettit, P. (1991). Decision theory and folk psychology. In M. Bacharach and S. Hurley (Eds.), *Essays in the Foundations of Decision Theory*. Basil Blackwell.

Portmore, D. (2011). *Commonsense Consequentialism: Wherein Morality Meets Rationality*. New York: Oxford University Press.

-------(2017). Transitivity, moral latitude, and supererogation. *Utilitas*, 29, 286-298. doi: 10.1017/S0953820816000364

-------(2019). *Opting for the Best: Oughts and Options*. Oxford: Oxford University Press.

Postow, B. (2005). Supererogation again. *Journal of Value Inquiry*, 39, 245-253. doi: 10.1007/s10790-006-1344-3

Pummer, T. (2016). Whether and where to give. *Philosophy and Public Affairs*, 44, 77-95. doi: 10.1111/papa.12065

-------(2019). All or nothing, but if not all, next best or nothing. *Journal of Philosophy*, 116, 278-291. doi: 10.5840/jphil2019116518

Rabinowicz, W. (2008). Value relations. *Theoria*, 74, 18-49. doi: 10.1111/j.1755-2567.2008.00008.x

Raz, J. (1975). Permissions and supererogation. *Philosophical Quarterly*, 12, 161-168.

Scheffler, S. (1982). *The Rejection of Consequentialism: A Philosophical Investigation of the Considerations Underlying Rival Moral Conceptions*. Oxford: Oxford University Press.

Sen, A. (1993). Internal consistency of choice. *Econometrica*, 61, 495-521.

-------(2017). *Collective Welfare and Social Choice: Expanded Edition*. Cambridge: Harvard University Press.

Sinclair, T. 2018. Are we conditionally obligated to be effective altruists? *Philosophy and Public Affairs*, 46, 36-59. doi: 10.1111/papa.12107

Slote, M. (1984). Morality and self-other asymmetry. *Journal of Philosophy*, 81, 179-192.

-------(1985). *Common-sense Morality and Consequentialism*. London: Routledge & Kegan-Paul.

-------(1991). Shelly Kagan's *The Limits of Morality. Philosophy and Phenomenological Research*, 5, 915-917.

Snedegar, J. (2015). Contrastivism about reasons and ought. *Philosophy Compass*, 10, 379-388. doi: 10.1111/phc3.12231

-------(2016). Reasons, oughts, and requirements. In R. Shafer-Landau (Ed.), *Oxford Studies in Metaethics, Volume 9*. Oxford: Oxford University Press (pp. 183-211).

Tadros, V. (2011). *The Ends of Harm: The Moral Foundations of Criminal Law*. Oxford: Oxford University Press.

Temkin, L. (2012). *Rethinking the Good: Moral Ideals and the Nature of Practical Reasoning*. Oxford: Oxford University Press.

Tungodden, B. and Vallentyne, P. (2005). On the possibility of Paretian egalitarianism. *Journal of Philosophy*, 102, 126-154. doi: 10.2307/3655706