Must We Be Perfect?: A Case Against Supererogation

Penultimate Draft

Abstract: In this paper we offer an argument against supererogation and in favour of moral perfectionism. We argue three primary points: 1) That the putative moral category is not generated by any of the main normative ethical systems, and it is difficult to find space for it in these systems at all; 2) That the primary support for supererogation is based on intuitions, which can be undercut by various other pieces of evidence; and 3) That there are better reasons to favour perfectionism, including competing intuitions about the good-ought tie-up, and the epistemic preference for theoretical simplicity.

Key Words: Supererogation; Perfectionism; Normative Ethics; Parsimony; Reasons; Intuitions

1. Introduction[1]

Supererogatory actions are a putative category of moral actions which go above and beyond what we are obliged to do. Since Urmson's famous 1958 paper, 'Saints and heroes', the topic of supererogation has benefitted from a surge of fascinating new literature. Most of us are familiar with many common examples of supererogatory actions: the soldier who throws himself on a grenade to save his friends; the doctor who decides to leave his practice behind to serve during a malaria outbreak in a third-world country; the millionaire who gives away his entire fortune to charity; etc. Many people find intuitive the thesis that such people have performed a task so demanding that we could not possibly expect it of someone, and hence that such acts must be supererogatory.

Yet this intuitive support is surprisingly fragile. There are undercutting defeaters or plausible alternative interpretations for the most prominent supporting intuitions, and supererogatory action is so poorly integrated into normative ethical theories that as soon as the intuitions are defeated, much motivation for including it in one's normative ethical theory dissolves. We will argue, on these grounds and others, that we ought to give up the idea of supererogatory action in favour of the more robust and well-supported view that the best and the required are one and the same.

2. Thesis Clarification

Let us begin with some clarifications regarding our thesis: this will allow precision of exposition as well as the refutation of certain philosophical objections to our view. We are arguing for moral perfectionism, the view that no actions are supererogatory – that there is no such thing

---

[1] We do not think it is possible to offer a comprehensive defence of our position in one paper. We appreciate, therefore, that some important objections to our thesis will inevitably go unanswered. It is our hope that we will not be misinterpreted as being unduly lenient on our own thesis. We have tried to identify the best arguments and objections for presentation here, with the hope of responding to further worthwhile objections at a later date.

as going 'beyond the call of duty'.[2] The best possible action is always obligatory, on our view. Defences of perfectionism in the literature are relatively rare except, perhaps, as part of general defences of consequentialism. Our work here is unique in the literature simply in virtue of offering an argument for this position that is not attached to a particular normative framework.

Defining Supererogation

'Supererogation' is often understood as meaning 'good but not obligatory'. This is inadequate. The reason for this is that actions may be *less* than obligatory but still, in some sense, good. This is difficult to comprehend on the ethically minimal views we are attacking, since on those views obligations tend to be easily achievable things which are *bad* if not done – like not killing innocent people. But if perfectionism is at least possibly true, then fulfilling obligations may be a much more stringent task than we might have considered. And in this case, an action which falls short of being the best possible action – and which is therefore not obligatory – might still be a (relatively) good one. For example, if utilitarianism is true, then the best use of money is probably donation to an incredibly effective charity like the Anti Malaria Foundation. This is, perhaps, obligatory. But donating it to a less effective charity – say, Compassion UK – might still be a good thing to do. But clearly in this case the utilitarian would say that donating to Compassion UK is not going beyond the call of duty, nor does it involve going the extra mile, or going above and beyond obligations. It falls short of our obligations. And so there are at least two *possible* ways in which an action can be good but not obligatory – by going beyond our obligations, or by falling short (though not too far short) of our obligations. We take it that supererogation is concerned only with the former, and we propose to christen the latter category 'subobligation'.

---

[2] There are many uses of the term 'perfectionism' in philosophy—we use the term as stated above. Of course, one may agree that it is impossible to act supererogatorily, but disagree that we ought to be perfect, given something like moral nihilism. Therefore, a general 'ought' to be perfect would follow from our argument only if moral nihilism is false.

Supererogatory acts are often understood as being morally praiseworthy acts that are optional. Benn (2018a) writes,

> Like morally required actions, supererogatory actions are not forbidden and are thus permissible; unlike morally required acts, they are also 'beyond duty'. As such, supererogatory actions are 'optional' in the sense that they are neither morally forbidden nor morally required. (p. 2)

Benn's use of 'optional' to refer to acts that are not morally forbidden or morally required follows McNamara (1996) in being a distinctly moral sense of optionality, and constitutes a feature of supererogatory acts that she claims all views of supererogation agree on. She further argues that, while such actions are praiseworthy by nature, they needn't require extreme amounts of sacrifice to be supererogatory; rather, they must only be comparatively costly (p. 9). We here want to agree with Benn and others on the optionality of supererogatory acts and remain neutral on the question of how costly they must be in order to be supererogatory, as such a question is tangential to our main argument.

We will postpone further detailed explication of our view. But we note the following features of perfectionism, which are important for our defence of it: firstly, perfectionism does not entail that there is only ever one right action. There may be many joint best actions available at any one time. Secondly, perfectionism, though it sets high standards, is still sensitive to varying degrees of goodness. Some non-obligatory actions may still be better than others. We will later argue that this spectral quality of morality actually fits better with perfectionism than with supererogationism. Finally, our version of perfectionism is consistent with a number of different normative ethical theories, and with a number of different views on blameworthiness, culpability, and freedom.

3. Where does supererogation live?

In surveying the moral landscape, it is helpful to ask ourselves where we might find the home of supererogatory actions. Much of the literature on supererogation uses the language of

4

'fitting' or 'making room' with respect to supererogation, due to the initial difficulty of situating the category among the rest of various frameworks' normative furniture (McNamara 1996, Heyd 2015b). Because talk of the supererogatory often involves talk of obligations, it is important that a theory provides an answer to the questions: What are we obliged to do, and how are we to distinguish these obligations from actions that are supererogatory? In our attempt to pin down which frameworks generate the category of supererogation, we will look at typical formulations of three prominent normative theories—Kantian deontology, Utilitarianism, and virtue ethics—as well as non-standard versions of these theories and some competing frameworks.

Five Aims: Methodological Preliminaries

This section of our paper has five aims which we think are important for our thesis. Aim 1) is targeted: it addresses those who hold to one of these normative theories and commends perfectionism to them. We take it that there are at least some reasons to hold to the theories we consider, and we think that insofar as they are plausible theories, for those who hold to them, there is commensurate reason to reject supererogation. Aim 2) is a related one: we take it that there is at least some justification for the theories we consider, and that insofar as there is, and insofar as most people have a degree of agnosticism about which ethical theory is true, to the same extent they should avoid supererogationism. Aim 3) is to try and show that, at best, viable moral theories tend to be *consistent* with supererogation rather than actually *generating* the category. Most theories we consider are *in tension* with supererogation, rather than just not providing a natural home for it. So even if we can find some theories which make room for supererogation, this will not necessarily be enough to give supererogationism a reasonable probability even given those theories. Aim 4) is to illustrate the theoretical complexity required for systems involving supererogation. In order for a normative theory to convincingly suggest supererogationism, it is important that it not do so merely in virtue of being extremely open to add-on normative frameworks. Aim 5) is to show that in most cases, belief in supererogation results primarily from

first-order, applied-level intuitions about the extent of our obligations, rather than being derivative of higher-order ethical theories. This is important because it highlights the vulnerability of these beliefs: rather than being firmly grounded in robust ethical theory, they result primarily from intuitions, which we think are significantly undermined by the considerations we identify later in the paper. Since we argue that *at best*, simple theories tend only to *permit* supererogation, we think this shows that belief in supererogationism is based primarily on first-order intuition rather than on broader normative theory.

Deontology

Given that supererogation is defined largely by its relation to moral obligations, it seems natural to try to locate this category in Kantianism. However, Kantians are largely in agreement that such a category is senseless (Baron 1987, p. 237-238).[3] The common Kantian argument against such a category goes as follows:

> P1: An action has moral worth in so far as it is done from duty to the moral law.
> P2: Supererogatory actions, by definition, are not done from duty to the moral law.
> C: Therefore, supererogatory actions do not have moral worth.

This is a problem, because moral worth – indeed, superior moral worth – is an essential feature of supererogatory actions. But traditional Kantianism has no concept of moral worth as being separate from the sense of dutiful conformance to the moral law with which an action is done. For Kant, no action is unqualifiedly morally good which is not done with good will.[4] 'In Kant's terms, a good will is a will whose decisions are wholly determined by moral demands or, as he often refers to this, by the Moral Law' (Johnson and Cureton 2016). Because supererogatory actions are, by definition, actions which are *not* morally demanded, they cannot be done with good will. It is

---

[3] 'Some, like von Herbert, believe that Kant's ethics asks too little, requiring only minimal decency and saying nothing about the 'higher flights of morality.' Others, reading the texts differently, believe that his theory asks too much, demanding total devotion to morality and treating everything worth doing (and perhaps more) as a duty. But, despite their differences, the two sets of critics concur in taking the central problem to be Kant's failure to recognize the category of the supererogatory.'

[4] 'There is nothing it is possible to think of anywhere in the world, or indeed anything at all outside it, that can be held to be good without limitation, excepting only a **good will**.' (Emboldened text our own). Kant, *Groundwork for the Metaphysics of Morals*, 4:393.

6

therefore difficult to see how supererogatory actions could be morally superior to obligatory actions, since only the latter are capable of containing the only unqualified moral good. The Kantian is forced to admit that there are no *morally* valuable acts that one is not bound by duty to perform. Many Kantians cite Kant's own unease with acts of exceeding virtue, fearing them to be acts of vanity, or self-congratulatory. Those who do not accept this interpretation of them suspect that Kantians may need to claim that these acts, generally thought to be supererogatory, are actually required. (Timmerman 2005)

Those who have tried to position supererogation within a Kantian framework typically turn to Kant's distinction between perfect and imperfect duties. (Hill 1971) An imperfect duty is, simply put, a duty to do something *sometimes*. Importantly, an imperfect duty is not the state of occasionally having a duty to do something; rather, it is a duty one always has, which is vague in its terms. In his paper 'Kant on Imperfect Duty and Supererogation', Thomas Hill Jr. proposes some criteria for a supererogatory act:

(a) is of a sort commended by a principle of wider imperfect duty,

(b) is motivated by a sense of duty (or, perhaps, respect for moral reasons),

(c) is neither forbidden nor required by another, more stringent duty.

(d) is in a context where no alternative is required by more stringent duty and there is at least one alternative that is neither forbidden by more stringent duty nor commanded by other principles of wide duty, and

(e) is done by an agent who has adopted the relevant principle of wider imperfect duty and has often and continually acted on that principle. (p. 71)

The common example of an imperfect duty is charitable giving. If I am required by duty to sometimes give money to charity, then it might be thought that any *particular* act of charitable giving is supererogatory – good to do, but not required.

However, it is not clear that frequently performing an imperfect duty could account for supererogatory actions. Hill's criterion that a supererogatory act be motivated by a sense of duty has been criticised by Marcia Baron (1987): 'The real difficulty is that supererogationists are trying

to get away from duty, but the Kantian analysis keeps pulling what is supposed to be 'beyond' duty back under the rubric of duty.' (p. 244) It is far from clear even on initial inspection that an action done according to a sense of duty (as Hill's supererogatory actions are) could be supererogatory. But the most alarming feature of Hill's case is seen when bearing in mind that supererogatory actions are supposed to be *better* than what duty requires.

Given this, Hill's conception of superogatory actions simply does not account for their status as morally superior. Take, for instance, an imperfect duty to feed your child. This is very plausibly an imperfect duty, since we are required to sometimes feed our children, but we ought not to always be feeding them, and it is not usually necessary to feed them *at any particular time*. Certainly feeding my child more and more often would not account for better and better actions – overfeeding my child could be as pernicious as underfeeding her. And it is clearly implausible that a given act of feeding my child should be considered above and beyond the call of duty! And so, even though handing my child an extra apple or feeding them *at any particular time* could perhaps be considered 'good but not required', this is not because they exhibit any peculiar virtue: it just reflects the fact that there is not always a unique, determinate course of virtue. We again appeal to the insufficiency of the good-but-not-required definition of supererogation, which, as we previously noted, could also pick out 'good' subobligatory actions or joint best actions.

A conception of supererogatory acts must be able to account for their *moral superiority* over actions which are obligatory. Kant's system can account for certain non-moral goods – for instance, aesthetic good, or happiness – which some action may aim at. Yet, again, because the only good in Kant's moral philosophy is good will, it is difficult to see how appealing to *non-moral* goods would be able to account for the perceived *moral* superiority of supererogatory actions.

Because the problems with finding supererogation in Kantianism seem to be the result of Kant's constraints on sources of goodness, some may look to other deontological systems to find the category. For example, a theory which does seem to allow room for supererogationism is Scanlon's contractualism; however, the primary reason this theory is consistent with

supererogation is simply because it has such limited scope. Scanlon (1998) has a theory of what makes something wrong, a theory which itself has some complexity. 'An act is wrong if its performance under the circumstances would be disallowed by any set of principles for the general regulation of behavior that no one could reasonably reject as a basis for informed, unforced, general agreement.' Similarly so for right actions: if an action wouldn't be disallowed by the same sort of principles, then the action is right to do. This alone, however, is not enough to generate supererogation. Part of our understanding of going above and beyond one's obligations is that the action is better than an action which merely fulfills an obligation. On Scanlon's framework of right and wrong action, there is no account of goodness. Therefore, it is hard to imagine how to conceive of supererogatory actions being better than obligatory actions. Because of its limited scope, in order to actually give an account of right action and of moral value, it looks like Scanlon needs to adjoin a completely independent account of an action's goodness, and the resulting hybrid will involve considerable theoretical complexity - which we take to be a theoretical vice later in the paper.

Utilitarianism

Utilitarianism is, of course, a framework known for its demanding moral obligations, arising from the lack of space for supererogation. A general utilitarian argument against supererogation can be given as follows:

> P1: Supererogatory actions, by definition, go beyond what is required of us.
> P2: We are required to choose actions which maximise good results over ones which do not (standard utilitarianism).
> P3: Doing the maximally good action in any situation is generally considered supererogatory (i.e., giving away the majority of our income to charity).
> P4: But there is a contradiction: supererogatory actions can't be required (P1).
> C: Therefore, no actions are supererogatory.

The standard 'good-ought tie-up', common to nearly all consequentialist frameworks, creates the extreme difficulty in positing a category of actions which are neither required, nor

prohibited.[5] What, on traditional utilitarianism, can we say is required, besides acting to maximize good results? And indeed, we ought not to do anything less. So actions which would typically be conceived of as being supererogatory will either maximize the good, or they will not, and we are required to do or not do them based on their outcome. Therefore, we cannot, on standard utilitarianism, conceive of actions which are very good to do, but not required.

Rule utilitarianism is sometimes thought to allow room for supererogationism, or perhaps even to generate the category. Rule utilitarianism holds that we are obligated to always follow rules that, generally speaking, maximize utility when adhered to. For instance, suppose a rule dictates that one should give 20% of their income to charity, because doing more than that would weary or frustrate them, causing them to have a bad attitude and to resent or cease all charitable giving, or because giving more than that would produce more bad than good due to resulting financial distress. Then, for Mary, there are two choices of actions which may be optimal: she could give 20%, or she could give more than 20%. The supererogationist might say that giving at least 20% is obligatory, while giving more than 20% is supererogatory.

Rule utilitarianism is sometimes understood in at least two ways: as something like a pragmatic version of act utilitarianism, or as a strict set of rules generated by rules-consequences relationships. Under neither conception is it easy to see how Mary could act supererogatorily. If Mary is a pragmatic act utilitarian and thinks she is vulnerable to altruism-weariness then giving more than 20% would be morally inferior since it can be expected to have worse consequences. In order for rule utilitarianism to not collapse into act utilitarianism, obligations must be based on expected, and not actual, utility.

> [I]f full rule-consequentialism claims that an act is wrong if and only if the *foreseeable* consequences of rules allowing that act are sub-optimal, rule-consequentialism cannot hold that an act is wrong if and only if the *actual* consequences of rules allowing that act will be sub-optimal. (Hooker 2016)

---

[5] Famously, Moore (1948).

If she is a pragmatic act utilitarian but is confident that she is not vulnerable to weariness then it seems as though she is obliged to give more, since the rule is there only to avoid the computational problems of calculating consequences for every individual act. But if better consequences from a rule-deviant action are readily recognised in a particular case, the pragmatic concerns disappear, and it would be supremely irrational for a pragmatic act utilitarian to preserve the rule when the pragmatic reasons for the rule are no longer germane.

If Mary is a strict rule utilitarian, on the other hand, then—since following the rules takes precedence over particular circumstances—she is obliged to follow the rule and give only 20% no matter what. Giving more constitutes a wrongful failing of obligation, not a morally superior supererogatory act. So, it does not seem as though rule utilitarianism will simply or clearly make room for supererogation.[6]

Virtue ethics

There has been a recent increase of philosophical interest in exploring supererogation within virtue theoretic frameworks. Although at this point virtue theory may appear to be the most promising avenue for accommodating supererogatory actions, here we will also struggle to find a natural place for the category. In the words of Heyd (2015b), 'It turned out that fitting supererogation into virtue-based moral theory proved to be a more difficult task than doing so in utilitarian and deontological theories.'[7] But seeing why this is requires understanding the unique virtue theoretical understanding the world. Virtue ethics calls good that which promotes flourishing, particularly the flourishing of an agent's character, to the fulfilment of her *telos*. Standard conceptions of right actions from a virtue theoretic perspective, such as Hursthouse

---

[6] J.S. Mill does recognise a distinction between obligatory actions and non-obligatory meritorious actions in 'Auguste Comte and Positivism'; however, he does not explain how this would fit into a utilitarian framework, and indeed Mill's utilitarianism is notorious for being (at best) at the fringes of what would today be labeled 'utilitarianism'.

[7] The reason for Heyd's difficulty was that the picture of the world assumed by virtue theories did not map on directly to the liberal conception of the post-enlightenment world from whence ideas of supererogation arose.

(1999), take a form similar to this: 'An action is right iff it is what a virtuous agent would characteristically do in the circumstances.' (p. 28).

But recent work on virtue ethics and supererogation have rejected Hursthouse's model, instead favouring a 'target-centered approach' (Swanton 2001). Rebecca Stangl (2015), who adopts this approach herself, believes that supererogationism makes sense on this account of right action. In Stangl's paper, 'Neo-Aristotelian Supererogation', she quotes from the Nicomachean Ethics:

> [T]here are many ways to be in error – for badness is proper to the indeterminate, as the Pythagoreans pictured it, and good to the determinate. But there is only one way to be correct. That is why error is easy and correctness is difficult, since it is easy to miss the target and difficult to hit it. And so for this reason also excess and deficiency are proper to vice, the mean to virtue; 'for we are noble in only one way, but bad in all sorts of ways.[8]

Roger Crisp (2013) argues that we do not have to interpret this passage to mean that, in any given circumstance, there is only one correct action. However, Crisp does think that this passage implies a strict dichotomy. Either an action of ours hits the mark, in which case it is virtuous, or it does not, in which case, it is vicious. And so, all virtuous actions are equally virtuous, and we are obliged simply to only perform virtuous actions. On Crisp's account, there is no room for supererogation in an Aristotelian account of virtuous action. Stangl responds to Crisp that it 'is implausible to think that all virtuous actions are equally virtuous.' (p. 353)

We agree that this seems implausible. But it also seems, on this view, as though there must be an action, or multiple actions, that are *optimally* virtuous. Stangl believes that actions which are more virtuous than they need to be in order to 'realize their end' or 'hit the target' are supererogatory. But conceive of it this way: the target is presumably either extended (in the sense of including non-best possible actions) or punctiliar. If punctiliar, then clearly there is no room for supererogation. If extended, then there might be. If the target is extended but admits no degrees of virtuosity, then it will allow no supererogation. But if it admits of degrees of virtuosity, then

---

[8] Aristotle, *Nicomachean Ethics*, 1106b, 29ff., quoted in Stangl, pp. 351-52.

there is room for supererogation. The problem here, however, is explaining why the real target is not the most virtuous option. Unless there is an adjunctive theory explaining why the boundaries of the target should be wider than simply the most virtuous possible option, and why they fall where they do, such wideness is arbitrary. For these reasons, the target-centered approach to virtue theory does not more easily allow for supererogatory actions than the standard approach does.

These arguments from some of the major normative ethics systems are not intended to be conclusive. There are, of course, several deontological, consequentialist, and virtue theoretical arguments that purport to make room for supererogatory acts (respectively, Brinkmann 2015, Hooker 2002, Kawall 2009). The point of this overview was just to show that, on the face of things, none of the three major normative theories seem to have a *natural* place for them. Perhaps space can be created in some or all of these systems; but now the question becomes: why make the space? What motivates us to try to fit such a category of actions in our normative ethical framework?

4. **Intuitions and supererogation**

We suggest that the main support for supererogatory actions comprises a variety of related intuitions. This is not to say that we think intuitions are the *only* reasons able to be given in favour of supererogationism; famously, the Catholic intellectual tradition accepts the category as a way of distinguishing ordinary people from saintly, or particularly excellent, people (Alston 1989, Wynn 2015). Additionally, one may have other reasons for finding a satisficing normative framework convincing, and conclude that, therefore, supererogatory actions must be possible. This is not necessarily a slur – we also take it that almost all ethical opinion is based on intuitions of varying kinds. But it is very plausible that the main basis for most people's belief in supererogation is their intuitive belief in propositions like the following:

'Forgiving my son's killer would be supererogatory'
'I am not obliged to give everything I own to the poor'
'The Good Samaritan went beyond the call of duty'

These intuitions are fairly widespread, and fairly easy to arrive at oneself. There is certainly a sense in which many intuitions which seem to favour supererogationism. These intuitions are of varying kinds: intuitions that morality is not as demanding as perfectionists suggest, that we are not blameworthy for failing to achieve the best, that the best moral acts are in some sense optional, and so on. We propose that, on further evaluation, any intuitive support for supererogation collapses on a number of levels.

We note firstly that even the general trend of intuitions supporting supererogationism is doubtful. As we will later show in the discussion of linguistic simplicity, it is difficult to see what obligation even *means* in supererogatory systems. It is therefore difficult to see how anyone could genuinely have the above intuitions with an adequate grasp of the concepts involved. Moreover, intuitions in support of perfectionism are common. Heyd (2015a) notes:

> This 'good-ought tie-up' is a theoretically attractive principle: whatever is good, ought to be done. If an action is good, then there must be reasons for doing it. If it is the best possible action, the reasons for doing it are conclusive, that is outweighing all other reasons for not doing it (or doing something else). Failing to do the best action cannot therefore be immune from blame or condemnation. Or, in other words, doing the best is always obligatory, never optional. Supererogation is impossible.[9]

This is related to the 'paradox of supererogation', which asks how it is possible that a good is not a duty. As we have explained, there is no natural answer for this in any of the major normative theories surveyed. Indeed, the main normative theories arguably lead us to expect that goods are always duties. We will reinforce this intuition with our discussion of theoretical simplicity anon. The intuitions favouring the theoretical simplicity of perfectionism are quite strong, and must be considered as evidence alongside the evidence from intuitions about supererogation.

As well as doubting the content of supererogatory intuitions and noting the strong intuitions in favour of perfectionism, we can offer a variety of undercutting defeaters against

---

[9] To avoid mischaracterizing Heyd, we note that this quote does not describe his own view (in his book *Supererogation* (1982) he defends a supererogationist framework) but rather the view of someone who holds to the good-ought tie-up.

supererogatory intuitions. We begin with some brief considerations against particular pro-supererogation intuitions.

It is thought by some that a good action's being optional adds some value to it. This is taken to support supererogationism since supererogationism makes certain acts morally optional. Heyd (1982) argues that this 'unqualified' freedom to act or not act is precisely what makes supererogatory actions so praiseworthy—this freedom entails that the action was performed, not from fear of sanction, but from personal preferences and values (p. 175). We appreciate this intuition and want to say two things in response. First, it might be the case that if supererogationism is true, then there is a *distinct* kind of value in its moral optionality that cannot be posited on perfectionism. We do not need to deny this here, for our aim in this section is not to show that supererogationism is false – only that the intuitions in support of it can be explained in a way that is consistent with perfectionism. For this, all we need to do is give a plausible explanation of why it *seems* very strongly to us that optional acts have distinct value in a way that is consistent with perfectionism.

Second, we submit that these intuitions can be explained without the kind of optionality that supererogationism provides. There are at least two other sources that are adequate to explain the goodness of these kinds of actions: libertarian free will and unrewarded goods. Libertarian choices for the good are thought to be good precisely because the agent did not have to perform it – there is a sense in which the action was optional, even if that is not a moral sense. We can think of this kind of optionality as the freedom from threat of sanction. For instance, one has the option of forgoing parental rights, even if doing so is morally wrong. If there is no threat of sanction, then we recognise that the act was not performed in order to avoid such sanction. This is augmented when the choice is unrewarded. It is very plausible that the main reason optional acts are good is because they are done truly selflessly – there is no reward in them, no obvious gain for the agent. So, it is perfectly plausible that our intuition that optional acts are good is derivative of our appreciation for selflessness, as the supererogationist admits. Where we believe the

supererogationist errs is in her assertion that such selflessness of motivation can only be found on a perfectionist framework. Indeed, this selflessness does not seem to have very much to do with supererogation.

It is objected at this juncture by Rebecca Stangl that these explanations of the exceptional goodness of more demanding acts are not the best way to account for certain mundane scenarios.[10] [They] point to the example of missing a flight through their own negligence and arriving at the destination at 2am instead of 8pm. The sister collecting them from the airport does better if she still agrees to come out at 2am rather than the previously agreed 8pm. Since neither case exemplifies a greater degree of libertarian freedom, and since the sister receives no reward in either case, these perfectionist explanations do not seem to explain the relative quality of arriving at 2am.

The above account is unclear, since it compares two actions in two different situations. Presumably the best (or joint best) action in both cases is to pick up the passenger when they arrive - after all, if the passenger arrived at 8pm as planned, it would hardly be morally superior for the sister to arrive at 2am! It just happens that the best thing to do in the punctual case is not too great a sacrifice for the sister. This should not surprise us: we already know that morality is sometimes easy, and sometimes more difficult. Perfection demands more at some times than others. So, although the perfectionist will say that they are equally good in some sense (in that they are both best actions given each circumstance), they can still grant that the non-punctual case involves a better action in another sense. The challenge is whether perfectionists can give a good explanation of this 'betterness' without recourse to supererogation.

In fact, we can: namely, the sacrifice and selflessness involved. As [our interlocutor] writes, 'the obvious answer is that the [2am] pickup costs my sister more.' They go on to conclude that in doing so, the sister 'goes the extra mile'. We agree fully with the initial diagnosis: indeed, the suggestion that some 'optional' actions are better on account of their being more selfless is already

---

[10] In comments on an earlier draft of this paper given at the Logos Workshop, University of Notre Dame, 2016.

present in our suggestion that the good of unrewarded virtue is derivative of the good of selflessness. This explains the situation in a way that is consistent with our *common* intuitions, and yet reveals no need to involve notions of supererogation.

One worry with cases like the airport scenario – and perhaps a subconscious worry with perfectionism more generally – is that suggesting that we are obliged to perform these demanding acts of kindness yields an aberrant view of moral life: it seems to suggest that the moral life consists of no more than a multiplicity of extravagant, unrelated and disproportionate acts of kindness and self-sacrifice – an intuition not consonant with most people's view of the ideally virtuous person. Susan Wolf famously raises a similar critique in 'Moral Saints', arguing that the lives of 'moral saints'—which she defines as people whose every act is as morally good as possible (p. 419)—neglect other goods and values in the attaining of solely *moral* perfection. Benn (2015) argues that the demandingness of perfection takes the form of a constraint on the *autonomy* of the agent, shrinking her set of permissible actions down to a small pool. She coins the 'Confinement Objection', which holds that a moral theory is implausibly demanding if it constrains the set of actions available to us in such a way. This is one sort of demandingness objection: that the demands themselves are too difficult or constricting to be required of, or even good for, the agent.

In another paper, Benn (2018b) argues for a different sort of demandingness objection: that perfectionism, when *demanded of* someone, undercuts itself by making it even harder for the agent to do good. '[T]he consequences of believing ourselves to be morally obliged to do what is morally best are likely to be depression and anxiety, especially given that moral standards are often viewed as having an external source.' (p.17) Such mental strain, she argues, can cause us to focus too much on avoiding sanctions. '[B]y trying to avoid failing to meet the standards, we tend to avoid meeting them.' (p. 18) This critique of what she calls 'psychological perfectionism' recalls Parfit's well-known critique of consequentialist moral theories (1984), where he argues that such theories are 'self-effacing' (part 1, section 10, pp. 24-25). That is, by trying or intending to fulfil the aims of some consequentialist theories (e.g. maximizing overall happiness) we prevent ourselves

17

from fulfilling the these aims as well as we might otherwise have done had we mentally accepted some other normative framework.

It would be ignorant to deny that moral anxiety can impede one's ability to fulfil their obligations. We hope we have made it clear throughout the paper, however, that perfectionism is not committed to any detailed view of what the perfect moral life looks like – any such caricatures are a function of a presumed set of values not central to (or even associated with) perfectionism. Perfectionism is compatible with the thesis, for example, that the perfect life involves considerable amounts of mundane actions oriented towards particular goals and callings, rather than a disconnected series of extravagantly and unnecessarily sacrificial actions. It does make room, of course, for extravagant sacrifice.

Additionally, Benn's picture of moral anxiety seems not to apply to normative frameworks like virtue ethics, which standardly denies deontic concepts like obligations, duties, rights, etc. Virtue ethics generally prefers a picture where there are no distinctly moral 'oughts', and holds that to say we ought to do or ought not to do something is just to describe how close to, or removed from, the ideal human such behaviour would be. Perfectionism may be true in the absence of obligations, so long as there is no way to go beyond or do better than what we ought to do.

Another pro-supererogation intuition is that we are not always blameworthy when we fail to be perfect. Likewise, this is easily explicable without the resources of supererogationism. A very simple reason exists for this: we are not always blameworthy when we fail to be perfect because there are exculpating factors other than an action not being obligatory. For example, exculpating factors include ignorance, practical inability, extreme demandingness, and so on. Perfectionists are not committed to the view that anything less than perfection is blameworthy, as we note later in the context of responding to various other putatively pro-supererogation intuitions.[11]

---

[11] In any case, why think that we are not always blameworthy? Maybe we are just really, really bad. We wish to leave this position open.

Must We Be Perfect?

Let us turn to a more global concern about moral justification. Concerns have been raised regarding the possibility of genuinely justified moral knowledge – most notably, on the basis of evolutionary debunking arguments.[12] We believe that there is considerable force to these arguments, in which case most – if not all – moral claims are suspect.

The global success of sceptical arguments is not required for our thesis, however. We can still grant that intuitions reliably grant us a significant amount of moral knowledge – about whether pleasure is a good, for example. But on the face of it, whether a good action is supererogatory or obligatory seems so obscure that it is incredibly difficult to see how we could have any reliable knowledge about this at all. If obligation does not fall at the end of the moral spectrum, how do we know where it falls? Are almost all good actions obligatory? Are virtually no actions obligatory? How are we supposed to judge? And how is it at all plausible that the selection pressures facing early hominids were ones that helped them to make these sorts of obscure metaphysical judgments? We will develop this point later in our discussion of maximal degree simplicity. But for now it is worth noting the ease with which scepticism arises when we ask how we know where the dividing line between obligatory and supererogatory sits and why it sits there.

A final intuition that seems to support supererogationism is the intuition that 'we're not really that bad'. Perfectionism is often associated with the idea that buying a frivolous cup of coffee constitutes a shirking of our obligations and is a blameworthy act (though later in the paper we will show why these claims do not necessarily follow from perfectionism). It seems implausible to many people that they fail to meet most of their obligations and fail so frequently. While it is consistent with our thesis that most people fulfil most of their obligations (though we consider it unlikely), we note here simply that this intuition about our general goodness is not as secure as we would like. In the first place, the empirical data are clear that most of us are not very good when it comes to certain evils. For example, most of us are content with living in luxury while others

---

[12] E.g. Street, 'A Darwinist Dilemma for Realist Theories of Value'.

live in abject poverty, even when these facts are well-known to us and small sacrifices of our own could relieve considerable poverty elsewhere in the world. Most of us are also perfectly at ease with apathetic and thoughtless contributions to corporate evils, and most of us are small steps away from being moral villains if only our life circumstances turned out differently. It is quite likely, in light of these, that the main reason we are convinced of our own decency is renormalising our moral calibration in light of the fact that everyone around us acts fairly similarly. More on this shortly.

There are empirical data supporting the claim that we are prone to overestimate our own rectitude. Illusory superiority is a well-known psychological phenomenon, according to which people have a tendency to overestimate their abilities in an enormous variety of domains – IQ, popularity, health, driving ability, and so on. This generates the amusing statistic that the large majority of people consider themselves to be generally above average. Particularly interesting is the Dunning-Kruger effect, where it is low-skilled experimental participants who have a particular tendency to overestimate their own abilities, and an inability to recognise their own inabilities. This indicates an analogue in the moral sphere: we would expect people in general to overestimate their own moral stature,[13] and in particular we would expect the most morally bankrupt to suffer from this effect the most. So our tendency to believe that we generally fulfil our obligations is entirely to be expected given the general empirical data, the intuitive plausibility of supererogationism and the results of moral psychology, which all mutually reinforce each other.

This is not the only relevant psychological evidence. Moral self-licensing is another well-documented phenomenon whereby subjects tend to act less morally when they have previously acted morally. Remarkably, people have a tendency to act less morally even just after *imagining* themselves acting morally (see Merritt *et al* 2010). Humans have a tendency to overestimate their own moral goodness, and even to use their own morality as an excuse for acting wrongly in future.

---

[13] This has, in fact, been documented. See Hoorens, 'Self-Favoring Biases, Self-Presentation, and the Self-Other Asymmetry in Social Comparison'.

Even thinking of themselves as moral makes them prone – unknown to them – to act wrongly. This constitutes a further category of evidence showing that humans are probably significantly worse than they generally consider themselves to be.[14]

Even in the absence of empirical evidence, we would probably expect humans to significantly underestimate the number of obligations they have. Why? If most humans are selfish and judgmental, and if they drastically fail to meet their obligations, it is very plausible that they would both lower their expectations and judge themselves relative to those around them. Humans have a very strong tendency to judge others and to compare themselves to others, and so the 'average person' becomes the relevant comparison class for moral self-judgment, rather than the actual standard of obligation. But since most humans are far from perfect, so they reduce their expectations accordingly. Indeed, retaining a high standard of obligations would make inter-human comparison very difficult, since almost all humans would fail drastically and so only be marginally different from each other. There is thus good practical reason for humans to renormalise their measure of obligation by limiting obligations to what might be expected of the average person. This allows a much more sensitive comparison of different humans where we really can insist that we are importantly better than others.

In sum, then, there is ample, independent, empirical plausibility to the thesis that we are not really as good as we are prone to think. This – though not required by perfectionism – is consonant with it, and lends credibility to it as well as undermining any objections predicated on our moral decency.

5. Perfectionism is theoretically simpler

Having undercut much intuitive or theoretical support for supererogationism, we now turn to our central positive argument for perfectionism: that perfectionism is theoretically simpler in a number

---

[14] To be clear, we are not claiming that these experimental results are impossible – or even difficult – to explain on supererogationism. We only suggest that they undercut the force of the objection to perfectionism that we're not so bad – humans are notoriously prone to overestimating their goodness.

of senses. We, along with others, have argued elsewhere that simplicity is an index of truth,[15] though we do not have space to defend that assumption here. Note, however, that in the absence of empirical evidence, simplicity is an *extremely* strong criterion for determining how likely a theory is to be true. So, all else (in particular, explanatory power) being equal, a simpler theory is more likely to be true. We have argued in the previous section that the explanatory power of perfectionism is no worse than that of supererogationism, so that if perfectionism wins on simplicity, it is a stronger theory overall.

There are at least three respects in which perfectionism is simpler than supererogationism: linguistic simplicity, ideological parsimony, and maximal degree simplicity. We will take each of these in turn.

There has been much debate about lexical and syntactic simplicity, mostly to do with the length of expression of some hypothesis or other in a given language. For example, it is often held that hypotheses which can be expressed by a shorter string in given computing languages are simpler and more likely to be true. There are many cases where simplicity in natural languages is preferred, or where there is a preference for predicates occurring in natural language over gerrymandered predicates like 'grue'. Linguistic simplicity is often indicative, though not constitutive, of conceptual simplicity.

Perfectionism is linguistically simple in the following sense: it maintains that 'ought' is univocal. This can perhaps best be seen by contrasting it with what supererogationism says about these kinds of words. According to supererogationism, 'ought' can imply an obligation, as in, 'you ought not to steal that ice cream'. But it can also mean something relating to supererogation, as in, 'you ought to be perfect'. This latter proposition is incredibly plausible. It is good to be perfect, and if moral reasons are intrinsically motivational or reason-giving, then an action's being good gives one a normative reason to bring it about. It is overwhelmingly plausible that the reasons in

---

[15] See, for example, Miller 2016.

favour of an action are intimately linked with the concept of 'ought', and so the more reasons there are in favour of an action the stronger the 'ought'. So, given the overwhelming moral reasons in favour of being perfect, it is highly probable that we ought to be perfect. But the supererogationist finds it much more difficult to explain what is meant by 'ought' in this case. They will have to give a meaning of 'ought' distinct from the special sense of 'ought' they give in assertions of obligation. And this multiplies the senses of 'ought' and so introduces needless linguistic complexity. By contrast, the perfectionist has a simple explanation of 'ought': to say that we ought to do something means that we have good moral reasons to do it, and the strength of the 'ought' varies in line with the strength of those reasons. This is a far simpler view of our moral discourse, and a far simpler view of the relationship between reasons and obligations.

This simplicity is arguably due to a deeper – and probably more relevant – simplicity, viz. ideological parsimony. Ideological parsimony has various aspects, the most important of which is the minimisation of concepts – we want to explain as much of the world as possible with as few conceptual resources as possible. Ideological parsimony may be quantitative—fewer concepts all together—or it may be qualitative—fewer *types* of concepts. These are what make for an elegant and genuinely explanatory system. But it is clear in light of the foregoing discussion that supererogationism involves a gratuitous inflation of both quantity and types of moral concepts employed in our deontic logic. Perfectionists have moral reasons and a conception of 'ought' that strengthens linearly with it and have just two fundamental categories – obligatory and non-obligatory. By contrast, supererogationists have obligatory actions sitting in the middle of the moral spectrum, with two completely different kinds of actions each side: on one side, non-obligatory blameworthy acts, and on the other side, non-obligatory praiseworthy acts. And their conception of 'ought' is tied in a complicated manner to both obligatory acts and supererogatory acts. Or, from a different perspective, they have a moral spectrum with heinous acts on one end and supererogatory acts on the other: a fairly simple system similar to that of the perfectionist. But a further obscure category is then imposed: the merely obligatory, where obligation is a fairly

enigmatic notion which is not easily reducible or linked to any other prior moral concept. On the perfectionist picture, moral reasons are straightforwardly linked to both praiseworthiness and obligations: you are obliged to do (one of) the action(s) you have the most moral reasons for, and the more moral reasons you act on, the more praiseworthy your actions is, and the more moral reasons you eschew, the more blameworthy your action is. On supererogationism, there is no such link between praiseworthiness, obligations, and moral reasons. That is, if moral reasons determine *obligations*, then the praiseworthiness of supererogatory acts cannot be determined by moral reasons. On the other hand, if moral reasons determine *praiseworthiness*, then obligations are not linked to (or are arbitrarily linked to a certain amount of) moral reasons. Assuming perfectionists can explain the moral landscape as well as supererogationists, this constitutes an offensive multiplication of kinds.[16]

The Supererogationist Picture

Less than obligatory	Obligatory	Supererogatory

Less Praiseworthy   More Praiseworthy

The Perfectionist Picture

Less than obligatory	Obligatory

Less Praiseworthy   ⟵ Moral Reasons ⟶   More Praiseworthy

Finally, there is a kind of simplicity which we call maximal degree simplicity. This was introduced by Richard Swinburne (2004, p. 55) and has been defended in more detail by the present authors (Miller 2016). The idea is that maximal degrees of properties are simpler than non-maximal, large degrees of properties. So, for example, omnipotence is held to be a simpler property

---

[16] See McNamara (1996) for an illustration of this point: his very sophisticated analysis of the structure and interrelationships of moral concepts is in part reflective of the multiplicity of concepts he needs to incorporate the supererogatory.

than megapotence. This is relevant because perfectionists hold that obligation is simply reducible to a maximal degree of goodness. This makes obligation a simple – almost trivial – addition to the already simple view of moral reasons and axiology. Supererogationism, on the other hand, has a relatively obscure view of moral reasons and axiology to begin with – it is not clear whether moral reasons generate a stronger 'ought', for example. In addition to this, there is the imposition of a new category, obligation, whose relationship with moral reasons and the strength of 'ought' is entirely unclear. It is not simply reducible to a maximal degree of goodness, but rather it lands fairly arbitrarily on the moral spectrum, with little explanation of why it lands there rather than anywhere else on the spectrum. This gives it a considerable degree of complexity in comparison with the perfectionist view.

We noted earlier that there is no natural place for supererogation in any of the main competing normative ethical traditions. This discussion of theoretical simplicity reinforces these arguments by showing what an unnatural and ontologically weighty theoretical resource the supererogationist view of obligation turns out to be. On perfectionism, obligation is easily reducible to prior theoretical resources. On supererogationism, it comes as a surprise, with no obvious explanation of its existence, its position on the moral spectrum, or any other of its features. This is true with respect to any given ethical paradigm, and as we have now shown, this is a symptom of a more fundamental theoretical problem.

6. Objections

Multiple Best Actions

It may be objected that we can never be obligated to perform two incompatible actions. But since there are often two equally good actions, which are also better than any other possible action, at least one (and perhaps both) must be supererogatory. We respond, first, that this would not suffice for supererogation, since the ostensibly supererogatory action would not be going beyond duty – only rising in parallel with it. A supererogationist may object that, because one may

perform either of the two acts, this suffices to make the act optional and, therefore, supererogatory. However, in this case, the optionality is shallow. On the perfectionist picture, for any given situation there is a set of best actions (perfectionism need not hold that there is always only one uniquely best act). While none of these are *uniquely* obligatory, their disjunction is.[17] That is, while there is a thin sense in which I have the option of performing action A, I do not have the option of not performing one of the actions in the set; I must perform (for example) either A or B.

Degrees of Imperfection

This discussion permits us to respond to a further objection: by making the demands of morality so stringent, perfectionism is not sufficiently sensitive to the variety of suboptimal actions, dismissing them all as failures. But there is no reason to suppose that perfectionism is any less sensitive to a spectrum of good vs bad than is supererogationism. Just as supererogationists with their ternary or quaternary categorisation of actions can still say that some actions within a category are better than others (for example, killing a child is worse than beating a child, even though both are strictly impermissible), so perfectionists can equally say that some non-obligatory actions are better or worse than others. Indeed, as we saw in the discussion of theoretical simplicity, the spectral quality of morality actually fits better with perfectionism than with supererogationism.

Optimising Agents are Bad Agents

There is another worry in this vicinity, which is that perfectionism's desire for moral optimisation is too cold and calculating to properly represent our sense of agency in the ethical arena. By focusing so intensely on the best possible action to do the most good, perfectionists become resistant to ordinary, less 'efficient' means of doing good, like buying a homeless person a cup of coffee at the expense of buying potentially life-saving malaria nets for vulnerable people

---

[17] This allows us to respond to the following objection: is the omission of a best action not a failure of obligation? We propose that this omission is simply morally indeterminate, as are many other under-described actions (like the action of removing someone's appendix, which is sometimes good but in most scenarios fairly perverse).

in Nigeria. Bernard Williams (1973) raises a similar concern for *utilitarianism*, arguing that if utilitarianism is true, then the right answer to various moral dilemmas should be obvious; but the answers are not obvious, and this is evidence against the theory. The objection to *our* thesis alleges that perfectionists want to say that buying coffee for a homeless person is a failure to meet our obligations – a wrong. And this is counterintuitive to many people.

We respond in four ways: first, as noted above, that perfectionism may still be sensitive to different ways of failing our obligations – there is still the world of difference between barely subobligatory good acts and heinous crimes. Second, we note that this problem is not unique to perfectionism. For supererogationists, it will be equally true to say 'you could have done better' and, arguably, 'you should have done better' to the generous coffee-giver. It is not clear why this is any more intuitively acceptable than the claim that the coffee-giver failed his obligations. Third, as we have already discussed, insofar as this is similar to the 'we're not so bad' objection, the force of this objection is undermined considerably by the empirical evidence presented above.

Fourth, we note that if this result really is unacceptable, it is not a result of perfectionism per se. Rather, it is a function of the conjugated ethical system in play. The objection assumes that the relevant ethical system is essentially utilitarian – where morality is tightly correlated with, and largely a function of, axiology. In this sort of utilitarian system, where any dividing line between obligatory and supererogatory is likely to be arbitrary and implausible, the obligatory action will always be the one which maximises the relevant axiological parameter (e.g. pleasure). And in this case, clearly anything less will be a failure of obligations. Since there is no natural place for a division between supererogatory and obligatory actions on these utilitarian views, utilitarianism naturally commends only the best action as obligatory. But then it is clear that the real problem with this result, as Williams points out, is *utilitarianism*, not perfectionism. Perhaps the following diagnosis is correct: because perfectionism and utilitarianism are both concerned with optimisation in some sense, they are thought by many people to naturally fit together. But there is no reason why perfectionists should be particularly keen on utilitarianism vis-à-vis other normative ethical

theories. Perfectionists are free to endorse any number of ethical systems, or even a collection of first-order moral judgments not easily reducible to any simple normative formula.

In utilitarian theories, the morality of action tends to be tightly linked with consequences and states of affairs – and in particular, the morality of an action is a proportional function of the axiological value of the resultant state of affairs. The better the (intended or actual) resultant state of affairs, the better the action. Clearly most ethical systems take resultant states of affairs into account, but they are not the only determinants of the morality of an action, and it is certainly not the case that the morality of an action linearly follows the value of the consequences. There is no clear reason in our own ethics, for example, to suppose that working for an hour and donating the £10 to the Anti Malaria Foundation is always better than spending that hour tutoring a struggling student at a local school. Which is better will depend on a number of other factors, and in some or many cases neither will be better than the other. The only reason we may seem forced to choose is if we have been tricked into accepting a consequentialist paradigm that is simply alien to our own ethics, and the ethics of many others.[18] Likewise, there is no reason to think that a perfectionist version of any given ethical system entails that the more effective donation is better than the less effective tutoring. So, perfectionism is certainly not committed to the wrongness of less than maximally efficient good actions.

Supererogation and Rights

Another objection is that moral intuitions are not the only support that supererogation enjoys. Benn (2017) argues that the existence of a category of supererogatory actions is entailed by some conceptions of natural 'paired' rights: 'Anyone who accepts the existence of paired rights on any plausible account of rights therefore also accepts the possibility of supererogatory actions' (p. 279). By 'paired rights' Benn refers to morally significant acts that are also optional—that is, non-

---

[18] This is not to say that consequences are irrelevant, or that we should not at all aim for efficiency in their charitable giving.

discretionary actions that we may permissibly do or refrain from doing. An example of such a right may be the right to demand the repayment of a loan. Accepting such a category entails the acceptance of supererogation, Benn argues, because while we accept the moral significance of the action, we lack a duty to either do or refrain from doing it. Therefore, deciding to forgo your right and forgive your debtors, while exceedingly morally good, is supererogatory.

We do not deny that other features of normative landscape that one accepts may lend support to supererogation. While Benn herself does not offer a defence of such natural rights in this paper, she argues that any evidence for their existence may also be evidence for supererogation. If one accepts these sorts of natural rights, Benn describes supererogation as appearing on the scene as 'spandrels', forming naturally from the rest of the normative architecture. We are not concerned, however, that such an entailment constitutes a threat to our argument, for two reasons.

First, we presume that a common conception of natural rights is Lockean in nature and, therefore, far more general than the examples of specific 'paired rights' that Benn describes—for instance, we have a 'right' to life, but do we really have a 'right' to forgive a loan? If a law were passed forbidding loaners to forgive loans, would this law trample on one of our natural rights? It is not clear to us that the word 'right' is being used univocally in these two cases.

Second, what is the nature of such rights? For the purpose of making an argument in favour of supererogation, it cannot be as simple as 'those things we have the moral option of doing or not doing'—this would beg the question in favour of supererogationists. Any conception of rights that produce supererogation as 'spandrels' would need to be more theoretically filled-out without relying on supererogation as a theoretical assumption. While we do not want to argue that such a task would be impossible, we are unsure ourselves how such an argument would go.

Perfectionism Mischaracterises Obligations

One may think that a necessary feature of something's being a moral obligation is that they may be forced, or may force others, to fulfil it. While this at first seems like a tidy way of delineating

obligatory actions from supererogatory ones, we find this notion highly uncompelling for a few reasons. The first is that the category of things we may be forced to do, or force others to do, is highly contingent on in the particular laws of land. Under the Third Reich, certain people groups could be forced into ghettos. Because we almost certainly do not want to say that these individuals had a moral obligation to comply, it is clear that not all things we can be forced to do are obligations of ours. The ability to compel, therefore, cannot be sufficient for something's being an obligation.

A more charitable way of understanding this sort of compulsion is to say that our obligations are those things we may be *legitimately* compelled to do (which would, of course, not apply to the case of those forced into ghettos). Narveson (1999) seems to hold a similar view about obligations of justice:

> Another important question, which has cropped up in some of our discussions but is nowhere more clearly relevant than here, is the distinction between justice and charity. By justice I here intend those things that we may, if need be, be forced to do—where our actions can be constrained by others to ensure our performance. (p. 144)

Narveson's use of this distinction applies here only to obligations of justice (he still holds that there are non-enforceable duties of charity), but one may imagine expanding the scope of this idea to distinguish all supererogatory acts from all obligatory acts. Still, it is relatively easy to think of cases where this seems wrong. For example, we may think that the state can legitimately prohibit the murder of innocents; however, depending on one's normative framework, we may also think that we are obliged to disregard this legitimate prohibition and murder one innocent for the sake of saving hundreds of other innocent lives.

Additionally, we doubt that ability to compel is even a necessary feature of moral obligations. Consider the following statement: fathers who have the means and ability ought to parent their children. This is a very plausible candidate for a moral obligation. Yet, a father cannot be forced to do so (though he may be forced to pay child support). Not only can the state not compel involuntary parenting, but we may also think that nobody can legitimately compel others to fulfil relational roles, even if such fulfilling is clearly obligatory.

Failure and Blame

Our final clarification is with regards to blameworthiness. We do not aim to discuss blameworthiness in detail in this paper, for lack of space. But there is an objection that perfectionists necessarily hold people to be blameworthy for performing anything other than the best, obligatory action. Indeed, some build blameworthiness into the very definition of supererogation: an action is supererogatory if it is praiseworthy to perform, but not blameworthy not to omit (for example, Mellema 1991, p. 13). We do not accept these entailments, because we do not accept a linear relationship between the goodness of an action and the degree of blameworthiness associated with it. There may be many exculpating reasons for any given action which are related to factors other than the goodness of the action itself. Perfectionism is not committed to affirming culpability for any failure to perform the obligatory action or disjunction of actions.

7. **Conclusion**

In sum, we have clarified our concept of supererogation as follows: supererogation does not relate merely to good but morally optional actions. An extra criterion is needed: supererogatory acts must be *better* than the mere fulfilling of obligations. Along with others, we have asserted that supererogatory actions are *costlier* in any given scenario than would be the mere meeting of obligations. We have also clarified our conception of perfectionism: perfectionism does not entail a uniquely best action in all circumstances. It is compatible with – and especially consonant with – the view that non-obligatory actions can still range anywhere from utterly heinous to nearly perfect. Finally, perfectionism is consistent with a variety of ethical theories and is not wedded to cold, impersonal calculations about how to pick the best of millions of possible options. While we wish neither to undermine the importance of striving for the very best, nor to exclude the possibility that there are very few best possible actions in any given case, our theory leaves it entirely open how large the set of best possible actions is.

In this paper, we have shown belief in supererogatory actions, so conceived, to be erroneous on several grounds: the lack of natural room for them in the standard versions of the three main normative ethical systems, as well as some non-standard and competing systems; supererogationism's dependence upon intuitions about our moral status; and its theoretical complexity vis-à-vis perfectionism.

We do not deny that supererogationism is intuitively appealing, and we do not seek to claim that intuitions never offer insight on important ethical issues. What we have sought to do is show that belief in supererogatory actions relies almost solely on intuitions, and especially intuitions about our *moral standing*, and that in this case this is very shaky ground. The preference for simplicity, our moral theories themselves, and the relevant empirical data all support the thesis that obligations are far more widespread than we might have otherwise conceived. Naturally, therefore, it is also plausible that we fail them more often than we might have otherwise conceived—a notion that should keep us both sombre and vigilant.

References

Alston, W. P., 1989. 'Some Suggestions for Divine Command Theorists', *Divine Nature and Human Language*, London: Cornell University Press.

Arisotle., & Irwin, T., 1985. *Nicomachean ethics*. Indianapolis, IN: Hackett Pub.

Baron, M., 1987, 'Kantian Ethics and Supererogation', *Journal of Philosophy*, 84:237-62.

Benn, C.,
-2018a. 'Supererogation, optionality and cost'. *Philosophical Studies* 175 (10):2399-2417. -
-2018b. 'The Enemy of the Good: Supererogation and Requiring Perfection'. *Utilitas* 30 (3):333-354.
-2017. 'Supererogatory Spandrels'. *Etica and Politica / Ethics and Politics* 19 (1):269-290.
-2015. 'Over-Demandingness Objections and Supererogation'

Brinkmann, M., 2015. 'Disjunctive duties and supererogatory sets of actions'. Royal Institute of Philosophy Supplement 77:67-86.

Clark, M., 1978. 'The Meritorious and the Mandatory', Proceedings of the Aristotelian Society 79, pp. 23-33.

Crisp, R., 2013, 'Supererogation and Virtue', *Oxford Studies in Normative Ethics*, 3:13-34.

Heyd, D.,
-2015a, 'Supererogation', *Stanford Encyclopedia of Philosophy*.
-2015b, 'Can Virtue Ethics Account for Supererogation?', *Supererogation*, ed. C. Cowley, Royal Institute of Philosophy Supplement, 77: 25-47
-1982, *Supererogation*, Cambridge, Cambridge University Press.

Hill, T., 1971, 'Kant on Imperfect Duty and Supererogation', *Kant-Studien*, 61:55-76.

Hooker, B.,
-2016, 'Rule Consequentialism', *The Stanford Encyclopedia of Philosophy* (Winter 2016 Edition), Edward N. Zalta (ed.).
-2002, *Ideal Code, Real World: A Rule-Consequentialist Theory of Morality*. Oxford University Press UK.

Hoorens, V., 1995, 'Self-Favoring Biases, Self-Presentation, and the Self-Other Asymmetry in Social Comparison', *Journal of Personality*, 63:793-817.

Hursthouse, R., 1999, *On Virtue Ethics*. Oxford: Oxford University Press.

Johnson, R. and Cureton, A., 2016, 'Kant's Moral Philosophy', *Stanford Encyclopedia of Philosophy*.

Kant (trans. Wood, A.W. and Schneewind, J.B.), 2002, *Groundwork for the Metaphysics of Morals*. New Haven, CT: Yale University Press.

Kawall, J., 2009, 'Virtue Theory, Ideal Observers, and the Supererogatory', *Philosophical Studies*, 146:179-96.

McNamara, P., 1996, 'Making Room for Going Beyond the Call', *Mind* 105. 419.

Mellema, G., 1991, *Beyond the Call of Duty*: *Supererogation, Obligation, and Offence*, SUNY Press.

Merritt, A.C., et al., 2010, 'Moral Self-Licensing: When Being Good Frees Us to Be Bad', *Social and Personality Psychology Compass*, 4:344-57.

Mill, J. S., 1969, 'Auguste Comte and Positivism', in *The Collected Works of J. S. Mill* (vol. 10), Toronto: University of Toronto Press.

Miller, C., 2016, 'Is theism a simple hypothesis? The simplicity of omni-properties', *Religious Studies* 52 (1):45-61.

Moore, G.E., 1948. *Principia Ethica*, Cambridge: Cambridge University Press.

Narveson, J., 1993; 2nd edition. *Moral Matters*. Peterborough, Ontario: Broadview Press.

Parfit, D., 1984. *Reasons and Persons,* Oxford University Press.

Scanlon, T., 1998, *What We Owe to Each Other*. Cambridge, MA: Belknap Press of Harvard University Press.

Smart, J. J. C. & Williams, B., 1973. *Utilitarianism For and Against*. Cambridge University Press

Stangl, R., 2016, 'Neo-Aristotelian Supererogation', *Ethics*, 126:339-65.

Street, S., 2006, 'A Darwinist Dilemma for Realist Theories of Value', *Philosophical Studies*, 127:109-66.

Swanton, C., 2001, 'A Virtue Ethical Account of Right Action', *Midwest Studies in Philosophy,* 112:32-52

Swinburne, R., 2004, *The Existence of God*. Oxford: Oxford University Press.

Timmermann, J., 2005, 'Good but Not Required?—Assessing the Demands of Kantian Ethics', *Journal of Moral Philosophy*, 2:9-27.

Urmson, J. O., (1958). 'Saints and heroes', In A. I. Melden (ed.), *Essays in Moral Philosophy*. University of Washington Press.

Wolf, S., 1982. 'Moral Saints'. *Journal of Philosophy* 79 (8):419-439.

Wynn, M. (2015), 'Supererogation and the relationship between religious and secular ethics: some perspectives drawn from Thomas Aquinas and John of the Cross.' *Royal Institute of Philosophy Supplement* 77:163-183.

Acknowledgments