

Hooker's rule-consequentialism, disasters, demandingness, and arbitrary distinctions

Fiona Woollard 

Department of Philosophy, University of Southampton, Southampton, UK

Correspondence

Fiona Woollard, Department of Philosophy, University of Southampton, Southampton SO17 1BJ, UK.

Email: f.woollard@soton.ac.uk

Abstract

According to Brad Hooker's rule-consequentialism, as well as ordinary moral prohibitions against lying, stealing, killing, and harming others, the optimific code will include an overriding "prevent disaster clause". This paper explores two issues related to the disaster clause. The first issue is whether the disaster clause is vague—and whether this is a problem for rule-consequentialism. I argue that on Hooker's rule-consequentialism, there will be cases where it is indeterminate whether a given outcome counts as a disaster such that it is permissible to infringe a given prohibition to avoid that outcome. I argue that it counts in favour of Hooker's rule-consequentialism that it makes this space for vagueness. The second issue is how to understand the disaster clause so that it does not make rule-consequentialism intolerably demanding—and more particularly whether avoiding over-demandingness requires the rule-consequentialist to place a counterintuitive limit on requirements to aid. I will argue that rule-consequentialism can avoid over-demandingness without placing a counterintuitive limit on requirements to aid.

KEYWORDS

Brad Hooker, demandingness, disaster clause, rule-consequentialism, vagueness

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2022 The Author. *Ratio* published by John Wiley & Sons Ltd.

According to Brad Hooker's rule-consequentialism:

Moral wrongness is determined by the code of rules whose internalization by the overwhelming majority of everyone everywhere in each new generation (not including generations after any new development that significantly reduces the costs of internalising more complex or demanding codes) has maximum expected value in terms of well-being with some priority for the worst off. (Hooker, 2005, pp. 268–269)

That is a bit of a mouthful, so we use the term “optimific code” to refer to that code of rules.

Hooker has argued that as well as ordinary moral prohibitions against lying, stealing, killing, and harming others, the optimific code will include an over-riding “prevent disaster clause”. This clause is needed to avoid cases where complying with these prohibitions will lead to disaster (Hooker, 2020, p. 450).¹

In this paper, I explore two issues related to the disaster clause.

The first issue is whether the disaster clause is vague—and whether this is a problem for rule-consequentialism. Hooker has argued that the disaster clause is vague but that this does not undermine rule-consequentialism. Rule-consequentialists can and should admit vagueness and imprecision in morality (Hooker, 2000, pp. 133–134). In a 2013 paper, Leonard Kahn presses a new variation of the vagueness objection based on an interesting thought experiment.² Kahn responds to his own objection, arguing that the sense of “disaster” relevant to rule-consequentialism is not vague. Kahn argues that for the rule-consequentialist, “a situation counts as a disaster if and only if the net value (benefits minus costs) of the consequences of allowing everyone to violate one or more of the rules in the ideal code, given that there is no other way to do so, is greater than the net value of not allowing everyone to do so” (Kahn, 2013, p. 234).

I will argue that Kahn is correct that the appropriate way for rule-consequentialists to understand “disaster” in this context is by reference to what the optimific rules would say about infringing the standard prohibitions. However, both Kahn's formulation of rule-consequentialism and Kahn's understanding of the disaster clause differ in important ways from Hooker's view. I will argue that the persuasiveness of Kahn's argument is a result of the wording of his account of disaster which in turn results from his formulation of rule-consequentialism. An account of disaster drawn from Hooker's formulation of rule-consequentialism does not give us reason to think that disaster is not vague. I will then argue that in fact we have reason to think that the Hooker's optimific rules will *not* provide a precise boundary number in the thought experiment presented by Kahn—and that this counts in favour of Hooker's version of rule-consequentialism.

The second issue is how to understand the disaster clause so that it does not make rule-consequentialism intolerably demanding—and more particularly whether avoiding over-demandingness requires the rule-consequentialist to place a counterintuitive limit on requirements to aid.

Hooker argues in *Ideal Code, Real World* that in order to prevent the disaster clause from becoming intolerably demanding, we require “some kind of line” that limits your duty to help strangers. Hooker argues that, even though it may seem counter-intuitively mean, such a line must imply that an agent, who has done what is required of them, is justified in refusing to make further sacrifices (Hooker, 2000, p. 168).³ In more recent work, he repeats this argument (Hooker, 2020, p. 451).

¹See also Hooker (2000, pp. 133–136; 165–169).

²Kahn understands vagueness in terms of indeterminacy: “the sense of a predicate is vague if and only if there are borderline cases with respect to whether or not this sense is truly predicated of its subject. So ‘disaster’ is vague if and only if there are some cases in which it is indeterminate whether or not, say, X is a disaster.” (Kahn, 2013, p. 230). I follow Kahn and use this understanding of vagueness throughout the paper. I thank Luke Elson for pressing me to clarify this.

³Hooker does leave support a requirement for extreme self-sacrifice in truly exceptional cases that is, “where such self-sacrifice is the only way to secure some good on the order of the saving of a significant proportion of humanity” (Hooker, 2000, p. 169).

I press Robert Noggle (2009)'s objection that Hooker, and others who try to avoid over-demandingness by putting an aggregate limit on beneficence, are committed to highly counterintuitive implications. I argue that to avoid other counterintuitive implications, we need to distinguish between different types of cases where we could save someone's life. We must recognise that "disaster" does not have a unified meaning across aid cases (just like it does not have a unified meaning across constraint-breaking cases). Once we have recognised such different categories, we can avoid over-demandingness without counterintuitive limits to requirements to aid.

There is a common thread connecting these two issues. I reject Kahn's no-vagueness argument on the basis that it makes an arbitrary distinction. The key challenge for any solution to the demandingness objection along the lines I suggest is to distinguish between the different types of cases where we could save lives without appeal to arbitrary distinctions. I finish by briefly explaining why I think this challenge can be met, despite the concerns I raised about arbitrary distinctions in the earlier sections.

1 | THE VAGUENESS OBJECTION: TORTURED LONDONERS

Kahn's version of the vagueness objection is a sorites style argument. We start with a thought experiment in which a demon will torture 8,000,000 Londoners for a year unless John lies to Harriet. Rule-consequentialism can appeal to the disaster clause to match the intuitive response that John should tell this lie. We then decrease the number of potential torture victims by 1. Intuitively, John should still lie to Harriet to save 7,999,999 people from torture. Rule-consequentialism can still explain this using the disaster clause. We continue decreasing the number of potential torture victims. Kahn argues:

At some point, the number of Londoners who would be tortured if John does not lie to Harriet will reach a point that is too low to qualify as a disaster. Under those conditions, let us say, it qualifies as only a "near disaster." But Richard Arneson asks, "If following the ideal code of rules even when doing so leads to disaster is irrational and morally wrong, why shouldn't we agree that following the sophisticated ideal code of rules even when doing so leads to near disaster is also irrational and morally wrong?" (2005). It seems to me that rule-consequentialists need either to provide a compelling answer to Arneson's question or to show that it is somehow misconceived. (Kahn, 2013, p. 228)

Kahn himself responds to this objection by arguing that the sense of "disaster" relevant to rule-consequentialism is not vague:

In fact, the distinction that [rule-consequentialism] makes between disasters and non-disasters is straightforward. According to [rule-consequentialism], whether or not a particular situation counts as a disaster or not is determined by whether or not everyone would be justified in violating one or more elements of [the ideal code] in order to prevent it, if there were no other way to do so. More to the point, a situation counts as a disaster if and only if the net value (benefits minus costs) of the consequences of allowing everyone to violate one or more of the rules in the ideal code, given that there is no other way to do so, is greater [than] the net value of not allowing everyone to do so. There is nothing inherently vague about this sense of "disaster" (Kahn, 2013, p. 234).

Kahn seems to be claiming that rule-consequentialism will give a precise boundary number for the torture-sorites puzzle described above. A number is a boundary number for the sorites puzzle if and only if it marks a sharp border between the cases where it is permissible to break the rule and the cases where it is impermissible to break the rule. More formally, n , is a boundary number if and only if it is permissible to break the rule to save n people from being tortured, but impermissible to break the rule to save $n-1$ people from being tortured.

I will argue later that Kahn is correct that the appropriate way for rule-consequentialists to understand “disaster” in this context is by reference to what the optimific rules would say about infringing the standard prohibitions. However, both Kahn’s formulation of rule-consequentialism and Kahn’s understanding of the disaster clause differ in important ways from Hooker’s view.

The first key difference between Kahn’s understanding of the disaster clause and Hooker’s understanding of the disaster clause is that Kahn seems to assume that there will be a single standard for whether a particular situation counts as a disaster. Hooker is clear that on his view what counts as a disaster is variable and will be different depending on which particular prohibition is at stake:

[What counts as a disaster is] variable (for example, what counts as a disaster the prevention of which would justify breaking a promise is far less than a disaster the prevention of which would justify imposing serious physical harm on someone). (Hooker, 2005, 272)⁴

Here, I think Hooker is clearly correct. If we require universal standards for something to count as a disaster, then the disaster clause will end up being utterly implausible. We will have to say either that a single death never counts as a disaster—implying that it is impermissible to break a promise to save a life—or that a single death always counts as a disaster—implying that it is permissible to cut off one person’s legs to save another’s life. It seems highly unlikely that the internalisation of such a code would produce the best consequences. Moreover, this would defeat the point of introducing the disaster clause in the first place—which was to prevent rule-consequentialism from having deeply counter-intuitive implications.

So, we should not expect a single standard for when a situation counts as a disaster that applies no matter what rule is to be violated. Instead, we should say that a situation counts as a disaster *relative to prohibition, p, and context, c*, if and only if the optimific rules would permit infringing prohibition *p* in that context in order to avoid that situation.

Recognising that what counts as a disaster is relative to a prohibition reveals a problem with the torture-sorites-puzzle as stated. Kahn argues: “At some point, the number of Londoners who would be tortured if John does not lie to Harriet will reach a point that is too low to qualify as a disaster” (Kahn, 2013, p. 234). But Hooker is clear that the death of one person counts as a disaster relative to at least some prohibitions such as the prohibition against a soldier deserting their post (Hooker, 2005, p. 272). I am not quite sure what kind of torture Kahn has in mind, but at least some forms of torture may be as bad as, or even worse than, death. And remember each of these poor people are going to be tortured for a year. It seems as if Kahn’s assertion could only result from an assumption that there must be a single standard for what counts as a disaster. Of course, Kahn might have extreme views about the strength of the prohibition about lying or there might be circumstances which make this lie especially heinous, but it seems as if in either of these cases Kahn should at least have included an explanation and noted that his claim would be controversial.⁵ It seems to me that, barring some special circumstances involving the lie, both Hooker’s view and the common-sense view will be that one single Londoner being tortured for a year is enough of a disaster that John should lie to Harriet to prevent this.

This does not dissolve the puzzle. Kahn can simply modify the example. He could say that the Londoners will be tortured unless John cuts off Harriet’s finger. But it does, I think, highlight the dangers of not being clear that “disaster” is variable when interpreting the disaster clause.

Second, and most importantly, Kahn’s account of when something counts as a disaster is based on a formulation of rule-consequentialism that makes no mention of internalisation costs. Kahn uses the following formulation of rule-consequentialism:

⁴See also Hooker (2000, p. 135).

⁵I thank Zsuzsanna Chappell, David Killoren, Nicholas Smyth, Daniel Star, David Sobel and David McNaughton for helpful discussion of this point.

Rule Consequentialism (RC): It is right for agent A to do action F in circumstances C if and only if A's doing F in C is in accordance with the ideal code.

The Ideal Code (IC): The set of rules which, if accepted by all, would have consequences which are better than any alternative set of rules (Kahn, 2013, p. 221).

Kahn explicitly recognises that there are various possible ways of filling in the details of rule-consequentialism, but states that he will not do so in this paper given that he is specifically interested in the disaster objections. He states that the statement he has given of rule-consequentialism "is sufficient for this purpose" (Kahn, 2013, p. 222). However, if I am right, the persuasiveness of Kahn's argument is a result of the wording of his account of disaster which in turn results from his formulation of rule-consequentialism. A Hooker-style account of when something counts as a disaster would not lead us to think that there should be a precise answer to whether something counts as a disaster for the rule-consequentialist.

Internalisation costs play an important role in Hooker's rule-consequentialism. For this reason, I would expect Hooker to include reference to internalisation in his account of whether a situation counts as a disaster (relative to a given prohibition, *p*, in context, *c*). I would expect Hooker to say something like:

A situation counts as a disaster, relative to a given prohibition, *p*, if and only if infringing prohibition, *p* in context, *c*, would be required by the code of rules whose internalization by the overwhelming majority of everyone everywhere in each new generation (not including generations after any new development that significantly reduces the costs of internalising more complex or demanding codes) has maximum expected value in terms of well-being with some priority for the worst off.

The crucial difference is that Kahn's formulation makes it sound as though whether something counts as a disaster is likely to be determinate. Remember, Kahn says: "a situation will count as a disaster if and only if the net value (benefits minus costs) of the consequences of allowing everyone to violate one or more of the rules in the ideal code, given that there is no other way to do so, is greater the net value of not allowing everyone to do so" (Kahn, 2013, p. 234). This makes it sound like there will be two determinate numerical values: benefits minus costs of acceptance of rule A and benefits minus costs of acceptance of rule B. The question of whether a situation counts as a disaster becomes the question of whether one of these determinate numerical values is bigger than the other. That seems like it should have a determinate answer.⁶

I do not think that the alternative Hooker-style formulation makes it seem that there should be a determinate answer to whether a situation counts as a disaster (relative to a given prohibition, *p*, in context, *c*). The Hooker-style formulation highlights that the question is what would be said about the situation by the optimific rules—the rules whose internalisation has maximum expected value. There seems to be no reason in advance to assume that the optimific rules will give a determinate answer about whether infringing prohibition, *p*, in context, *c*, would be required to avoid that situation. In particular, it seems unlikely that the optimific rules will provide us with a boundary number in the torture-sorites puzzle.

Why is it unlikely that the optimific code will provide us with a precise boundary number in the torture-sorites puzzle? One issue is the internalisation costs. It would be cognitively very difficult to grasp a rule which gave precise boundary numbers for each different rule, sacrifice and context. More than this, it seems likely to be psychologically difficult. We would need to come to see a massive moral difference between cases

⁶It might be that Kahn's approach would not lead to a determinate answer because it would not yield determinate numerical values. There may be indeterminacy in calculating the net value of accepting each rule due to indeterminacy in for example, which animals count morally and how to weigh different aspects of well-being. I thank Luke Elson for pointing this out. My argument is compatible with, but does not depend upon, this objection to Kahn's argument.

involving n people and cases involving $n-1$ people. But there seems to be barely any non-moral difference between the cases.

Kahn argues that rule-consequentialist distinction between disasters and non-disasters is not morally arbitrary. He argues that a distinction is morally arbitrary if and only if there is not sufficient moral reason to make that distinction. According to Kahn, the rule-consequentialist does, according to their own theory, have sufficient moral reason to make this distinction: they do it based on what would be required by the optimal moral code (Kahn, 2013, p. 234). Kahn might offer an analogous response to my worry about psychological difficulty: there is a morally significant difference between cases involving n people and cases involving $n-1$ people—the difference is that one is specified as a potential disaster by the optimific moral code and the other is not.

I do not think this response works. Rule-consequentialism does hold that appeal to the code is the ultimate basis for right and wrong. But surely there must still be room for a charge that some rules appear arbitrary and for that reason codes which include these rules will be difficult to internalise.

Perhaps one way to think about this is that internalising the rules of a code often involves coming to see certain non-moral features as morally significant and to develop ways of balancing those morally significant features of a situation when they conflict.

I think there are features which are not the kind of things that can be treated as morally significant without further explanation. To draw on a well-used example, it seems like it would be difficult to internalise a moral code that treated the fact that pain would occur on a Tuesday as changing the moral significance of that pain.⁷ It would be hard to bring people to see this feature was morally relevant—unless you could point to something that made Tuesdays different from other days. The claim that the optimific code would treat Tuesday pain as different does not seem to be the right kind of explanation—because it immediately raises the question of why the optimific code would have this odd feature. I think it would be hard to internalise a code which treated such features as morally relevant without the right kind of further explanation.

I am slightly unsure whether this line of argument is legitimate for a rule-consequentialist. Hooker does often appeal to what humans naturally find easy (or difficult) to accept in his discussion of internalisation costs: for example, he argues that our natural attachments to friends and family mean that rules requiring impartial concern would have very high internalisation costs (Hooker, 2000, p. 137). My argument is slightly different, for it appeals to an idea that we will naturally want our moral code to make a certain kind of sense to us, and that there will be some features of the moral code which naturally seem to stand in need of explanation. Is appeal to this kind of internalisation cost legitimate?

One concern with appeal to this kind of internalisation costs is that it may seem that if internalisation costs are to play the required role within rule-consequentialism, the calculation of internalisation costs must not appeal to moral judgments. According to rule-consequentialism, moral judgements are the output of the optimal code. This surely means that the inputs which determine which code is optimal cannot include moral judgments.⁸ Perhaps this concern can be mitigated by rephrasing the requirement on putative moral codes as a requirement for coherence: a code which treated as morally significant whether pain occurred on a Tuesday, without a further explanation for why Tuesdays were morally different from other days, would lack coherence.

There is a danger here. It seems as if any code will give some features basic moral significance: it will treat these features as morally significant without giving any further explanation. There seem to be some features which are suitable to be treated as of basic moral significance. The fact that an act would cause pain to a sentient creature seems to be one such fact. There are two possible ways to respond to this. I could say that, other things being equal, the coherence of a code is a function of how many features it gives basic moral significance: more features

⁷This example is inspired by Derek Parfit's Future Tuesday Indifference case (Parfit, 1984, p. 124).

⁸I thank Richard Arneson for pressing me on this.

with basic moral significance equals less coherence (other things being equal). Alternatively, I could argue that the optimal code can give basic moral significance to facts about pain without loss of coherence due to the role that such facts play in the code as a whole.

If the above line of reasoning about internalisation costs is legitimate, then I think it will rule out a precise answer in the sorites puzzle. The issue is not just that the difference between n and $n-1$ seems to be the kind of feature that should not be treated as morally relevant without further explanation. The cases seem to be so similar that there does not seem to be enough of a difference for us to get a grip on, either when it comes to seeing features that we can see as morally significant or when it comes to balancing the morally significant features.

Of course, you might argue that there is a clear, morally relevant difference between the n case and the $n-1$ case: one more person is tortured in the n case. That an additional person will be tortured is surely the type of fact that can be treated as morally relevant without further justification. But there is just the same difference between the $n-1$ case and the $n-2$ case. Why does the additional victim not lead to the same moral difference between the $n-1$ and the $n-2$ case?

It might seem as if the world is full of examples of arbitrary distinctions that have moral significance. The speed limit on the M3 motorway is 70 mph. Does not that create an arbitrary moral distinction between driving at 70 and 71 mph? I am not completely sure that we do draw a sharp moral distinction between 70 and 71 mph. But insofar as we do so, it is because having a clear speed limit is useful: it makes it easier for people to avoid driving dangerously; it makes sanctions fairer etc. It does not seem as if having a precise boundary number in the torture-sorites case is useful in the same way. After all, there is not a large number of people regularly facing such cases who need a clear rule to prescribe their natural temptation to lie too easily or who must be treated consistently to avoid unfairness.

It seems as if the best way to respond to the torture-sorites puzzle is to recognise the existence of genuinely hard cases. There will be cases where the consequences are clearly not bad enough to outweigh the prohibition. As the number of potential victims increases, the pressure on the prohibition increases. It will become increasingly difficult to make a judgement. We may reach a point where it becomes less difficult, or even when the number of potential victims becomes clearly enough to outweigh the prohibition.

Hooker's ideal code is more than just a set of rules which give rise to first order judgments about what is forbidden and what is permitted. It also includes dispositions to blame, to feel guilty, to punish (Hooker, 2000, p. 76). This aspect of Hooker's rule-consequentialism provides the resources to recognise hard cases. I see hard cases as cases where the first order classifications are genuinely indeterminate: infringing the prohibition is not determinately permissible in the hard case, nor is it determinately forbidden. If a moral code consisted only of first order judgments, then this indeterminate status may seem impossible—for it might seem as if any action that is not determinately forbidden is *de facto* permissible.⁹ But we can avoid this implication by taking second-order responses into account. Hard cases may be associated with a distinctive set of complex second order responses. For example, in hard cases blame may be inappropriate, but a certain kind of moral regret may be required.

Hooker argues: "...common moral consciousness seems similarly imprecise about how bad the consequences of keeping rules have to be to warrant breaking them. Indeed, perhaps we would distrust any moral theory that was precise about just how bad consequences had to be in order to justify normally abhorrent acts necessary to prevent those consequences" (Hooker 2000, p. 136). I think we would, rightly, distrust a moral code which drew a sharp line somewhere in between one potential torture victim and 8 million torture victims. For this very reason, I think that the costs of internalising such a code would be very high and thus that the optimific code is unlikely to draw such a line on Hooker's version of rule-consequentialism. Like Hooker, I see this lack of precision as a virtue of his version of rule-consequentialism. This version of consequentialism reflects the messy moral reality.

⁹I thank Robert Frazier and Penelope Mackie for pressing me on this.

Although I have criticised Kahn's argument that the sense of "disaster" relevant to rule-consequentialism is not vague, I agree with him on two key points.

First, I agree with Kahn that the appropriate way for rule-consequentialists to understand "disaster" in this context is by reference to what the optimific code would say about infringing the standard prohibitions. I say that appeal to what the optimal code would say is the appropriate way to understand "disaster" *in this context* because there may be different uses of the term "disaster" which are relevant to rule-consequentialism. For example, suppose that people could not be brought to care about the welfare of those who will live in the far future. This might mean that there were very high internalisation costs for a code which treated the welfare of far future people as morally important. With this in mind, we might ask whether the rule-consequentialist should include an "escape clause": including either requirements or permissions to act that were not part of the optimal code *where these are necessary to prevent disaster*.¹⁰ This is a very interesting question. However, this "escape clause" is considered in a very different context from that of Hooker's original disaster clause. The "escape clause" is supposed to kick in when the optimal code gives us an intuitively wrong answer: we thus need to understand "disaster" in the context of the escape clause in a way that does not appeal to the optimal code. In contrast, Hooker's original disaster clause is part of the optimal code: Hooker is arguing that the optimal code itself matches our intuitions that there are times in which the standard rules should be broken. In this context, to understand whether a given potential outcome counts as a disaster, the key thing to consider is whether the optimal code would treat it as such. Moreover, for the optimal code to treat a potential outcome as a disaster relative to a given prohibition just is for the optimal code to permit or require infringement of that prohibition to avoid that potential outcome.

Second, I agree with Kahn, that the vagueness objection does not undermine rule-consequentialism. Indeed, I agree with Kahn that the rule-consequentialist is in no more difficult situation than any other moral theory when it comes to the vagueness of "disaster". If I am right, a precise understanding of "disaster", and in particular a precise boundary number in the torture-sorites case would be difficult for humans to internalise. It would be difficult to internalise because it would require us to treat a distinction as morally relevant without a satisfying explanation. And, if I am right, for this reason, the optimific code is unlikely to draw such a line on Hooker's version of rule-consequentialism. Maybe I am wrong. Maybe there are good reasons to have a precise boundary number in the torture-sorites case. But if there are such good reasons, then there is *no* good reason to think that the rule-consequentialist cannot make use of them.¹¹ It seems as if rule-consequentialism is likely to provide a non-arbitrary firm boundary if and only if there should be a non-arbitrary firm boundary.

2 | DEMANDINGNESS

In *Ideal Code, Real World* and elsewhere, Hooker raised a concern that the disaster clause may make rule-consequentialism overly demanding.

For the starvation of innocent people counts as a disaster. And often you can save some of these lives by contributing to the best charities. But there is an overwhelming number of lives to be saved. Thus an unlimited requirement to prevent disaster seems to entail that you should keep

¹⁰I thank Melinda Roberts for pressing me on this.

¹¹Kahn considers an argument that the rule-consequentialist is not at disadvantage relative to any other plausible theory when it comes to the disaster clause and vagueness because: (P1) Any plausible ethical theory requires a disaster clause or relevantly similar alternative and (P2) Disaster is vague—as are all similar alternatives. He rejects this response because he rejects (P2) on the grounds that the sense of "disaster" relevant to the rule-consequentialist is not vague (Kahn, 2013, p. 230). I would suggest that the sense of disaster relevant to rule-consequentialism is not vague *if and only if* the sense of disaster relevant to the disaster clause is not vague.

making contributions—at least to the point where you yourself are impoverished. (Hooker, 2000, p. 165)

Hooker resolves this problem by arguing that the optimific rules will not include an unlimited requirement about preventing disasters. Getting people to internalise such a rule would be extremely costly (Hooker, 2000, p. 166). Instead, he recommends:

Over time, agents should help those in greater need, especially the worst off, even if the personal sacrifices involved in helping them add up to a significant cost to the agents. The cost to the agents is to be assessed aggregately not iteratively. (Hooker, 2000, p. 166)

There are a lot of interesting things to be said about the demandingness objection and Hooker's response to it. Unfortunately, there is not space here for a discussion of all these aspects. The issue that I want to focus on is whether the optimific code will include a limit on sacrifice which means that agents are permitted to refuse to make any further sacrifices on the grounds that they have made a big enough overall sacrifice already.

Hooker admits that any such limit is subject to the objection that it can seem “counterintuitively mean”: “How could it ever be morally permissible to ‘shut the gates of mercy on mankind?’” (Hooker, 2000, p. 168)¹² Nonetheless in both *Ideal Code, Real World* and, again, twenty years later, he argues that this is the only way to avoid over-demandingness.

I cannot see how there can be a limit to the demands of self-sacrifice that morality can reasonably require without there being some point where refusing to make further sacrifices is justified. (Hooker, 2020, p. 451)

Robert Noggle argues that Hooker, and others who try to avoid over-demandingness by putting a cap on beneficence, are committed to highly counterintuitive implications:

Suppose that Alice has already made beneficent sacrifices that equal or exceed [the limit]. *Ex hypothesi*, Alice is not required to incur any further costs in the name of beneficence. Suppose that Alice now encounters a child drowning in a shallow pond whom she could save at the cost of muddy clothes. Since Alice has already reached the cap, incurring this cost to save the drowning child is purely optional. A theory with such a cap would permit Alice to ignore the child and still congratulate her for being beneficent. (Noggle, 2009, p. 3)

As Noggle argues: “given the vital role that intuition plays in justifying Hooker's theory, it is worth asking whether an alternative construction of the moderate position might avoid these counterintuitive implications.” (Noggle, 2009, p. 4).

Any view that does not distinguish between different types of case where we can save lives is going to end up with extremely counterintuitive results. This is because intuitively we should be willing to make really very significant sacrifices in some cases. Consider the famous Bugatti case from Peter Unger.

Bob's Bugatti: Bob is a successful 70 year old engineer about to retire. With all but a \$3000 emergency fund, he has invested all his money in a mint-condition Bugatti automobile which he intends to sell soon. The Bugatti cannot be insured, thus if it is damaged Bob will lose his life-savings. Instead of a comfortable retirement Bob will have a hard time just making ends meet for the remaining 15

¹²Here Hooker borrows words from Thomas Grey's “Elegy Written in a Country Churchyard.”

to 20 years he can expect to live. Bob has carefully parked his Bugatti near a disused side-track of a trolley system. A runaway trolley is rushing down the main track to where a small child is trapped. Bob can press a switch, diverting the trolley to the side-track, but if he does so his precious Bugatti will be destroyed. (Unger, 1996, pp. 135–136)

Intuitively Bob is required to sacrifice his life savings to save the child in the Bugatti case. Yet a view that we are required to give our life savings to famine relief charities is certainly counter-intuitively demanding.

So, we need to distinguish between different types of cases where we could save someone's life. We must recognise that "disaster" does not have a unified meaning across aid cases (just like it does not have a unified meaning across constraint breaking cases). What counts as a disaster should also vary depending on (a) what kind of sacrifice is required; (b) the relationship between the agent and the potential victim and (c) the relationship between the sacrifice and the disaster prevention.

Recognising that whether an agent must aid and how much they must sacrifice depends upon the details of the case is needed to recognise the sacrifice required in the Bugatti case without endorsing a very demanding account of duties to aid. Once we have recognised such different categories, we can avoid over-demandingness without "shutting the gates of mercy". A requirement for Alice to save the child (however much money she has given to charity) does not entail a requirement to give unlimited amounts to charity. The Pond case is in a different category than the charity case.

Of course, taking this approach requires us to give an account of what makes different cases fall into different categories. As Unger puts it, we need to find some morally significant differences between the Pond case and the charity case (Unger, 1996, p. 9). On Hooker's view, this would require finding differences which the optimific code would treat as morally significant. Such an argument would need to bear in mind the points I made above about the difficulty of internalising a code which treats apparently arbitrary features as morally significant.

I have argued elsewhere that agents have both a duty to make regular contributions in response to ongoing need and a separate duty to make substantial sacrifices when personally involved in a crisis (Woollard, 2015, part III). Whether the agent is personally involved depends on various factors including proximity, whether they have had a personal encounter with the victim, and whether they are the unique potential saviour. These features are not morally significant in their own right but are naturally suited to being morally significant *given the need to pick out a limited set of cases where we should make substantial sacrifices*.

My argument for the need to pick out a limited set of cases where we are required to make substantial sacrifices is that it is required to balance the need for aid and the need to ensure that a person's body and other resources are mostly at her own use. The latter is in turn needed to ensure that the agent's body belongs to her. I have argued elsewhere that Hooker's rule-consequentialism can defend the need to respect that a person's body belongs to her (Woollard, 2015, chapter 9). However, Hooker could, of course, simply give his standard argument that the optimific code cannot be too demanding due to internalisation costs (Hooker, 2000, p. 166).

This approach has avoidance of over-demandingness built-in: the criteria are chosen specifically to pick out a set of cases which the agent can expect to face rarely if at all. In addition, I also hold that when an agent is faced with repeated unusual cases *then* they may argue that they have sacrificed enough. This is a more complex—and less counterintuitive—view than holding that enough aggregate sacrifice over a lifetime of any type can justify turning one's back in a Pond case.

I argue that the features picked out are naturally suited to being morally significant given the need to pick out a limited set of cases where we should make substantial sacrifices. It is this that allows me to avoid the concerns about arbitrary distinction raised earlier about the torture-sorites puzzle. The features that make an agent count as personally involved in a crisis are the kind of features that make the victim and their need loom large from the agent's point of view. Treating these cases as special works with, rather than against, our natural tendencies to see certain features as morally relevant.

On my account there are other features that affect how much the agent is required to sacrifice. For example, I argue that it may matter whether the costs to the agent are direct costs rather than opportunity costs and how the costs to the agent come about (Woollard, 2015, pp. 157–158). It is this latter feature that I appeal to in order to explain the extremely high sacrifice required in the Bugatti case:

... the Bugatti itself contributes to the overall problem. If the Bugatti were not present, Bob would have no reason to object to turning the trolley. If we accept the potential destruction of the Bugatti as a reason not to turn the trolley, then we allow the Bugatti to shape the normative scene despite the fact that a child's life is at stake. (Woollard, 2015, p. 157)

Again, it seems as if these explanations should be available within Hooker's rule-consequentialism.

I have focused on my own account of duties to aid. However, if that account does not suit, there are other promising avenues to explore. Noggle (2009) goes on to offer an elaboration of beneficence as an imperfect duty which may well have the resources to distinguish cases where sacrifice is required. My key aim in the section is to not to argue that Hooker should accept one particular account of our duties to aid. My aim is to argue that, on pain of counter-intuitive implications, he should accept an account that distinguishes between different cases in which we could save others. Once different categories of case are recognised, we can avoid over-demandingness without "shutting the gates of mercy" in a counter-intuitive way.

On the account I have proposed here, "disaster" is a term of art. In aid cases, whether a potential outcome counts as a disaster (such that the agent is required to make certain sacrifices to prevent the outcome) depends on (a) what kind of sacrifice is required; (b) the relationship between the agent and the potential victim and (c) the relationship between the sacrifice and the disaster prevention. In prohibition breaking cases, whether a potential outcome counts as a disaster (such that the agent is permitted to infringe a given prohibition to prevent the outcome) depends on the context and on the prohibition in question. This means that a potential outcome could count as a disaster in one sense and as not a disaster in another sense *for the very same agent, the same behaviour and the same action*. Suppose John could prevent one Londoner from being tortured for a year, only by both lying to Harriet and submitting to be tortured himself for 3 months. It may well be permissible, but not obligatory, for John to prevent the torture of the Londoner by this means. So on this account, the torture of the Londoner would be a disaster (relative to the prohibition on lying) and not a disaster (relative to a requirement for John to submit to 3 months' torture). In addition, the torture of the Londoner may well be bad enough that it is reasonable for John to make this sacrifice. This gives us a third type of disaster claim: the claim that a potential outcome is a disaster such that the agent has a strong reason to prevent the outcome and that therefore the agent would be reasonable/rational to make such sacrifices.¹³

I admit that things are getting pretty complicated here. Nonetheless, I think recognising these different ways something can be a disaster is useful. As Frances Kamm has noted, talking about the strength of a duty to save can be misleading. Kamm notes that an agent may be required to make great sacrifices to save A and not required to make such sacrifices to save B, but still permitted to save B rather than A. The duty to save A is stronger than the duty to save B in some respects but not others (Kamm, 2007, pp. 388–391). Talking about the strength of duty *without qualification* tempts us to overlook these complexities. I think that talking about a disaster *without qualification* also tempts us to overlook important complexities. When we say that a potential outcome is a disaster, we should always check whether we need to qualify what type of disaster claim we are making.

¹³ I thank Chris Woodard who prompted me to reflect on this issue. See Woodard (2022) in this volume for excellent discussion about how someone who accepts Hooker's rule-consequentialism can, and should, think about reasons for action.

3 | CONCLUSION

I have explored two issues related to the disaster clause in Hooker's rule-consequentialism: first, whether the notion of disaster is vague in a way that undermines rule-consequentialism and second, whether rule-consequentialists can avoid demandingness without implausibly holding that enough overall sacrifice entitles an agent to refuse to respond to an emergency. I have agreed with Hooker—and disagreed with Kahn—that when it comes to infringing prohibitions to prevent disaster, the sense of “disaster” relevant to rule-consequentialism must be both relative to the prohibition in question and vague. I have argued that the optimific code should not permit agents to refuse further sacrifices after their overall aggregate reaches a certain level. Instead, we should distinguish between different types of cases where we could save someone's life. What counts as a “disaster” is also relative when it comes to requirements for agents to make sacrifices to prevent disaster. Drawing these distinctions is required anyway to avoid counterintuitive implications. Once such distinctions are drawn, we are able to avoid over-demandingness without setting limits to overall sacrifice in a counter-intuitive way.

ACKNOWLEDGMENTS

This paper was presented at Bradfest or Rules to Live By: A Conference in Honour of Brad Hooker, University of Reading, 24th April 2021. A later version was presented at the Claims in Distributive Justice Workshop, Institute for Future Studies, Stockholm, 16th June 2021. I thank Charlotte Newey and Luke Elson for organising Bradfest and Julia Mosquera and Anders Herlitz for organising the Claims workshop. I thank audiences at both events for helpful comments. I thank Chris Woodard for an insightful commentary. Most of all, I thank Brad Hooker for both his deep and significant contributions to a breathtaking range of topics in philosophy and for his apparently infinite generosity. I would not be where or who I am today without the support he has given me over the years. I know that many others would say the same.

ORCID

Fiona Woollard  <https://orcid.org/0000-0001-5144-3379>

REFERENCES

- Arneson, R. (2005). Sophisticated rule consequentialism: Some simple objections. *Philosophical Issues*, 15, 235–251. <https://doi.org/10.1111/j.1533-6077.2005.00064.x>
- Hooker, B. (2000). *Ideal code, real world*. Oxford University Press.
- Hooker, B. (2005). Reply to Arneson and McIntyre. *Philosophical Issues*, 15, 264–281. <https://doi.org/10.1111/j.1533-6077.2005.00066.x>
- Hooker, B. (2020). The role of rules in consequentialist ethics. In D. Portmore (Ed.), *The Oxford handbook of consequentialism* (pp. 441–462). Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780190905323.013.12>
- Kahn, L. (2013). Rule consequentialism and disasters. *Philosophical Studies*, 162, 219–236. <https://doi.org/10.1007/s11098-011-9756-8>
- Kamm, F. (2007). *Intricate ethics: Rights, responsibilities, and permissible harm*. Oxford University Press.
- Noggle, R. (2009). Give till it hurts? Beneficence, imperfect duties, and a moderate response to the aid question. *Journal of Social Philosophy*, 40, 1–16. <https://doi.org/10.1111/j.1467-9833.2009.01435.x>
- Parfit, D. (1984). *Reasons and persons*. Oxford University Press.
- Unger, P. (1996). *Living high and letting die: Our illusion of innocence*. Oxford University Press.
- Woodard, C. (2022). Reasons for rule consequentialists. *Ratio*, 35. <https://doi.org/10.1111/rati.12352>
- Woollard, F. (2015). *Doing and allowing harm*. Oxford University Press.

How to cite this article: Woollard, F. (2022). Hooker's rule-consequentialism, disasters, demandingness, and arbitrary distinctions. *Ratio*, 00, 1–12. <https://doi.org/10.1111/rati.12354>