Infinite Options, Intransitive Value, and Supererogation

Daniel Muñoz
Monash University

Supererogatory acts are those that lie "beyond the call of duty." There are two standard ways to define this idea more precisely. Although the definitions are often seen as equivalent, I argue that they can diverge when (i) options are infinite, or when (ii) there are cycles of better options; moreover, each definition is acceptable in only one case. I consider two ways out of this dilemma.

1. Introduction

Many good deeds lie "beyond the call of duty." Think of heroic sacrifices, thoughtful favors, and gifts of forgiveness. These are all wonderful things to do, and yet they seem optional rather than obligatory. Such acts are called *supererogatory*. They are apparently everywhere.

A century ago, supererogation was neglected by leading lights like G.E. Moore (1903). Since then, philosophers have brought it out of the shadows. We now have two clear and illuminating definitions of supererogation. The first says that supererogatory acts are those that are *permissible* and *incompatible with the moral minimum*.[1] To call an act "permissible" is to say that it may be done, that it would not be wrong. The idea of a "moral minimum" may be left undefined, but it roughly means "doing the least good that one can permissibly do." For example, keeping my spare kidney is part of the moral minimum that I could do right now; donating it is incompatible with keeping it; but it is also perfectly permissible to donate. By this first definition, the donation is therefore supererogatory. That is exactly the right result.

The second definition says that a supererogatory act is *permissible* and yet *better than a*

---

[1] "Despite disagreement in the details, there is wide agreement that acts of supererogation are both morally optional and morally better than the minimum that morality demands" (Archer 2018: 5). See Ferry 2013: 574; Benn 2017b: 271; McNamara 1996: 429, Portmore 2019: 11. In later work, McNamara (2011a: 155, 2011b: 215) calls this notion "acting beyond the call," reserving "supererogation" for acts that also merit praise. I will not say much about praiseworthiness.

permissible alternative.[2] Lo: giving the kidney is permissible and yet better than keeping it. Eventually

we should clarify what we mean by 'better'—nicer consequences, more praise due to the agent, more

reason in favor of the act?—but whatever the details, this definition gets the kidney case right, too.[3]

Our two definitions thus coincide in our case of self-sacrifice. Same for normal cases of

favors, forgiveness—indeed, any boring normal finite choice. But what about choices from infinite

options? What if our options A, B, C, … form a cycle of betterness, where A > B > C > … > A?

In such cases, I argue, not only do the definitions come apart: they are both unacceptable,

though in different ways. One can handle only infinite cases; the other can handle only cycles.

Assuming for now that we care about both cases, we face a dilemma. If we can't solve it,

supererogation may slink back into the shadows of indefinability.

The official goal, then, is to find a more rigorous definition of supererogation. But *en passant*

I also hope to convince you of a methodological point: that infinities and cycles deserve to be seen

as valuable test cases in deontological ethics, as they are seen in epistemology (e.g. Builes 2020 on

hyperintensional evidence for infinite events), decision theory (see e.g. Briggs 2019: Section 3.2.3,

Nover & Hajek 2004 on infinite expected utilities), and economics (see e.g. Hansson & Grüne-

Yanoff 2018: section 2.3; Sen 2017: Chapter 1*; and Ahmed 2017 on cyclic preference). Funny cases

can yield serious insights.

2. Infinite Options

Here is a familiar kind of infinite case. You are handed a *Deontic Dial* that can be set to any real

number between 0 and 1. If left at either extreme, disaster ensues. If set between 0 and 1, however,

---

[2] See e.g. Benn 2017a: 2400n.2, Benn and Bales forthcoming. See also Archer (2018: 4–5), who freely switches between the two definitions, as does Portmore (2017: 286n.1, 287, 291).
[3] My topic is "moral" supererogation (which is morally better), as opposed to "rational" supererogation (which is better for the agent); see Benn and Bales forthcoming. I will not be able to discuss some other definitions of supererogation; see e.g. Dorsey 2013, Chisholm 1963: 109.

better and better things will happen, the higher the number you pick.[4]

In deontic logic, this kind of case is known to make trouble for the definition of obligation. The tempting definition is that I am obligated to pick the *best* option (see McNamara 1995: 311). But here I have infinitely many options, which get better and better with no upper limit, violating what is known as the "limit assumption" (Lewis 1973, Fehige 1994: 44). The tempting definition is silent on such cases; it cannot even recommend avoiding the calamitous 0 and 1. This problem has inspired a lively debate (Danielsson 1968: Chapter 3.2, Åqvist 1987, McNamara 1995), with no clear victor yet. As Fehige (1994: 42) puts it: "When the best options are lacking, then so are flawless accounts of the lack."

But morality is not just a matter of obligations. Beyond the classic "threefold classification" of actions into obligatory, wrong, and optional, we have the concept of the supererogatory.[5] Supererogation splits the optional acts between good deeds, which surpass the call of duty, and decent deeds, which merely answer it. It is precisely this split that we are here to explore. To that end, here is a new puzzling case. What if there is no worst permissible option?

Imagine a *Tradeoff Dial* that can be set to any real number between 0 and 1—except 0. If you set it to x, you will be subjected to x days of quite serious pain, and you will also relieve Ingrid, an innocent stranger, of $2x$ days of even worse pain. The bigger your sacrifice, the greater Ingrid's benefit. Clearly, since you have a right not to incur pain, you are not obligated to ratchet up the dial. But it sure would be nice of you.

In the *Tradeoff Dial*, you have infinitely many permissible options—setting the dial to 0.5, to 1, to 0.000001, etc.—and I think each of them is supererogatory.[6] Let me explain why.

---

[4] The case is due to Blake Barley (cited in McNamara 1995: 310n.1). Fehige (1994: 45) gives a related case: John can speak the name of a natural number n and thereby relieve n people of starvation.
[5] On the "threefold classification," see Urmson 1958.
[6] To clarify: the key to *Tradeoff Dial* isn't that *all* of the options are supererogatory, though they are. What matters for our purposes is that all *permissible* options are supererogatory.

Imagine that you have tentatively chosen to turn the dial to 0.5, and you are considering whether to move to something higher, like 0.6. Either option is permissible, but it is clearly better, in every relevant sense of 'better', to turn the dial up. We are talking about a big benefit for Ingrid at a smaller cost to you. We would be right to praise you if you spared her pain as an expression of altruism. You would be doing something you have more moral reason to do. You would also be producing better consequences, impartially considered. Indeed, the more you turn up the dial, the better things get, in all of these ways. It is therefore supererogatory to set the dial to 0.6 instead of 0.5. But the same reasoning shows that *any* setting is supererogatory, since any such setting will be better, in every relevant sense, than any of the infinitely many (permissible) settings below.

That's the intuitive take on our infinite case, *Tradeoff Dial*. What do our definitions say? Start with the most common definition:

> *Better Than a Permissible Alternative (BPA)*
> An option A is supererogatory just if A is permissible and, for some permissible option B, A is better than B.

This fits the case perfectly. Each permissible option is supererogatory because any way of setting the dial is better than (permissibly) setting it lower. It's a good day for BPA.

The other definition says:

> *Incompatible with the Moral Minimum (IMM)*
> An option A is supererogatory just if A is permissible and doing A is incompatible with the moral minimum.[7]

This trickier definition says that *no* option in *Tradeoff Dial* is supererogatory. Let's see why.

The first thing to clarify about IMM is its notion of "incompatibility." Two options are incompatible if doing one entails not doing the other. So when IMM says that a permissible option is supererogatory if it is "incompatible with the minimum," the point is that doing the minimum

---

[7] McNamara (1996: 429) puts it in terms of propositions rather than options: $\mathbf{SU}p =_{df} \mathbf{PE}p \,\&\, \mathbf{MI}\neg p$. (Where '$\mathbf{SU}p$' means "it is supererogatory that p," '$\mathbf{PE}p$' means "it is permissible that p," and '$\mathbf{MI}\neg p$' means "doing the moral minimum entails not-p.")

entails not doing that option. This is very different from the idea that permissible options are supererogatory if they include things that aren't entailed by the minimum. If I must pay my $20 phone bill either with $5 notes or $10 notes, the moral minimum doesn't entail that I pay in tens. But paying in tens is not supererogatory given IMM, because it is still *compatible* with the moral minimum that I pay this way. By contrast, if the moral minimum is donating exactly $5, then donating $10 would be supererogatory; it is inconsistent with donating only $5.

Now the other tricky term: the "moral minimum." There are two ways to think of this; specifically or generally. We could see the minimum as a fully *specific option*—intuitively, the worst permissible one (see Ferry 2013: 574 on "the minimally permissible act"). In this case my specific options are to set the dial to this or that number. But there is no worst permissible number. Every number is permissible—the lower the worse, and they infinitely descend. This means there is no minimally good way to discharge your obligations.[8] Without a minimum, nothing can be inconsistent with the minimum. If so, IMM wrongly says that no option here is supererogatory.

You might worry that there is something fishy about a missing minimum—isn't there always a ground floor? If you think there must be, you could instead say that the minimum in our case is simply this: setting the dial above 0! This is not a specific option. It's a *general* option, which can be done in different ways. But this general option doesn't rule out any permissible options in our case, since the permissible options *just are* the various ways of setting the dial above 0. You can set it to 0.1, to 1—any positive real on the unit interval that your heart desires. Since these are all compatible with the general minimum, none will count as "beyond the call."

However we interpret IMM, it says that no option is supererogatory in our infinite case,

---

[8] In informal moments, McNamara glosses over this possibility: "if it is possible for me to discharge my obligations in a supererogatory way (in a better than minimal way), then it ought to be possible for me merely to discharge them in a minimal way" (2011b: 426). But he is keenly aware of the problems raised by infinitely descending options (2011a: 157n.9), and I will be invoking his solution in section 5 below.

which does not seem like the right answer.[9] Indeed, the opposite seems true. Setting the dial anywhere above 0 ought to count as supererogatory. BPA gets this right; IMM does not.[10]

Nor is there any clear way to fix IMM, if we take the "minimum" to be an essential part of the definition. For we aren't just dealing with the technical problem of *how to define the minimum* when the permissible options lack a lower bound to the permissible. The problem is that, even given a minimum, we can't use it to define the supererogatory. One big part of IMM is that there can't be supererogatory options without an incompatible minimal option. But in our infinite case, *Tradeoff Dial*, every permissible option is supererogatory. There is no room for any other permissible option, general or specific, to play the complementary role of the moral minimum. IMM fails.

Next we turn to a quite different case, where IMM shines—and BPA struggles.

3. Cycles

Suppose you must donate $20, but there is no particular charity that you must donate to. You have 26 options for where to donate: Charity A, Charity B, and so on, all the way to Charity Z. Whimsically enough, these options form a cycle: donating to A is better than donating to B, which is better than donating to C, etc.—but donating to Z is better than donating to A.

Now, before we ask which acts are supererogatory in this *Charity Case*, we have to ask: why grant that there could even *be* cycles of betterness? Admittedly, cycles are weird. I doubt that they

---

[9] The infinite case is also problematic for Hansson's (2013: 448) definition: p is supererogatory *iff p* implies q, q is obligatory, p is optional, and p is better than q and not-p. This definition does not work in *Tradeoff Dial*. Picking 0.5 is optional; picking some number is obligatory (let's suppose); but we can't say that picking 0.5 is better than picking some number that *isn't* 0.5. For there are many such numbers. Smaller ones are worse; bigger ones are better.

[10] I admit that there is something paradoxical about this; no matter how hard I try to choose worse permissible options, I cannot fail to supererogate. Suppose that the dial starts off at 0.5, and I turn it down to .00000001, in hopes of being more selfish. Even this is supererogatory, given BPA. Indeed, no matter how low I go, I still soar beyond the call. Worse still, in *Tradeoff Dial*, there is *nothing* I can do except supererogate. How then can we attribute my going "beyond" to my agency at all? I leave this question open. (I owe the ideas in this footnote to a very supererogatory referee.)

are genuinely possible, and so I won't ask you to assume that they are genuinely possible. But I think

we should at least entertain their possibility, in this case, for two reasons.

First, even if cycles are impossible, we might still be able to reason about what *would* be true

if some options *were* to form a cycle. We are often good at counterpossible reasoning, as shown by

its uses in moral philosophy—where the deontologist may ponder what follows from the utilitarian's

view, and vice versa—as well as its use in sciences like relative computability theory, where people

ask what would follow from the decidability of certain problems (which may in fact be necessarily

undecidable; see Jenny 2016). Even though I suspect that cycles are impossible, I find myself with

strong intuitions about cycling beyond the call.

Second, because there are arguments for the possibility of cycles, we should be willing to

explore the views of people who find those arguments convincing. One classic in this genre is the

"Spectrum Argument" (Temkin 1996, 2012: Chapter 7; see also Rachels 1998, 2001: 216).[11]

Here is a Spectrum Argument involving our case.[12] Suppose that Charity Z will do lot of

good for a tiny group—say, 260 units of goodness each for 10 people. Charity Y will help

considerably more people, but will give each of them a bit less—say, 250 units each for 20 people.

Intuitively, it is better to give to Charity Y; the increase in *total* benefit matters more than the drop in

the *average* benefit to each. Now suppose that each successive charity is like this—bigger total

benefit, smaller average benefit, intuitively better overall—all the way up to Charity A, which gives a

---

[11] I should clarify the relation between cycles and transitivity. To say that 'better than' ('>') is *transitive* is to say the following: for any A, B, and C, if A > B and B > C, then A > C. To say that there are *cycles* of betterness is to say that there exist some A, B, C, … such that A > B > C > … > A (Sen 2017: 62). How are cycles and transitivity linked? Here is the key. If 'no worse than' is transitive, so is 'better than', and if 'better than' is transitive, there can't be cycles of betterness—but these implications each go only one way (Sen 2017: 62–3, 66; see also Temkin 1987: 144–45). That means that we could in principle give up transitivity without embracing cycles. Indeed, some arguments against transitivity make no attempt to establish cycles (see e.g. Temkin 1987, and Kamm 1985, 1996: Chapter 11 (most of Kamm's cases are explicitly not meant to involve cycles)).
[12] My case is based on Derek Parfit's Repugnant Conclusion (1984: Chapter 17).

paltry 10 units of goodness each to 260 people. Charity A is better than B, which is better than C, which is better than D, and so on all the way down to Charity Z. But A does *not* seem better than Z; in fact, the reverse seems true. It is better to hugely benefit a small group than to barely benefit each among a giant mass. We have a cycle: A > B > … > Z > A.

That is the Spectrum Argument. It isn't irresistible, but I think it has real force. It would be interesting if its conclusion ruled out an otherwise attractive view of supererogation.

Now, which options are supererogatory in our cycling case? Here, donating to charity is a duty. But my intuition is that no specific way of donating is supererogatory. Even though each option is better than an alternative, we can no longer see this as a straightforward improvement, as we did in the infinite case. Why? Because switching to better and better charities can (and eventually *will*) make things worse. Suppose you are considering giving to Z, then you switch to Y, in hopes of going beyond the call. If you keep "upgrading" to better charities, you will eventually wind up with A, which is *worse* than what you started with! This just isn't a problem in the infinite case, where helping more is always better.

Here's the bottom line. In our A–Z cycling case, no permissible option is supererogatory. This is great news for IMM. Since there is no worst way to give, the moral minimum is either undefined, or it is defined as the general option of giving to some or other charity. Either way, the minimum doesn't rule out giving to any particular charity, so no particular gift is supererogatory. BPA, meanwhile, wrongly says that *every* permissible option is supererogatory. Each choice of charity is permissible, and each is better than a permissible option: A > B > … > Z > A.

4. Better Than BPA?

We have a dilemma. On the one hand, IMM tells us that supererogation is permissible yet inconsistent with the minimum; this view correctly predicts that no option is supererogatory when

all of our permissible options form a cycle. On the other hand, BPA says that supererogatory acts are permissible and better than a permissible alternative; this view correctly predicts that every option is supererogatory when all of our permissible options form an infinite ascending chain. But neither definition gets the right answer for both chains and cycles.

What to do? We have two options: amend BPA to deal with cycles, or amend IMM to deal with infinite descent. Let's start with BPA. Can we fix it?

There (misleadingly) appears to be an easy fix. Just add a restriction:

Acyclic BPA
An option A is supererogatory just if it is better than a permissible alternative B, and there are no options C, … such that A > B > C > ... > A.

On this view, it's not enough for the supererogatory act to be better than a permissible alternative; we also require that two acts not belong to a cycle. The rough idea: if you're going in circles, you aren't going beyond the call.

But let me clarify something. Acyclic BPA does *not* say that supererogatory options can't belong to *any* cycles. Recall the *Charity Case*, where you choose between Charities A–Z, which form a cycle: A > B > … > Z > A. Suppose we add a further option—giving to Charity Ω—that is permissible but worse than giving to any of A–Z; the hobnobbing honchos of Charity Ω will squander some of your cash on lavish fundraisers. In this *Ω Case*, it should clearly be supererogatory to give to one of the other charities, even though they cycle with each other.

Acyclic BPA can allow for this. What the principle says, to put it more carefully, is that acts can't be supererogatory in virtue of being better than options with which they cycle. If you start from a non-supererogatory permissible act (like giving to Charity Z, in our original case), and move to better and better options along the arc of a cycle (to Y, then X, etc.), you will never arrive at a supererogatory option. Moving along a cycle cannot *take* you beyond the call of duty. We can accept this and still say that giving to any charities from A–Z is supererogatory once the inferior Ω is

thrown into the mix.

That said, Acyclic BPA is vulnerable to another kind of counterexample: the *Sweetening Case* (cf. Chang 2002). Instead of adding Ω, suppose we add an option that is better than one of the originals while still cycling with the others. For instance, suppose we add A+, which is a slightly improved version of A (perhaps it involves volunteering a bit extra for Charity A). Since A+ is an improvement, we have A+ > A, and since the improvement is slight, we still have Z > A+. So there is a longer cycle of permissible options: A+ > A > B > … > Z > A+. According to Acyclic BPA, then, A+ is *not* supererogatory. That is bad news, since A+ seems to lie beyond the call.[13]

Why does A+ seem supererogatory? The natural answer, I think, is that A+ is better than A and compares at least as well as A to any alternative: if A is better than a certain option, so is A+; if A is no worse than a certain option, so is A+. Instead of myopically focusing on the pair itself, we need to ask how the options compare to the rest of the menu—the "bigger picture."

This suggests another definition, which I tentatively endorse:

Big Picture BPA
An option A is supererogatory just if (i) A is permissible; (ii) A better than a permissible alternative B; and (iii) for any C, B > C implies A > C, and ¬(C > B) implies ¬(C > A).

This rightly tells us that A+ soars beyond the call. It is better than the permissible A, and it fares at least as well as A does when compared to any alternative.

But what does Big Picture BPA say about the simple A–Z case? It depends. For even though we have said that A–Z form a cycle—which tells us how each option compares to its alphabetical neighbors—we haven't laid out any *other* pairwise comparisons. Is A better than C? Is M better than Z? The answers will determine what Big Picture BPA counts as supererogatory.

To illustrate, suppose '>' is *incomplete*, except for the cycle A > B … > Z > A. This means

---

[13] This verdict has a curious consequence. If A+ is supererogatory, then we have a supererogatory option that is *worse* than a *non-supererogatory* option. A+ is worse than Z. Awfully strange. But with cycles, would we expect anything less?

that, for any pair of options that aren't adjacent in the cycle, such as A and C, neither option is better than the other: ¬(A > C) and ¬(C > A). Given so much incompleteness, Big Picture BPA says that no option is supererogatory, which seems right. Take A, B, C, and D. Although B is better than C, B does *not* fare as well with respect to D: we have C > D and ¬(B > D). Moreover, B fares worse than C with respect to A: we have A > B and ¬(A > C). In light of this, it is hard to say that B purely improves on C. The same goes for the middle pair within any other quartet of adjacent options.

But suppose we tweak two pairwise judgments: we add that B > D and that A > C. Now B *does* seem supererogatory. For B is better than C; it is better than everything that C is better than (viz. D); and everything that is better than B (viz. A) is also better than C. In a way, we are thinking of B as "C+," and so just as Big Picture BPA called A+ supererogatory in the *Sweetening Case*, it here says the same of B. Another right result. What's more, Big Picture BPA gets the right answer in our other examples, too: *Tradeoff Dial*, the *Ω Case*—it's our best principle yet.

5. Beyond IMM?

Back to IMM—can it be salvaged?

As a reminder, IMM says that supererogatory acts are permissible and incompatible with the moral minimum. The problem with this, as we saw, is that there can be supererogation without minimality. In *Tradeoff Dial*, there is no minimally decent number at which to set the dial, and yet there are supererogatory ways to set it (indeed, *all* settings are supererogatory).

But what if we take out the idea of a "minimum" from IMM? Can we wring a suitable definition from the bare idea that supererogatory acts are incompatible with mediocrity? It seems we can. The key isn't that all supererogation is incompatible with a *particular* bad option, but that each way of supererogating is incompatible with some or other bad options. Consider:

Incompatible with Descending Levels (IDL)
An option A is supererogatory just if A is permissible and, for some permissible option B, doing A is incompatible with all permissible options worse than, or as good as, B.[14]

This definition gives up on the idea that, for a given choice, we can pick out a particular option (or set of options) as the minimum. There needn't be a minimum. The intuitive idea is, instead, that for any supererogatory option S, we can draw a cutoff along the ranking of options from worst to best, such that S is superior to every permissible option at or below the cutoff. (There needn't be *one* cutoff that every supererogatory option hovers over.)

This view does better with *Tradeoff Dial.* For any setting *n* of the dial, there is a cutoff below which every permissible setting is worse than *n*. Setting the dial to .5 is supererogatory because it is better than setting the dial to .4 or lower. Setting the dial to .4 is supererogatory because it is better than setting the dial to .3 or lower—you get the idea.

IDL succeeds in the infinite case. But can it handle cycles?

Take the simple *Charity Case* of A–Z. As IMM correctly predicts, no option is supererogatory in this situation. But what does IDL say? Surprisingly, it says the opposite; on a natural way of filling out the case, IDL dubs *every* charitable option supererogatory.[15]

Consider the option of giving to Charity A. By hypothesis, this is better than permissibly giving to Charity B; let's make B our cutoff. And let's suppose that giving to B is better than, or as good as, each of the remaining permissible alternatives (giving to any of C–Z). All of these (B–Z) are incompatible with giving to A. The gift to A is therefore supererogatory; it is better than giving to B,

---

[14] I owe the ideas in this paragraph and the next, including the name "IDL," to an extraordinary referee who suggested many other improvements (such as using the "deontic dial" instead of my original, baroque infinite case, along with the ideas in fn. 10); I am truly grateful. The origin of the principle I call IDL is in McNamara 2011a: 169.

[15] There is also the verbal problem that, if 'better than' is intransitive, we do not strictly speaking have a *ranking* on which to draw a cutoff. (Is anything "below" anything on a cycle?) This is why I have not used "cutoff" talk in the official statement of IDL, though I find it useful in informal explanation.

and incompatible with everything worse than or as good as giving to B. By similar reasoning we could show that every option in the cycle is supererogatory. The problem here, intuitively, is that IDL counts A as supererogatory simply because it beats B, which is no worse than C–Z, without considering whether Z beats A. IDL thus does *worse* than IMM in the case of cycles. It is in the same boat as BPA.

Can IDL be patched up? One idea, analogous to Acyclic BPA, is:

Acyclic IDL
An option A is supererogatory just if (i) A is permissible; (ii) for some permissible option B, doing A is incompatible with all permissible options worse than, or as good as B; and (iii) there are no options Z, … such that A > B > … > Z > A.

The point of this is to notch an exception into IDL. An option A is normally supererogatory when it is better than a permissible alternative B (and incompatible with anything as good as or worse than B); beating B isn't enough, however, if A and B form a cycle.[16]

But Acyclic IDL, like Acyclic BPA, struggles with the *Sweetening Case*. Although the sweetened A+ should be supererogatory, Acyclic IDL says that it isn't (supposing that A is as good as, or better than, each of C–Y). True, A+ is better than A, and A+ is incompatible with all permissible options worse than, or as good as, A (namely, B–Y). But A+ and A form a cycle: A+ > A > B > … > Z > A+. This means that the "no cycles" exception in Acyclic IDL kicks in, and A+ does not count as supererogatory.

The natural way to handle the *Sweetening Case*, in my view, is to advert to Big Picture BPA.[17]

---

[16] Like Acyclic BPA, Acyclic IDL is not as hostile to cycles as it might sound: it allows that giving to any of A–Z is supererogatory in the *Ω Case*, where Charity Ω is permissible and worse than giving to any of A–Z. (The principle's third clause is idle, because there is no cycle involving Ω.)

[17] A determined friend of IDL could try another tweak:

Acyclic IDL*
An option A is supererogatory just if (i) A is permissible; (ii) for some permissible option B, doing A is incompatible with all permissible options worse than, or as good as B; and (iii) if there is a Z, … such that A > B > … > Z > A, *then* Z > B.

And so, given the extravagant assumption of cycles, I conclude that Big Picture BPA is our best hope for defining supererogation. It covers lots of cases with a minimal dip in elegance.

6. Conclusion

Two standard definitions of supererogation—BPA (better than a permissible alternative) and IMM (incompatibility with the minimum)—lead to a dilemma when we go beyond acyclic finite options. My solution is to replace BPA with Big Picture BPA: an option A is supererogatory just if (i) A is permissible, and (ii) A is better than a permissible alternative B, where (iii) A compares at least as well as B to the rest of the menu. On this definition, every permissible option is supererogatory in the infinite case of *Tradeoff Dial*, and nothing is supererogatory in the cycling *Charity Case*.

But it's worth emphasizing that this tweak to BPA does more than accommodate intuitions. We are also capturing a theoretical insight: given cycles, betterness becomes a rather blunt notion. We need something sharper if we want to carve out a space for supererogation, to single out the sense in which it goes "beyond" the call. This is a step forward in the more general quest for a rational way to choose from cyclic options; "opt for the best!" is no longer enough.

I have also argued that Big Picture BPA can handle certain tricky cycles (Ω, A+). That is not to say that the principle is airtight. People may devise even trickier cases, which demand further tinkering. Still, I think we have already reached a significant conclusion. We have shown that BPA is

---

Which says: the "no cycles" exception doesn't kick in if, for every way of drawing the cutoff, the options at or below the cutoff *themselves* form a cycle—the sign of sweetening. For in the *Sweetening Case*, this is what happens with A–Z. A+ beats A, and there is a cycle from A to Z to A+, but A–Z *themselves* form a cycle, because Z > A. (It is not just that Z > A+ > A.)

Alas, an asterisk is not enough. Not only is Ayclic IDL* less elegant than Big Picture BPA; Big Picture BPA handles more cases. Suppose we change the *Sweetening Case* so that A fares better than A+ with respect to C: we have A > C and yet ¬(A+ > C). To my mind, this should disqualify "A+" as supererogatory, since it now fares worse than A against C. Big Picture BPA embodies this reasoning exactly. But Acyclic IDL* still counts A+ as supererogatory. In effect, Acyclic IDL* tells us to consider *part* of the bigger picture (viz., how A and A+ fare against Z) while ignoring the rest (e.g., how A and A+ fare against C). It seems to me the whole picture ought to matter.

more deeply right than IMM. The notion of a moral minimum is useful, but not as fundamental or flexible as the idea that supererogation compares favorably—in some sense—to other permissible ways of acting.

This conclusion, of course, depends on intuitions about exotic choices. I have assumed that our definition of supererogation must hold even when there is no moral minimum—either because our permissible options loop around, or because they infinitely descend into the mediocre. Can we really trust our judgments about such recherché scenarios?

I take the point. I don't assume a priori that we can trust our hunches about infinities and cycles, nor do I want to claim that my intuitions here are irresistible. But I am struck by how robust they are (as I hope you have been struck by your own reactions). In *Tradeoff Dial*, I strongly feel that every nonzero setting is supererogatory, and insofar as I can wrap my head around cycles, I am confident that giving to Charity A isn't supererogatory just because it beats giving to Charity B.

To be sure, many philosophers would prefer not to have cycles anywhere near their headspace; they believe that cycling is conceptually impossible.[18] These readers may resist my take on the *Charity Case* (as they resist the Temkin/Rachels Spectrum Argument). But I hope that even they will find value in the discussion above. After all, cycles aside, we showed that BPA succeeds in choices from infinite options, where IMM must be replaced by IDL. That is not quite as exciting as the dissolution of a new dilemma. But it is progress. At the very least, we have shown that "going beyond the call" is not the same as "going beyond the minimum," because when options are infinite, there may be no end to the worse and worse ways of discharging obligations.[19]

---

[18] For some important recent doubts about cycles, see Voorhoeve 2008 and Nebel 2018.

[19] This paper began as a joke. Specifically, a text from David Balcarras, who came up with a case that parodied my love of supererogation and David Builes's recent paper on infinite coin flips. My thanks to both Davids for getting me to think less finitely. For comments and encouragement, I thank David Builes, Al Hájek, Liz Jackson Withorn, Chris Tucker, Doug Portmore, and three referees. One referee in particular—mentioned in the footnotes—deserves special thanks for multiple rounds of brilliant comments. I do not know what minimally decent refereeing looks like, but not like that.

Works Cited

Ahmed, Arif (2017). "Exploiting Cyclic Preference," in *Mind* 126 (504): 975–1022.

Åqvist, Lennart (1987). *Introduction to Deontic Logic and the Theory of Normative Systems*. Naples: Bibliopolis.

Archer, Alfred (2018). "Supererogation," in *Philosophy Compass* 13: 1–9.

Benn, Claire (2017a). "Supererogation, Optionality, and Cost," in *Philosophical Studies* 175 (10): 2399–417.

-------(2017b). "Supererogatory Spandrels," in *Ethics & Politics* 19 (1): 269–90.

Benn, Claire and Bales, Adam (forthcoming). "The Rationally Supererogatory," in *Mind*. URL = <https://academic.oup.com/mind/advance-article-abstract/doi/10.1093/mind/fzz055/5580317>.

Briggs, R. A. (2019). "Normative Theories of Rational Choice: Expected Utility," in Edward N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy* (Fall 2019 edition), URL = <https://plato.stanford.edu/archives/fall2019/entries/rationality-normative-utility/>.

Builes, David (2020). "A Paradox of Evidential Equivalence," in *Mind* 129 (513): 113–27. doi: 10.1093/mind/fzz046

Chang, Ruth (2002). "The Possibility of Parity," in *Ethics* 112 (4): 659–88.

Chisholm, Roderick M. (1963). "Supererogation and Offense: A Conceptual Scheme for Ethics," in *Ratio* 5: 1–14.

Danielsson, Sven (1968). *Preference and Obligation*. Uppsala: Filosofisha Föreningen.

Dorsey, Dale (2013). "The Supererogatory, and How to Accommodate It," in *Utilitas* 25 (3): 355–82.

Fehige, Christoph (1994). "The Limit Assumption in Deontic (and Prohairetic) Logic," in Georg Meggle and Ulla Wessels (eds.), *Analyomen 1: Proceedings of the 1st Conference*. Berlin: Walter de Gruyter: 42–56.

Ferry, Michael (2013). "Does Morality Demand Our Very Best? Moral Prescriptions and the Line of Duty," in *Philosophical Studies*, 165 (2): 573–89.

Hansson, Sven Ove (2013). "Representing Supererogation," in *Deontic Logic Corner* 25 (2): 443–51.

Hansson, Sven Ove and Grüne-Yanoff, Till (2018). "Preferences", in Edward N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy* (Summer 2018 Edition). URL = <https://plato.stanford.edu/archives/sum2018/entries/preferences/>.

Jenny, Mattias (2016). "Counterpossibles in Science: The Case of Relative Computability Theory," in *Noûs* 52 (3): 530–60.

Kamm, Frances (1985). "Supererogation and Obligation," in *Journal of Philosophy* 82 (3): 118–38.

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

McNamara (1995). "How to Terminate Your Mom With Her Trust," in *Analysis* 55 (4): 310–13.

-------(1996). "Making Room for Going Beyond the Call," in *Mind* 105 (419): 415–50.

-------(2011a). "Praise, Blame, Obligation, and DWE: Toward a Framework for Classical Supererogation and Kin," in *Journal of Applied Logic* 9: 153–70.

-------(2011b). "Supererogation, Inside and Out: Toward an Adequate Scheme for Common-sense Morality," in Mark Timmons (ed.) *Oxford Studies in Normative Ethics, Vol. 1*.: 202–35.

Moore, G.E. (1903). *Principia Ethica*. Cambridge: Cambridge University Press. Revised edition, T. Baldwin (ed.), 1993.

Nebel, Jacob M. (2018). "The Good, the Bad, and the Transitivity of *Better Than*," in *Noûs* 52 (4): 874–99.

Parfit, Derek (1984). *Reasons and Persons*. Oxford: Oxford University Press.

Portmore, Douglas (2017). "Transitivity, Moral Latitude, and Supererogation," in *Utilitas* 29 (3):

286–98.

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

Rachels, Stuart (1998). "Counterexamples to the Transitivity of *Better Than*," in *Australasian Journal of*

Philosophy 76 (1): 71–83.

-------(2001). "A Set of Solutions to Parfit's Problems," in *Noûs* 35 (2): 214–38.

Sen, Amartya (2017). *Collective Choice and Social Welfare: Expanded Edition*. Cambridge: Harvard

University Press.

Temkin, Larry S. (1987). "Intransitivity and the Mere Addition Paradox," in *Philosophy and Public*

Affairs 16 (1987): 138–87.

-------(1996). "A Continuum Argument for Intransitivity," in *Philosophy and Public Affairs* 25: 175–210.

-------(2012). *Rethinking the Good: Moral Ideals and the Nature of Practical Reasoning*. Oxford: Oxford

University Press.

Urmson, J. O. (1958). "Saints and Heroes," in A. I. Melden (ed.), *Essays in Moral Philosophy*. Seattle:

University of Washington Press: 198–216.

Voorhoeve, Alex (2008). "Heuristics and Biases in a Purported Counter-example to the Acyclicity of

'Better Than'," in *Politics, Philosophy & Economics* 7 (3): 285–99.