

Susan Wolf on Supererogation and the Dark Side of Morality

Nora Grigore*

Received: 22 December 2020 / Revised: 11 July 2021 / Accepted: 1 December 2021

Abstract: Wolf proposes supererogation as a solution for curbing the exaggerated demands of morality. Adopting supererogation is supposed to prevent us from considering that all morally good deeds are obligatory. Supererogation, indeed, makes some morally good deeds merely optional, saving the agent from the requirement of behaving as much as possible like a Moral Saint. But I argue that Wolf cannot use supererogation in service of her overall project, for two reasons. First, because implied in the concept of supererogation is that going beyond duty adds to our humanity rather than detract from it (as Wolf argues). Secondly, after analyzing attempts to acclimatize supererogation in other theoretical frameworks, I conclude that supererogation can limit morality’s claims only if Wolf’s reasons of “individual perfection” can defeat moral reasons. I argue that a common scale of evaluation between moral and non-moral reasons is needed for their comparison, but Wolf explicitly rejects this way out.

Keywords: Morality; moral saints; supererogation.

1. Introduction

In “Moral Saints” Susan Wolf famously depicts a rather bleak image of the moral saint. A moral saint, claims Wolf, is someone “whose every action

*  <https://orcid.org/0000-0002-4035-5282>

 noragrigure@utexas.edu



is as morally good as possible, a person, that is, who is as morally worthy as can be" (Wolf 1982, 419). At the same time, she claims, there is a dark side to this moral excellence, mainly because dedicating so much time and resources to morality would have catastrophic consequences for the personal, private side of the moral agent.

One may protest the idea that living one's life to the moral extreme has such bleak consequences: the saint does not have to be this eviscerated self Wolf describes. Adams (1984), for instance, claims that if some suppositions from Wolf's picture are removed (e.g. the assumption that a moral saint should always maximize the moral good), then the resulting image is quite different. However, this is not the path I'll take. Rather, I'll show that, given Wolf's dark view of moral demands, her solution for restricting them (i.e. supererogation) doesn't work.

According to Wolf, following moral ideals has catastrophic effects for our personal lives. She therefore wishes to limit the influence of moral recommendation in favor of a personal, individual point of view, that she calls "the point of view of individual perfection." Wolf thinks that the best instrument for moral theories to make this limitation is supererogation.¹ Actions are usually called "supererogatory" when they are considered to be morally excellent, but not obligatory (paradigmatic examples being saintly or heroic deeds). The idea is that supererogation establishes a threshold (because supererogation is thought as going *beyond* duty) for what is obligatory; any moral behavior above it is merely permitted even if morally laudable (e.g., saintly and heroic action). I will focus on the relation between the image of a very demanding kind of morality and supererogation. And

¹ The usual general characterization of supererogation goes along these lines: "Supererogatory acts are those which lie 'above and beyond the call of duty'. Such acts characteristically enjoy a very high degree of value, probably more value than any other act available to the agent. (...) actions which is not wrong of the agent not to do (Dancy 1998, 173). Recent books covering the topic of supererogation start by this rough characterization: "It is often said that works of supererogation involve going beyond the call of duty, doing good in a way that transcends the requirement of moral obligation" (Mellema 1991, 3) or "Supererogation is the technical term for the class of actions that go 'beyond the call of duty'. Roughly speaking, supererogatory acts are morally good although not (strictly) required" (Heyd 2002).

I'll argue that supererogation cannot be accommodated in the framework adopted by Wolf, in which morality is seen as overly demanding, and two separate, independent scales for moral and non-moral values exist.

What I hope to show is that the failure to accommodate supererogation is not of merely local interest, a glitch in the bigger picture painted by Wolf about life and morality. There is a certain conception about morality as very demanding, (one might say "life-denying") that is so prevalent, so natural and by-default-adopted, that even authors who are not sympathetic towards it, who are critical and opposed, still espouse some of its basic assumptions. I think Susan Wolf is such an author and therefore a very relevant illustration of this pervasive image of morality. In short, Susan Wolf is famous for saying that moral saints cannot but have a diminished quality of life, and a diminished humanity. In the fight between 'life' and morality, she is no doubt, on the side of life. However, my point is that she shares this presupposition with her adversaries, that normal human life and the "higher flights of morality" are engaged in a struggle or, at least, in zero-sum game. She shares this presupposition with many other illustrious philosophical names. After all, it is a quite a Nietzschean claim that morality is "life-denying" (to be exact, "morality of slaves" has this role). Korsgaard (1996), quoting Nietzsche, considers that the proper role of morality is that of imposing or "forcing" values upon nature and life, thereby restricting and shaping their course. My present point is that this is a rich philosophical tradition, and one in which supererogation (with its knack for lifting some moral obligations) has never been at ease. Nevertheless, Wolf seems to want to have both: a conception of demanding, obligations-imposing morality and supererogation. My aim is to explain why this is an impossible philosophical mission.

In short, my aim is to argue for the following conditional claim: if Wolf wishes to keep (what she takes to be) the commonsensical image of morality, then it will be very difficult to *also* maintain the theoretical solution she proposes, namely supererogation. The conception of morality Wolf wishes to keep isn't welcoming towards the concept of supererogation. In that sense, Wolf doesn't ultimately "respond to the unattractiveness of the moral ideals that contemporary theories yield" (Wolf 1982, 434).

2. The dark side of moral saintliness

Wolf's main point is that we spontaneously judge moral ideals to be unpleasant and damaging *from another point of view than the moral one*, namely from what she calls "the point of view of individual perfection."² In other words, she thinks that if one follows the recommendations of commonsense morality (promoting at all times other people's good and disregarding one's own interest) and moves asymptotically towards a moral ideal, then one will end up hurting some nonmoral, personal values. One will end up, that is, with a mutilated life in one respect or another. Susan Wolf acknowledges that commonsense morality doesn't make the saintly moral ideal into an *obligatory* path to take. But she thinks that, even if not obligatory, this path is recommended as *the best* path one could take. She objects to that, saying that this can be seen as the best *moral* path and, at the same time, as a *bad* choice for the agent and those close to him in many other important respects (e.g. from the point of view of one's personal life). Her recommendation for solving this tension between the moral point of view and the point of view of individual perfection is to somehow restrict the claims that morality places on us and to give personal ideals legitimacy in our evaluations:

If we are not to respond to the unattractiveness of the moral ideals that contemporary theories yield either by offering alternative theories with more palatable ideals or by understanding these theories in such a way as to prevent them from yielding ideals at all, how, then, are we to respond. Simply, I think, by admitting that moral ideals do not, and need not, make the best personal ideals. (Wolf 1982, 434–35)

I want to underline that Wolf mentions here *two other possible ways* of preventing the undesirable consequence of self-mutilating moral sainthood: a) modifying our views about morality (such that its demands are no longer at odds with goals of personal perfection); or b) making the pursuit of an ideal

² Wolf (1982, 427): "Let us call the point of view from which we consider what kinds of lives are good lives, and what kinds of persons it would be good for ourselves and others to be, the point of view of individual perfection."

something of an undertaking outside morality. In other words, one might think that our conception of morality should change in such a way that moral excellence would not involve a mutilated self; or, alternatively, that our conception of morality would include only rules about what is obligatory and forbidden, and would place aspirations or ideals outside morality.

She gives counterarguments to taking route a), which implies that she wants to keep both the claims of what it means to be (commonsensically) moral and the moral ideals that they engender. However, she also wants some limits placed on moral claims when they go against personal goals, and she mentions *supererogation* as a helpful theoretical instrument. Supererogation may seem helpful because it establishes limits for moral obligation, a kind of threshold above which the agent isn't morally obligated to act. Therefore, heroic or saintly actions would not be morally required, and their omission could not be regarded as a fault (it can be seen as an appropriate instrument for pushing back against the threat of moral claims). I will argue, however, that even though supererogation is usually regarded as an instrument for pushing back against maximization requirements (for instance), her way of seeing morality is an unwelcoming theoretical environment for the concept of supererogation.

3. Morality's demandingness

In other words, morality should not be the only or even the most important set of values guiding our actions: "morality itself should not serve as a comprehensive guide to conduct" (Wolf 1982, 434). Once we admit that there are certain nonmoral values that one is right to attend to, then the claims of morality to be the most important guide to action are limited, mitigated so as to not lead to the extremes exhibited in the image of the moral saint.

Following her recommendation that one should not take moral evaluation and action too far, Wolf also recommends amending moral theories in order to fit this limitation of powers regarding moral claims. Namely, she recommends that moral theories should use *supererogation* as a helpful instrument:

From the moral point of view, we have reasons to want people to live lives that seem good from outside that point of view. If, as I have argued, this means that we have reason to want people to live lives that are not morally perfect, then any plausible moral theory must make use of some conception of supererogation. (Wolf 1982, 438)

Supererogation seems like a good choice in this respect because it implies that saintly and heroic deeds cannot and should not be deemed obligatory. However, I will claim that other aspects of supererogation come into conflict with some parts of Wolf's story, and specifically with her way of seeing morality.

Wolf claims (rightly, I think) that commonsense morality is heavily other-oriented: according to its recommendations one is supposed to help others on each and every occasion, regardless of the sacrifice imposed on the agent.

Notice that these two features—the expectation to help others, and the agent's sacrifice not counting as a valid moral consideration against giving that help—jointly constitute what authors usually call the *demandingness* of morality, i.e. the heavy burden placed on the moral agent to act sometimes against her own interest for the general good. In Wolf's case, she sees demandingness as unjustifiably affecting the agent's private life. Her solution *isn't to change our conception of morality but to mitigate some of its effects on us*. This move, I will argue in the next section, doesn't work because once one admits that morality is overly demanding, the effects of conceiving morality as a harshly demanding enterprise are difficult to avoid.

4. What is problematic in Wolf's solution

Supererogation seems to be a good solution and the right theoretical tool for Wolf. She wants to be able to say that in some (but not all) circumstances the agent may ignore morality's recommendations in order to attend to personal perfection. Supererogation, as usually presented, grants this permission even though it doesn't specify exactly why or what kind of reasons the agent is allowed to follow instead. The concept of supererogation only allows that there are circumstances when we are able to rightly ignore the morally right thing to do, without punishment, blame or justified reproach

(even if, subjectively, one might feel regret). It seems, therefore, that the concept of supererogation can provide what Wolf is looking for: some limitation of morality's grip on our life and values, making room for the point of view of individual perfection.

I agree that supererogation is an instrument able to provide all of these things. But its task is made difficult (if not impossible) by the particular conception of morality that Wolf keeps (even if she pushes against its perceived excesses). Supererogation can be deployed, as a theoretical instrument, against favorable or unfavorable theoretical backgrounds: some moral theories may make less room or no room at all for supererogation, depending on various other factors. For example, an obviously unfavorable environment for supererogation is one in which morality presupposes maximization of the good, as in act-utilitarian or act-consequentialist theories. If one is required to maximize the overall good on each occasion, then there cannot be acts that are good but not required, or good acts that can be omitted without blame; so no supererogatory actions exist. Basically, any theory claiming that what is morally good has to be covered by some kind of obligation or has to stay under an "ought," will threaten supererogation.

My first claim is that Susan Wolf has a specific conception on which morality is—and should be—demanding and imperative.³ My second claim is that it is very difficult to integrate supererogation into such a way of seeing morality.

One might think that my objection to Susan Wolf misses the point because the imperative character of morality was the whole purpose of employing supererogation. Precisely *because* morality is imperative and demanding (threatening to invade the personal domain, as Wolf sees it) we need something to curb its claims; had morality not been so demanding, then there would have been no need for supererogation in the first place. Supererogation is presumed to bring much needed permissions for the agent in the austere environment dominated by moral imperatives. My response is that not all theoretical landscapes can be balanced just because one needs some balance in them. Sometimes, the theoretical devices embedded in the theory simply exclude the possibility of supererogation. (E.g., the

³ Wolf claims she is tapping into the commonsense view of morality.

maximization required by some utilitarian theories can promptly exclude supererogation even if one might think supererogation is needed in those theories in order to make them more intuitively plausible.) In other words, I agree that morality being perceived as demanding is the circumstance where one is more likely to need the theoretical help provided by supererogation. However, this doesn't mean that it can always be *successfully* deployed and integrated within a certain theory. I think Wolf's theory is one of the unsuccessful cases. Let me explain why.

Susan Wolf clearly states that, according to commonsense morality, saintliness or ideal behavior isn't obligatory. However, both common sense and her own account of morality tend to go against this thesis. This is important because a morality that admits that some morally good things are not obligatory doesn't threaten supererogation. On the other hand, a morality that at least tends towards making all morally good things obligatory (what I have been calling an *imperative* kind of morality) is usually a threat for supererogation. When Wolf says that moral ideals are not obligatory, she seems to regard commonsense morality as being the former kind of morality. However, when she actually pushes against the claims of morality as she understands it, that morality seems closer to the latter kind. I am going to argue that both her account of morality and the commonsensical one have aspects similar to the imperative way of seeing morality.

First, morality according to common sense is far from being a coherent set of beliefs. It is true that, usually, we do not regard moral ideals as obligatory, and it is also true that supererogation (going beyond duty) is a commonsensical notion. So it would seem that there are some obligatory and some non-obligatory (e.g. supererogatory, saint-like) types of actions according to common sense. However, it is also a commonsense belief that anyone who can help, *should* help others in a difficult situation. If this "should" is translated into moral obligation, then each time one helps, one is simply fulfilling a moral obligation—something that is not, morally speaking, optional. I think this is the basic intuition behind the good-ought tie-up, namely that *one has a moral obligation to help others* in need because this is what constitutes the moral good—one might say that this is what morality is all about. The corollary of this thought is that one cannot invoke the inconveniences, or the losses suffered by oneself in order to opt out of

moral obligation: this is how morality works, by foregoing one's own interests to altruistically care about other people.

Consequently, there is a tension here and we can see that commonsense morality may indeed seem, at times, very demanding. If one interprets it as saying that every morally good deed *should* be done, one can also interpret this as requiring an ascent to moral ideals. Wolf herself mentions in passing something resembling a tension in the commonsense view, but she attributes it to different contexts and doesn't elaborate on the relation between ideals and contexts. She says that "outside the context of moral discussion" we consider it natural to reject the model of the moral saint, because we agree that we aren't blamable if we don't always act following the highest moral recommendation). But in the context of moral discussion, however, we also want to claim that "one ought to be as morally good as possible" and it would be at least shameful not to aim at that:

In other words, I believe that moral perfection, in the sense of moral saintliness, does not constitute a model of personal well-being toward which it would be particularly rational or good or desirable for a human being to strive. Outside the context of moral discussion, this will strike many as an obvious point. But, within that context, the point, if it be granted, will be granted with some discomfort. For within that context it is generally assumed that one ought to be as morally good as possible and that what limits there are to morality's hold on us are set by features of human nature of which we ought not to be proud. (Wolf 1982, 419)

I think this is a quite clear expression of the tension I want to point at: on the one hand it is obvious that we are not required to be saints; on the other hand, it seems equally obvious that we *do* want to say that we are required to pursue what is morally best. Of course, one can choose one or the other, by accepting the moral obligation implied in the requirement to always do your morally best, or by limiting it. My point is simply this: that *Wolf seems to want to embrace both at the same time*. This is because she wants to keep the very demanding version of morality that makes every good deed required of us and, at the same time, to have supererogation limit these requirements. Or, in other words: to accept that we always have to aim at doing our morally

best and that we do not always have to. In this respect, Susan Wolf's position is in keeping with commonsense morality by taking in its inner tensions. But if her take is ambivalent in this respect, then one part of her image of morality (the one describing morality as demanding and imperative) is the one that makes supererogation difficult to accommodate.

My second point about the tension between requiring and not requiring that all morally good deeds be performed, regards the way Susan Wolf herself chooses to depict morality. Especially when she argues against morality's demanding ideals and in favor of limitations being imposed (in order to make room for the legitimacy of the point of view of individual perfection), Wolf is presenting a morality that has a strong, imperative character, one that overrides and demotes other concerns:

[T]he desire to be as morally good as possible is apt to have the character not just of a stronger, but of a higher desire, which does not merely successfully compete with one's other desires but which rather subsumes or demotes them. The sacrifice of other interests for the interest in morality, then, will have the character, *not of a choice, but of an imperative*. (Wolf 1982, 423–24, my italics)

This image, of an imperative morality, should not come as a surprise if we consider one other aspect of the problem, namely that Wolf sees morality as engaged in a zero-sum game with the domain of the private, personal life of the agent. This point is made clear by the disturbing picture of the moral saint: whatever one does for the moral good of others is a loss for the personal self; and, conversely, whatever one does good from the point of view of individual perfection is a rejection of the relentless demands of morality.⁴ A consequence of this way of thinking is that morality is seen as a difficult, demanding, and unpleasant (to say the least) to follow. Such a view is (not necessarily, but likely) going to have to rely on obligation in order to see its

⁴ “The normal person's direct and specific desires for objects, activities, and events that conflict with the attainment of moral perfection are not simply sacrificed but removed, suppressed, or subsumed. The way in which morality, unlike other possible goals, is apt to dominate is particularly disturbing, for it seems to require either the lack or the denial of the existence of an identifiable, personal self” (Wolf 1982, 424).

tasks fulfilled and its recommendations followed, because it is unlikely that people would want to undertake such unpleasant tasks voluntarily (especially since sacrifice is often involved). And this shows, against our moral intuition that not all moral deeds should be demanded, that certain moral frameworks end up with extending (at least some form of) obligation to the whole domain of morality.

There is, I think, a third argument for my tenet that Susan Wolf inclines towards an imperative view of morality, even if she doesn't say it explicitly. She says so herself, in her passionate plea for the personal domain: morality *should not* play the role of supreme scale of values, and the agent should not ask permission for omitting to always do the morally best thing. But this protest means that she believes that moral value *is* the value that trumps any other kind of value (moral values "subsume or demote" other values⁵), and the agent might feel that she has to ask permission in order not to do the morally best thing. It is not only when she *opposes* these tendencies that she recognizes them (the tendencies of *requiring* any morally good deed, of following any moral good with an "ought"). It is also when she approvingly characterizes morality that she says the following: "A moral theory that does not contain *the seeds of an all-consuming ideal* of moral sainthood thus seems to place false and unnatural limits on our opportunity to do moral good" (Wolf 1982, 433, my italics).

Therefore, I think one should not be surprised if Susan Wolf ends up with an imperative version of morality, one where obligation plays a central role and usually does not accommodate supererogation. However, it can be argued that there are examples of such moral theories that have tried to accommodate supererogation, and what Wolf ends up with is not a straightforward contradiction, but rather a puzzling tension. I'll now further pursue the charitable assumption that Wolf's view could be one of them. The main task of the next section is to present various theoretical possibilities and evaluate them to see if they could be a path Wolf could take.

⁵ Wolf (1982, 424).

5. How supererogation may be accommodated

In a way, what Susan Wolf recommends is quite banal and, moreover, it is something that we already do routinely: we limit moral demands when they threaten other parts of our lives. The problem for a moral theory, however, is to find a justification for this limitation in its own terms. The problem is: can it be morally justified to limit morality's demands? And how exactly will that justification look for a particular theory?

When speaking specifically about supererogation, the problem is already famous: the puzzle, paradox,⁶ or simply the problem of supererogation—they all refer to a number of difficulties for various theories in justifying the agent's permission to sometimes omit the morally best action. For Susan Wolf, in particular, the problem of supererogation is the following: *How can one justify that we are sometimes allowed not to follow moral prescriptions, and instead act for the good from the point of view of individual perfection?* It could seem that she offers an answer when she worries about the possible objection that not pursuing moral ideals in order to attend to individual perfection is just an excuse for pursuing a selfish agenda. Differently put: how do we know that when we reject morality's claims on us, we do this for the right reasons? She replies that there are nonmoral values and nonmoral virtues involved in individual perfection. These give rise to valid nonmoral reasons to sometimes reject the claims of moral reasons.

In other words, some of the qualities the moral saint necessarily lacks are virtues, albeit nonmoral virtues, in the unsaintly characters who have them. In advocating the development of these varieties of excellence, we advocate nonmoral reasons for acting, and in thinking that it is good for a person to strive for an ideal that gives a substantial role to the interests and values that correspond to these virtues, we implicitly acknowledge the goodness of ideals incompatible with that of the moral saint. (Wolf 1982, 426)

⁶ Cf. Archer and Ridge (2015) and Horgan and Timmons (2010).

On the image Wolf offers here, there are moral reasons in favor of pursuing supererogatory acts (or saintly acts), and they are sometimes opposed by nonmoral reasons (belonging to individual perfection) that sometimes win the confrontation between reasons.

There are two immediate problems with this response that I can discern. First, leaving aside the problems raised by confronting moral with nonmoral reasons, is the missing common scale of comparison. The moral and the individual point of view are independent points of view, according to Wolf, without an overarching framework to encompass them both. It is true that the moral point of view gives some weight to the individual point of view and the other way around. But when they are in conflict, there are no means to decide which one will prevail. Wolf explicitly rejects the possible construction of a common framework out of fear that it will become one that will again make moral value the ruling, deciding value:

The philosophical temperament will naturally incline, at this point, toward asking, “What, then, is at the top—or, if there is no top, how are we to decide when and how much to be moral?” In other words, there is a temptation to seek a metamoral—though not, in the standard sense, metaethical—theory that will give us principles, or, at least, informal directives on the basis of which we can develop and evaluate more comprehensive personal ideals... I am pessimistic, however, about the chances of such a theory to yield substantial and satisfying results. For I do not see how a metamoral theory could be constructed which would not be subject to considerations parallel to those which seem inherently to limit the appropriateness of regarding moral theories as ultimate comprehensive guides for action. (Wolf 1982, 438–39)

This is only a general observation regarding her solution. But, more to the point, if Wolf wants to employ supererogation in her theory, this comes with some additional complications. When trying to justify why an agent is allowed to omit some morally excellent actions, the justification cannot be merely prudential—it has to carry moral weight. It is obvious why, for *prudential* reasons, the agent can omit heroic or saintly deeds: they involve heavy self-sacrifice. What is difficult to do, and what the problem of supererogation asks, is what *moral* reasons one could have not to act saintly or heroically.

And this is a justified demand if we consider that, on the commonsense notion of supererogation, it is not only disadvantageous to place the agent under an obligation to act heroically, but it is first of all *morally wrong*—we feel—to make sacrifices of this kind *obligatory* (Urmson 1958). The obligation itself seems in these cases immoral, for in most cases something is wrong with being constrained to give your life or limb for the greater good. The problem is—to give a theoretical support for this impression that “something is wrong.” So there must be some *moral* reasons justifying the fact that we are *not obligated but merely permitted* to act in a saintly or heroic manner.

Of course, Wolf’s reasons for foregoing saintly actions are explicitly *non-moral*, personal and partial to the agent. She insists that they are not exactly prudential reasons, as they do not have the agent’s interest in view, but they are something else, namely concerned with the agent’s individual perfection. Therefore she cannot provide, in her theory’s own terms, a *moral* justification for sometimes disobeying morality and following one’s own plans.⁷ Consequently, she doesn’t have a good answer to the problem of supererogation, even if she says that the claims of the personal, individual point of view are recognized by morality to some extent. The problem for her theory is that these claims, when seen from within the moral point of view, don’t carry much weight according to her own evaluation. Therefore, they cannot be reliable in “defeating” moral reasons that would recommend heavy sacrifices on the part of the agent.

6. Morality and supererogation

To recap, Wolf’s image about morality is that the moral domain is at odds with the personal domain, that it is other-oriented and has an imperative character. Wolf wants to keep these features, as she believes that this is what morality should look like, but at the same time she wants to restrict moral claims such that one would not be under an obligation to give up personal plans in order to attend to helping others. For this task she proposes that moral theories make use of supererogation.

⁷ Jonathan Dancy raises a similar objection to her theory in *Moral Reasons*.

I have argued, first, that supererogation is difficult to accommodate within moral theories that are obligation-based and have an imperative character because these theories usually tend to assume that all morally good deeds stay under an “ought” (and therefore come into conflict with the idea that some excellent moral deeds are merely permitted—as supererogatory). In Section 3, I have argued that both commonsense morality and Wolf’s own position fall into this category of obligation-based morality, despite the fact that they do contain some opposing intuitions in this respect. However, because supererogation can be accommodated even in unfriendly environments by making appropriate theoretical adjustments, I have looked into the possibility of making such adjustments in Section 4. In order to be able to use supererogation, Wolf’s theory should be amenable to a credible strategy for morally justifying the omission of saintly or heroic actions. The justification should explain why or how, sometimes, nonmoral reasons from the personal side are able to defeat moral reasons. I’ve argued that, because Wolf doesn’t have a common scale for moral and nonmoral reasons, she has no way of explaining how such a confrontation may be decided.

In the end, the issue of being able to use supererogation within Wolf’s framework is this: Supererogation, as a conceptual structure, has two main components,⁸ namely that some excellent moral deeds are not (and should not) be obligatory, and that the same moral deeds are praiseworthy or good from a moral point of view. I have argued that neither of these components fits with what Wolf wants to say.

The first component tends to be undermined directly by the image of an imperative, obligation-based morality because this kind of morality tends to make all morally good deeds obligatory. Even if and when limits are imposed to obligation in these theories, they look like concessions made to human weakness or to everyday intuitions⁹ rather than limits imposed by morality itself. For the state where *all morally good deeds are obligatory* is regarded as the *default rational* state from which one can depart by various technical means or by making concessions to the fact that humans cannot

⁸ Dreier (2004) and Hurka and Schubert (2012).

⁹ Slote (1984) and Scheffler (1994) argue for integrating supererogation in act-consequentialist frames in order to accommodate commonsensical intuitions about morality.

live up to this high (and very demanding) moral standard. For example, Thomas Nagel claims that allowing the omission of morally good acts by invoking supererogation is a compromise due to “human weakness,” compromise in which “[w]e must so to speak strike a bargain between our higher and lower selves in arriving at an acceptable morality.”¹⁰ Wolf vigorously protests this position; for her, our omission of morally saintly deeds should need no permission and no excuse:

It is misleading to insist that one is permitted to live a life in which the goals, relationships, activities, and interests that one pursues are not maximally morally good. For our lives are not so comprehensively subject to the requirement that we apply for permission, and our nonmoral reasons for the goals we set ourselves are not excuses, but may rather be positive, good reasons which do not exist despite any reasons that might threaten to outweigh them. (Wolf 1982, 436)

So, on Wolf’s view, supererogation cannot come as an excuse for moral weakness. And, indeed, once one admits that supererogatory actions exist, one should also admit that some moral deeds can be omitted without an apology needed from the agent. But then Wolf’s theory should afford means (even if not explicitly given) to limit morality’s claims by using *moral* reasons, or at least reasons that can be given significant moral weight. If not, her nonmoral personal reasons, that are supposed to justify our omission to behave saintly, can look like just another *prudential* reason to not risk too much in the service of morality. To show how nonmoral personal reasons can defeat moral ones, Wolf would need something like a common scale of values, i.e., she would need some theoretical device that would allow for a comparison between the two kinds of reasons. And she explicitly rejects this possibility.

The second component of supererogation, the one claiming that actions that go beyond duty are morally excellent, praiseworthy actions, doesn’t fare much better than the first component, since Wolf famously claims that saints are repulsive, defective human beings. Recall this is because Wolf thinks that the more one improves morally, the narrower one’s mental horizons become (due to an obsessive concern with helping others), the less time one has for

¹⁰ Nagel (1986, 202).

oneself and the more unpleasant and lacking in humanity they become. This is also a feature of her view about morality as being an extremely demanding enterprise, one engaged in a competition with the personal domain, such that each time one acts morally the personal domain loses, and the other way around. For Wolf, there is definitely such thing as “too much morality.”¹¹ It isn’t clear if this is a case of “too much of a good thing” or a case of “something better to have only in moderate quantities.” Considering her tone, I would venture to say that it is the latter.

Regardless of how one chooses to interpret her position in this respect, Wolf’s view regarding moral saints paints a very different picture from the one promoted by supererogation regarding saintly and heroic deeds. For, in the case of supererogation, saintly actions are presented as praiseworthy and overall good. Wolf’s reply to the objection that her image of the saint is not very appealing is that a saint is morally excellent, but a rather unpleasant figure from another point of view, that of personal perfection. However, I believe this doesn’t address the discrepancy between her image of the saint and the commonsense one regarding supererogatory action. Our tales of heroic and saintly deeds are not cautionary tales about how the hero was a morally excellent person, but nevertheless they ruined their humanity out of lack of moderation regarding the moral good and, therefore, one should be careful not to do the same. Quite the contrary. The main character from *Schindler’s List* who, at the end of the story, is tortured by remorse thinking that he could have done more, earn more money and buy more lives—is hardly someone to whom we could reasonably recommend moderation because he could be seen as a glutton for morality, having lost part of his humanity in the process. Therefore, I think that one cannot employ the usual concept of supererogation while at the same time denying that agents who act supererogatorily are morally excellent *and better human beings overall*. This part of supererogation, claiming that agents who go beyond duty are not only partially admirable but also overall better human beings, will always be in conflict with the image of the moral saint depicted by Wolf.

I believe it is important to see that Susan Wolf’s way of seeing morality is part of a venerable tradition, and more importantly, part of a tacitly held

¹¹ Wolf (1982, 483): “In other words, there seems to be a limit to how much morality we can stand.”

opinion that morality is and should be demanding, harsh, life-denying. Once we see this as a philosophical presupposition (and not as a “fact of moral life” as Nagel (1986), for example would claim), we may begin to wonder if another view of morality is possible, and if it would be a better fit for the concept of supererogation.

References

- Adams, Robert Merrihew. 1984. “Saints.” *The Journal of Philosophy* 81: 392–401. <https://doi.org/10.2307/2026294>
- Archer, Alfred, and Michael Ridge. 2015. “The Heroism Paradox: Another Paradox of Supererogation.” *Philosophical Studies* 172: 1575–92. <http://www.jstor.org/stable/24704233>
- Dancy, Jonathan. 1993. *Moral Reasons*. Oxford: Blackwell.
- Dancy, Jonathan. 2004. “Enticing Reasons.” In *Reason and Value: Themes from the Moral Philosophy of Joseph Raz*, edited by Wallace, Jay R., Philip Pettit, Samuel Scheffler and Michael Smith, 91–118. Oxford: Clarendon Press.
- Dreier, James. 2004. “Why Ethical Satisficing Makes Sense and Rational Satisficing Doesn’t.” In *Satisficing and Maximizing*, edited by Byron, Michael, 131–54. Cambridge: Cambridge University Press. https://doi.org/10.1111/j.1468-0149.2007.00449_9.x
- Heyd, David. 1982. *Supererogation: Its Status in Ethical Theory*. Cambridge: Cambridge University Press.
- Horgan, Terry, and Mark Timmons. 2010. “Untying a Knot from the Inside Out: Reflections on the ‘Paradox’ of Supererogation.” *Social Philosophy and Policy* 27: 29–63. DOI:10.1017/S026505250999015X
- Hurka, Thomas, Esther Schubert. 2012. “Permissions to Do Less Than the Best: A Moving Band.” *Oxford Studies in Normative Ethics* 2: 1–27. DOI:10.1093/acprof:oso/9780199662951.003.0001
- Korsgaard, Christine. 1996. *The Sources of Normativity*. Cambridge: Cambridge University Press.
- Nagel, Thomas. 1986. *The View from Nowhere*. New York & Oxford: Oxford University Press.
- Scheffler, Samuel. 1994. *The Rejection of Consequentialism: A Philosophical Investigation of the Considerations Underlying Rival Moral Conceptions*. Oxford: Oxford University Press.
- Slote, Michael. 1984. “Satisficing Consequentialism – I.” *Proceedings of the Aristotelian Society* 58: 139–63. <https://doi.org/10.2307/2940825>

Urmson, James Opie. 1958. "Saints and Heroes." In *Essays in Moral Philosophy*, edited by Melden, Abraham I., 198–216. Seattle: University of Washington Press.

Wolf, Susan. 1982. "Moral Saints." *Journal of Philosophy* 79: 419–39.

<https://doi.org/10.2307/2026228>