UTILITARIANISM, CONSEQUENTIALISM AND MAKING ROOM FOR SUPEREROGATION

NORA GRIGORE

Abstract. I am discussing three strategies of fitting supererogation within consequentialist frameworks, namely Slote's (1984) and Scheffler's (1994). My main claim is that not only the utilitarian or consequentialist framework is modified to accommodate supererogation, but also the concept of supererogation suffers transformations in the process. It is therefore questionable if the theories discussed manage indeed to make room for the commonsensical moral intuitions carried by the concept of supererogation.

Keywords: *Ethical Theories, Supererogation, Consequentialism, Optimization, Satisficing.*

Ethical theories, of both utilitarian and deontological persuasion, tend to have difficulties in accommodating the commonsensical intuition that some morally good deeds cannot be required, *i.e.* in accommodating supererogation. As Heyd (1982) notes, in their "pure but crude forms", the two kinds of ethical theories have difficulties in accommodating supererogation for different reasons: deontological theories because they tend to assume that the domain of the morally good is exhausted by duty (in its various forms); utilitarian theories have difficulties because they tend to require uncompromised maximization of the good, not leaving any space for extraordinary good deeds that should not be required. This does not mean, of course, that more refined versions of the two cannot attempt to accommodate supererogation (the degree of their success is another matter).

In focus here are act-consequentialist[1] theories and their relation to a commonsensical concept of supererogation. I hope to show that, in trying to accommodate supererogation, not only these theories suffer modifications but also the concept of supererogation emerges as different from the commonsensical one.

To 'accommodate' supererogation means that some plausible explanation has to be given within the theory for the commonsensical intuition that morally excellent deeds cannot be required. In order to accommodate this basic intuition about supererogation, utilitarians and consequentialists will usually change not only the theoretical setting of their theory, but also the meaning or the sphere of the concept of supererogation.

I. Supererogation under Utilitarian lenses

The commonsensical notion of supererogation has, I claim, deceptive clarity and simplicity. This might be because it constitutes only a broad outline of a possible problem: the details are to be filled in by anyone trying to figure out an explanation for the tenets of supererogation. Hurka and Schubert (2012), for example[2], paint this broad outline in the following manner:

> The concept of supererogation has two sides. On one side, a supererogatory act isn't morally required; on the other side, it's somehow better than its alternative, or "beyond" duty in a sense that connotes superiority. (Hurka and Schubert, 2012:8)

[1] I will call them "consequentialist" from now on for brevity of expression.

[2] A similar one: "Supererogatory acts have at least two essential features: i) they must be morally optional in the sense of being neither obligatory nor forbidden and ii) they must be in some sense morally superior to some other act that the agent may permissibly do instead. Differing accounts of supererogation typically vary in terms of what they take to be the relevant sense of moral superiority. "(Portmore, 2003: 326)

This two-parts view of supererogation seems straightforward. And yet a multitude of questions are left unanswered: What makes the supererogatory deed superior? Is it superior because of the sacrifice of the agent? Could one imagine deeds that go beyond duty with little or no sacrifice? There are many other features that can be added to the simple two-parts view of supererogation in order to make possible a full theoretical explanation. The features that appear in consequentialist discussions of supererogation sometimes fill in the general picture of supererogation, and sometimes transform it; for while sometimes consequentialist and utilitarian theories adapt themselves to accommodate supererogation, sometimes the concept of supererogation is adapted to fit an utilitarian frame.

To see this process of reciprocal influence unfolding, one needs to look at the outline of a consequentialist/ utilitarian theory and notice which basic traits of these theories come into conflict with which basic traits of the commonsensical concept of supererogation.

In order to chart the differences, the conflicts and the transformations that follow the reconciliation attempts, I will identify the main fault-lines of this conflict, flagging the most promising attempts at resolution.

I.1. The impersonal scale measuring the good

First, one might notice that in characterizing the two sides of supererogation (being neither obligatory nor forbidden and being morally better than alternatives) the second part is left vague; we are told that the supererogatory action is "better in some sense" but one is left to fill in the details as one chooses. This is exactly where consequentialist and utilitarian theories supply promptly detailed theoretical constructions, so that "better" tends to acquire a quite technical sense. A favored, well-known technical device is to appeal to an impersonal ranking[3] of states of affairs as a scale in determining the rightness of a moral deed.

[3] Act-consequentialism is generally characterized as a certain sort of view about the relation between an act's rightness and its consequences.

I.2. Required optimization or maximization

The commonsensical understanding of supererogation lacks not only explanations about what "better" moral deeds might mean, but also the idea of an optimal or maximally good moral deed. This idea is one of the hallmarks of consequentialism:

> Among ethical theories, those that I call 'act-consequentialist' may be characterized roughly as follows. Such theories first specify some principle for ranking overall states of affairs from best to worst from an impersonal point of view. (…) After giving some principle for generating such rankings, act-consequentialists then require that each agent in all cases act in such a way as *to produce the highest-ranked state of affairs that he is in a position to produce.*[4] (Scheffler, 1982:1)

The idea of a maximum of good is not, in itself, incompatible with supererogation. However, the whole point in setting a maximum is that, by the lights of consequentialism/utilitarianism, the deed with the best outcome is *obligatory*. Obviously, this is one of the features of consequentialist and utilitarian theories that comes into conflict with supererogation. It is already a commonplace[5] to point out the incompatibility between the requirement of maximization and supererogation: if supererogation is mainly about doing more than required, then to require the maximum one can do leaves no place for going beyond what is required. As the two (supererogation and the maximization requirement) seem to exclude each other,

---

An act-consequentialist holds that states of affair (outcomes, consequences) can be objectively or impersonally ranked according to their goodness and that any given act is morally right or permissible if and only if its consequences are at least as good, according to the impersonal ranking, as those of any alternative act open to the agent--the doing of an act being itself included among its consequences. (Slote, 1984: 139)

4 My italics

5 See Heyd (1982), Mellema (1991), New 1974, Portmore (2003), Scheffler (1982), Slote (1984).

one is forced to say either that supererogation does not actually exist (and to explain away the commonsensical moral intuition upholding its existence[6]) or to say that consequentialist and utilitarian theories should give up optimization or maximization, respectively. The route usually taken is to accept that supererogation represents a robust intuition of common sense morality and the consequentialist inspired theories should adapt in order to accommodate it. In this sense, authors who want to stick to the basic intuition of consequentialism while being able to accommodate supererogation[7], propose giving up the maximization requirement and adopt instead a satisficing requirement, *i.e.* the view that what is morally required is to do something "good enough" by some adopted standard, not something that is best. If this proposal indeed reconciles the commonsensical notion of supererogation with consequentialism, will be discussed in the following sections.

There are authors such as Zimmerman (1993) and Vessel (2010) who disagree with this widely held verdict of incompatibility between supererogation and maximization requirement. Their strategy will be to try to keep the maximization requirement while introducing other changes in the conceptual frame surrounding the problem of compatibility of the two.

I.3. The threshold of supererogation as satisficing

Obviously missing from the core of any consequentialist or utilitarian theory is the idea of a threshold of what is required, beyond which one may permissibly act in order to obtain even better outcomes. On the other side, supererogation is actually defined by the existence of such a threshold. As a consequence, consequentialist theories trying to accommodate supererogation will usually adopt a kind of threshold of what is required below the optimal or the maximum possible. The kind of threshold to be

6 See New (1974) and Vessel (2010).

7 Slote (1984), Hurka and Schubert (2012), Dreier (2004).

discussed is the one given by the concept of *satisficing*, where the outcome is not supposed to be the best possible one, but one that is "good enough" in the circumstances. This proposal will obviously create room for going "beyond what is required". Nevertheless, it will also leave some room for doubt and debate regarding the success of capturing the commonsensical moral intuition of supererogation by this theoretical device.

I.4. The outcome of the action and the sacrifice of the agent as a cost

The sacrifice of the agent and the overall (presumably) good result of his deeds are not quantified or measured against each other according to the commonsensical view of supererogation. In a consequentialist or utilitarian frame naturally the outcome will be important and it will be important for it to be measurable/ quantifiable. The impersonal maximization of the good is an idea usually aiming at something like the "greater good" of all. The problem with supererogatory action is that the pursuit of the greater good by the agent might come with a heavy cost for the agent. The two-part image of supererogation (that is, permission and superiority of the supererogatory action) does not mention the sacrifice of the agent as a condition for something to be considered supererogatory. It only says that one is permitted not to do it; it does not say why. However, often times the classical paradigmatic examples of supererogatory actions are saintly and heroic deeds, which tend to have a heavy cost for the agent and to bring very good outcomes (for the rest). Even if sacrifice is not always present with supererogatory action, it is enough to have some cases of supererogation that bring very good results with heavy cost (*e.g.* the soldier who saves many lives by sacrificing his own life) in order to generate objections to theories of utilitarian inspiration. For, in a classical "crude" version of utilitarian theory, such heavy sacrifices will be required in case they maximize the result. Therefore, this trait of utilitarianism has become a classical objection usually labeled as the *demandingness objection*: it is too

much to require from an agent to always maximize the goodness of the outcome. Critics point out that this is where the concept of supererogation, properly integrated, would bring some relief from exaggerated utilitarian demands by making sacrifice and heavy cost for the agent optional. The manner in which authors attempt to make this integration is by leaving the agent some latitude regarding the allocation of her time and effort: this is where the agent-centered prerogatives make room for agent-favoring permissions and agent-sacrificing permissions in Scheffler (1982), Slote (1984b) and Hurka and Schubert (2012).

These points from I.1 to I.4 are sensitive points where supererogation and consequentialist theories collide. Authors who will try to make supererogation sit comfortably with consequentialism and utilitarianism will adopt strategies following these fault-lines. They will propose to give up the requirement of "the best outcome" (*i.e.* of optimizing/maximizing) and settle for "good enough", or they will propose to have an agent-centered approach rather than an impersonal one, thereby calling into question the impersonal scale for measuring the goodness of the outcome of a moral deed. A third option will be to enlarge the area of permissions usually available in an utilitarian setting (making more actions permissible for the agent in accordance with commonsensical moral intuitions).

I will present two strategies of this kind:

A) *Giving up the optimization/maximization requirement* and adopting a *satisficing* requirement instead. (Slote (1984))

B) *Giving up the impersonal scale measuring the outcome* (and adopt a more agent-centered approach). (Scheffler (1982)).

II. Satisficing as the Reconciliatory Solution between Supererogation and Consequentialism: Michael Slote

Usually, when supererogation is taken into account in consequentialist contexts, the main concern is not the problem of supererogation, but some larger theoretical point. In Slote's case,

supererogation, seen as a carrier of powerful intuitions of ordinary morality, serves as a test for the consequentialist position, a test meant to reveal if these powerful moral intuitions can be acclimatized in a consequentialist environment.

There are, famously, a number of moral verdicts upheld by common sense morality and denied by consequentialist inspired theories. Slote himself quotes Bernard Williams and Samuel Scheffler as prime examples of authors criticizing consequentialism for its disconnect with important moral intuitions:

> Moreover, critics of optimizing consequentialism have recently tended to focus on one particular way in which such consequentialism implausibly offends against common-sense views of our obligations of beneficence. They have pointed out that (optimizing) act-consequentialism makes excessive demands on the moral individual by requiring that she abandon her deepest commitments and projects whenever these do not serve overall impersonally judged optimality. For example, it has been held by Samuel Scheffler (and others) that it is unfair or unreasonable to demand such sacrifice of moral agents, and by Bernard Williams (and others) that such requirements alienate individuals from their own deepest identities as given in the projects and commitments they hold most dear, thus constituting attacks on their integrity (integralness) as persons. (Slote, 1984:157)

Slote accepts that the traditional form of consequentialism (*i.e.* the one requiring an optimization of the overall outcome) cannot accommodate some of these commonsensical moral intuitions. It is clear that the common moral sense will judge the optimizing requirement to be too demanding, especially in cases of great sacrifices. For example, it might be optimal, in the aggregate, for one soldier to die in order to save all the others in his company. However, the verdict of the common moral sense will be that such sacrifice cannot be demanded and the soldier failing to optimize the outcome in this way cannot be blamed.

According to Slote, this does not mean that consequentialism as such should be abandoned, for one may abandon the optimality requirement without abandoning the basic intuition of consequentialism, namely that there is a connection between the rightness/wrongness of a moral deed and its consequences. His claim is that at least this one intuition of the common sense morality regarding supererogation, namely the intuition that one is not always required to act in a manner producing the best outcome, may be accommodated in a consequentialist setting.

Slote remarks that act-consequentialism has been regarded as a "unitary moral conception", according to which "the rightness of an act depends on whether it produces the best consequences impersonally judged"[8]. However, this conception includes two claims that are conceptually separable: first, that rightness of an act depends only on how good its consequences are, and second, "that the rightness of an act depends on its having *the best*[9] consequences (producible in the circumstances)"[10]. One may uphold the first without endorsing the second. That is, one may maintain that only its good consequences make an act right, yet these good consequences need not be the best possible ones. They need only be "good enough" to make an act right:

> Could not someone who held that rightness depended solely on how good an act's consequences were also want to hold that less than the best was sometimes good enough, hold, in other words, that an act might qualify as morally right through having good enough consequences ,even though better consequences could have been produced in the circumstances? (Slote, 1984: 140)

Therefore, Slote advocates a new kind of consequentialism, one that does not have the optimizing requirement. His strategy is

8 Slote, 1984: 140.
9 My italics.
10 Slote, 1984: 140.

to argue that a 'satisficing consequentialism' (rather than the 'optimizing consequentialism') would agree more with intuitions of common-sense morality of benevolence and would be, in this way, more plausible:

> And since the plausibility of various forms of consequentialism partly depends on how far their implications diverge from the deliverance of ordinary moral intuition, this new form of consequentialism may turn out to have some distinctive advantages over traditional optimizing forms of consequentialism. (Slote, 1984: 152)

Slote borrows the notion of "satisficing" from economics[11] where an action is said to satisfice rationality inasmuch as its outcome is less than the best but nevertheless "good enough".

[11] Dreier (2004) criticizes Slote for his assumption that examples taken from the domain of economics could be made analogous to the ones from the moral domain: Dreier claims that the "good enough" of someone selling a house on the spot for a lower price is not the same as the "good enough" of someone offering *not the best* room, but *a room* to a homeless family, even though they might present a *prima facie* similar structure. The similarity consists in the fact that both are examples of agents choosing an option that, they admit, weighs 'less' on the scale of good than another option available to them. Briefly, the reason why he thinks rational satisficing does not work is that in normal, rational cases, agents will always maximize their preference, never satisfice. What Slote and others have described as cases of satisficing are actually, according to Dreier, cases of maximizing one's preferences. For let us take the example of the person selling their house. If they sell it for a lower price because they do not want to wait longer in incertitude, then this is their preferences, so they are actually maximizing their preferences even if they accept a lower price (because the utility is higher for a satisfied preference, even if the price obtained is lower). In order to truly satisfice, Dreier claims, the person selling the house would have to have a certain preference to which they attach the highest utility (e.g. to obtain a higher amount of money) and then to go against that very preference (*i.e.* to accept less). In short, in order to truly satisfice, one would have to prefer more money and then accept less money (which has an irrational air about it). Instead, what typically happens in examples like this one, claims Dreier, is that one

His first example of an agent plausibly settling for less than the best, but nevertheless "good enough", is the example of someone selling their house: they might accept not the best price but something that is deemed a satisfactory price and is offered more promptly. The reason for accepting the lower price, claims Slote, is not an anxiety about not being able to sell or an indifference towards money, but simply being content with "good enough". A second example is "the snacker": the person who chooses not have an extra snack offered for free even though they know they would enjoy it and they do not fear any bad consequences. They simply decide they had enough to eat: Slote considers this to be a kind of moderation, one that is not a form of asceticism and which "it is difficult to see why it should count as irrational"[12] . By analogy, he claims, we can make equal sense of cases of moral satisficing. For example, a hotel manager helping out a homeless family by giving them a spare room has done nothing wrong if she did not offer them the best room. While she did less than the best, her commendable gesture of benevolence was good enough in the circumstances. Another example[13] used by Slote is the doctor who volunteers to go to a country in need of medical help. However, he is not required to go to the country which is most in need; no matter where he chooses to go (maybe following a personal interest), his gesture would be commendable even if the outcome is not the best possible, but only good enough.

Slote thinks that a consequentialist theory of the good enough can accommodate supererogation. It can do so because it creates a

prefers time and so accepts less money, which is perfectly rational and this is why the example sounds plausible, but it is a *maximizing* example, not a *satisficing* one; in order to satisfice one would have *to accept less of the same thing that one prefers*.

12 Slote (1984), p. 145.

13 For how our intuitions might incline towards supporting a "maximizing" or "satisficing" verdict depending upon the kind of example and context we use, one can read Jenkins and Nolan's article "Maximizing, Satisficing and Context" (2010). Their thesis, in short, is that one will side with "satisficing" if "best" is understood in context as meaning "the few at the top" and with maximizing if "best" is understood as meaning "the one at the top".

threshold for what is required – the good enough – one that the agent may go beyond. Satisficing[14] makes room for supererogation because it is permissible for both the doctor and the hotel manager to optimize, to do more than the required "good enough". Going beyond the established threshold would count in these circumstances as supererogatory:

> One of the chief implausibilities of traditional (utilitarian) act-consequentialism has been its inability to accommodate moral supererogation. But a satisficing theory that allows less than the best to be morally permissible can treat it as supererogatory (and especially praiseworthy) for an agent to do more good than would be sufficient to insure the rightness of his actions. Thus, if the person with special interest in India sacrifices that interest in order to go somewhere else where he can do even more good, then he does better than (some plausible version of) satisficing act-consequentialism requires and acts supererogatorily. But optimizing act-consequentialism will presumably not treat such action as supererogatory because of its (from a common-sense standpoint) inordinately strict requirements of benevolence. (Slote, 1984: 157)

The viability of Slote's solution[15] is not the important point here, so I am not going to try to decide if his proposal of

[14] Hurka (1990) is critical of Slote here because, Hurka claims, Slote is equivocating between two possible meanings for "satisficing": the "absolute" and the "comparative". The absolute satisficing simply establishes a threshold of good enough action without reference to other alternatives present to the agent. The comparative satisficing demands that the outcome of the action be "reasonably close to the best" and so it makes reference to the maximum possible. Hurka believes only the absolute kind of satisficing would work for Slote's version of consequentialism because in a bad situation only absolute satisficing would make sure that enough has been required from the agent, enough to improve the situation significantly; the comparative satisficing, he claims, would only ask for some kind of improvement on the bad situation (which would be some percentage of the maximum).

[15] Slote admits that his solution needs elaboration. The "good enough" of satisficing needs specification. He reviews some conceptions about

reconciliation succeeds. From the perspective of accommodating supererogation, I believe two points are important.

First, the conflict he is trying to resolve is not about the demandingness of the optimizing consequentialist theory. The demandingness objection is a fairly common one against traditional utilitarian and consequentialist theories. The charge against theories of this kind is that they ask too much from the agent by making the optimal or the maximal result obligatory because sometimes the best result comes with a very large cost for the agent (for example, paradigmatic cases of saints and heroes involve extreme sacrifices). And common sense morality objects that such large costs cannot be demanded. This objection will be discussed in detail in the next section, where the impersonal aspect of consequentialism comes into play. Here it is important to notice that the objection discussed by Slote against optimizing consequentialism is not the same as the demandingness objection: the incongruity between common sense morality and optimizing consequentialism discussed here is simply that common sense morality accepts without problems that the agent might do less than the best (even when there is no significant sacrifice involved on the part of the agent):

> So the divergence between common-sense morality and standard (utilitarian) act-consequentialism with regard to such cases cannot be accounted for in terms of a disagreement over whether one can correctly require an agent to sacrifice his own desires, projects and concerns in the name of overall optimality. (Slote, 1984:151)

Slote points out that his example of the hotel manager helping out the homeless family is an illustration of this idea: there is no major sacrifice in this case on the part of the hotel manager, but common sense morality and the traditional consequentialist

---

satisficing from Popper and Bentham and reaches the conclusion that a certain percentage below the maximum would probably be an appropriate conception of "good enough".

view still diverge regarding the optimization requirement. Therefore, the agent may do less than the best not because the optimization requirement asks too much (much more than it is reasonable to ask) but simply because it seems perfectly acceptable and reasonable to do less than what optimization requires. In short, one is allowed to satisfice, not only when there is a large cost for the agent (and because of that) but simply because the action appears to be a perfectly reasonable action.

Secondly, and more important, Slote points out that the concept of supererogation obtained by giving up the optimizing requirement *is not the same supererogation concept* featured in the commonsensical view about supererogation:

> Such consequentialism in effect then allows various sorts of compromise between the demands of impersonal morality and personal desires and commitments. To that extent, it allows greater scope for personal preferences and projects than traditional optimizing act-consequentialism does. However, it offers less scope than would be available on most common-sense views of what an agent may permissibly do. For ordinary morality would presumably allow an agent (capable of doing better) to pursue projects that do not contribute very much to overall human well-being, and satisficing consequentialism – unless it maintains a very weak view about what it is to do enough good – will rule such projects out. (Slote, 1984:158)

The main point here, I think, is that even when one has made some room for going beyond what is required (*i.e.* when satisficing is required, optimizing is not), the obtained concept of supererogation might have a narrower scope than the one attributed by the commonsensical view. The doctor volunteering to take care of patients in a country that she prefers, will behave supererogatorily if her plans change and she agrees to go to a different country, where help is most needed – this will be the verdict of satisficing consequentialism. However, the doctor *is*

required to reach this satisficing threshold (that is, to go to a country in need), which means that satisficing consequentialism will *require* some acts of benevolence that might be declared by common sense morality to be entirely *optional*.

An equally important point is that restricting the concept of supererogation is the result of compromise between the commonsensical notion of supererogation and the impossibility of supererogation (within optimizing consequentialism). This compromise also makes room not only for supererogation but also for a more personal approach to morality. To recall, the consequences of an action were to be measured on an impersonal/impartial scale of goodness for states of affairs. Slote is not arguing, like Scheffler (1982), that one may ignore at times this scale (because of an "agent-centered prerogative") but he is saying that given more choice (by not being under the obligation to optimize), agents may choose something closer to their own plans and aspiration when they are only under a satisficing obligation. The doctor in Slote's example may choose, in this way, a country she is interested in, instead of the country most in need of medical assistance.

In conclusion, Slote's attempt to reconcile supererogation with consequentialism has modified both consequentialism (by giving up the optimization requirement) and the concept of supererogation (which will have a narrower domain than the commonsensical one).

III. The impersonal trait of consequentialist evaluations

When confronted with the conflict between utilitarian maximization requirement and the intuition that some excellent moral deeds cannot be required, some authors have chosen to drop the maximization requirement. That resulted in adopting satisficing as a solution. Once the requirement is established at the 'good enough' level, the agent may be said to go beyond what is required and therefore some room is made for supererogation (even though

it might not have the same wide scope as the commonsensical notion of supererogation; some things required by satisficing might end up as purely optional from the point of view of common sense morality).

This is not the only strategy available in order to acclimatize supererogation in consequentialist and utilitarian environments. Another strategy, sometimes involving satisficing, is based on an objection against the *impersonal and impartial* kind of evaluation demanded by consequentialist and utilitarian frames. At first sight, it does not seem to be a connection between supererogation and the impersonal aspect of consequentialist/utilitarian evaluation of states of affairs. However, once the agent is allowed to depart in various ways (*e.g.* by being allowed to give more weight to his own preference) from the impersonal/impartial way of measuring the overall goodness of the outcome, the obligation to maximize is implicitly dropped and this usually makes room for a notion of supererogation.

The objection against impartiality has a tradition. Bernard Williams describes impartiality in morality in general as being

> (…) something which, indeed, some thinkers have been disposed to regard as the essence of morality itself: a principle of impartiality. Such a principle will claim that there can be no relevant difference from a moral point of view which consists just in the fact, not further explicable in general terms, that benefits or harms accrue to one person rather than to the other. (…) from the moral point of view, there is no comprehensible difference which consists just in my bringing about a certain outcome rather than someone else's producing it. (Williams, 1981: 96)

Impartiality is not a trait of utilitarians only; theories of Kantian and deontological descent have also claimed that being impartial and impersonal are main ingredients of being moral. Nevertheless, utilitarianism has a specific way of demanding impartiality, one implied by the way the sum of individual utilities

is supposed to be calculated. Nobody's happiness is supposed to matter more and the only thing that matters is the maximization of the total, a total in which any kind of individuality is lost (as in any sum total). The acute problem facing this view becomes evident when the utilitarian demands that any resource of time and energy be dedicated to the maximization of the sum total. It would seem that there is no room left for the agent to be involved in personal projects that do not bring sum total maximization of utility.

This kind of critique of utilitarianism, about the tendency of utilitarian theories to require the atrophy of the personal, has prompted attempts of reconciliation between the impersonal approach and an agent-centered approach. Scheffler's agent-centered prerogative is such an attempt.

III.1. Scheffler's agent-centered prerogative and making room for supererogation

Scheffler (1982) lists two classical objections to theories of consequentialist/utilitarian inspiration, one targets the "view from nowhere" and regards the distribution relations between agents in a utilitarian scenario. The other focuses on the point of view of the agent making decisions in a consequentialist manner.

The first objection is that utilitarianism prescribes ignoring the unhappiness of a few if this leads to a maximization of overall happiness. The second is Bernard Williams' worry that living in accordance with utilitarian prescriptions results in alienation from one's own life projects. Scheffler argues that Williams' worry about alienation is rather vague. If it means that utilitarianism demands that sometimes we give up our own plans when these would extract a large cost or impose too much damage on the others, then this does not seem wrong and it is something required by pretty much all non-egoistic moral theories, not only by utilitarianism. However, a charitable interpretation of Williams' objection would be, according to Scheffler, that alienation takes place, not because the agent is required to *give up to some* of his projects, but because

the agent is required to evaluate *all* personal projects by reference to the impersonal scale measuring the capacity to increase the overall goodness or happiness; this is quite unnatural because this is not how we usually evaluate our projects:

> Utilitarianism thus requires the agent to allocate energy and attention to the projects and people he cares most about *in strict proportion* to the value from an impersonal standpoint of his doing so, even though people typically acquire and care about their commitments quite independently of, and out of proportion to, the value that their having and caring about them is assigned in an impersonal ranking of overall states of affairs (Scheffler, 1982: 9)

Notably, Scheffler considers this second objection as an objection not only against utilitarian theories, but also against any kind of consequentialist theory because they share the theoretical feature of impartial ranking of overall states of affairs.

Scheffler's proposal for tackling this difficulty is the introduction of an agent-centered prerogative which is meant to make it permissible for the agent to spend resources on projects evaluated *out of proportion* with the impersonal scheme:

> On a plausible view of this kind the answer to the question of whether an agent was required to promote the best overall outcome in a given situation would depend on the amount of good he could thereby produce (or evil he could avert), and on the size of the sacrifice he would have to make in order to achieve the optimal outcome. More specifically, I believe that a plausible agent-centered prerogative would allow each agent to assign a certain proportionately greater weight to his own interest than to the interests of other people. It would then allow the agent to promote the non-optimal outcome of his choosing (...) (Scheffler, 1982: 20)

Some but not all kinds of projects may receive this permission, of ignoring the impersonal ranking. I will not enter into the details of his proposal here, like the circumstances in which it is permissible to ignore the impersonal ranking, how the greater weight for the agent's plans will be assigned, difficulties of this view and so on. Whatever the details of this proposal, one thing becomes clear, there is room made for supererogatory conduct:

> Since it would permit people to devote energy and attention to their projects and commitments out of proportion to the weight from the impersonal standpoint of their doing so, the view would lack the feature that generates that objection. But at the same time, it would certainly on such a view always be *permissible* for an agent to bring about the best available state of affairs. Thus there might be an agent who willingly sacrificed his own projects for the greater good; on this view his conduct would be supererogatory. (Scheffler, 1982: 22)

Scheffler main focus here is, obviously, not supererogation. Rather, the main concern appears to be finding a way to meet certain objections (related to the impersonal/impartial way of evaluating outcomes) raised by common sense morality against consequentialist and utilitarian theories. This is an enterprise similar to Slote's (1984), but while Slote was *keeping the impersonal ranking* of states of affairs and was proposing a lower threshold for what is morally required *on that scale*, Scheffler wants to proclaim the agent's *independence from the impersonal scale* in certain circumstances (circumstances in which the scale may be legitimately ignored).

According to Scheffler, the agent is allowed sometimes not to take into consideration the obligation of optimization/maximization because the agent is allowed sometimes to evaluate the outcome of his action independently from the evaluation on the impersonal scale. This is the agent's prerogative, to be morally permitted in certain circumstances to ignore the demand of producing the impersonal best ranking outcome and thereby to ignore the maximization demand. In circumstances where the agent has this

prerogative, she might nonetheless choose to take into account the demands of the impersonal ranking at a cost for her own plans. This is the spot where supererogation falls into place: when the agent is allowed to choose between ignoring and taking into account the demanding impersonal scale, the agent might choose the „greater good" of the impersonal demands and act in this way supererogatorily. Examples are easy to find: Slote's doctor may choose to go to a country where her medical help is going to have the greatest impact (impersonally judged) or in a country where the impact is more modest, but where the doctor might also have a personal interest in being there; in this case, going to the country where the impact is greatest (impersonally judged) is supererogatory.

There are several assumptions here that Scheffler does not discuss[16]. In a comparison of the outcomes of personal plans with outcomes of the actions prescribed by the impersonal ranking, the assumption seems to be that the impersonal ranking is going to be more demanding, *i.e.* it is going to ask for more things to be sacrificed. This needs not always be the case. A personal plan may sometimes involve many more sacrifices of well-being than the impersonal demands of morality: the impersonal evaluation of outcomes might require the doctor to go to the country most in need, but joining a religious order might require going to the county most in need, celibacy and asceticism.

Another question is whether a sacrifice on the part of the agent is an indication of morally better outcomes impersonally judged. This needs not be always the case either[17]. The larger picture of presuppositions seems to be divided in two: on the one hand, actions made according to personal plans would presuppose less sacrifice from the agent and moral outcomes scoring lower on the impersonal scale of the good (*i.e.* the outcomes would be less good for others but better for the agent); on the other hand, actions

[16] This is not a criticism, nor, for that matter, surprising since supererogation was not his main topic of discussion.

[17] For an illuminating discussion of how sacrifice might lead to worse overall results, one might see Jean Hampton, "Selflessness and the Loss of the Self "(1993).

made by accepting the demands of the impersonal scale would presuppose more (or more serious) sacrifices from the agent and they would score higher, even highest on the scale (*i.e.* the outcomes would be best for others but less good for the agent). These are common assumption to make, but by no means unassailable ones. Especially in a consequentialist framework, one cannot take for granted that the greater the sacrifice of the agent is, the greater the value of the outcome will be.

However, Scheffler's claim that there is room made for supererogation by his agent-centered prerogative is not affected by the issue I have just raised. What is new and remarkable about his way of seeing supererogation is that there is no threshold, properly speaking, that the agent would go above and beyond. The agent is presented with a choice (in the circumstances where the prerogative is allowed) of doing more for herself (*i.e.* personal plans) or more for others (paying heed to the impersonal ranking). She is permitted to do either and there is an option where she would do more by impersonal standards. But in choosing to take into account and obey the obligation presented by the impersonal scale, she does not go beyond an established threshold of obligation that was lower o*n the same scale*. She has simply chosen *another scale* (a more demanding one).

As the commonsensical concept of supererogation involves a threshold of duty beyond which actions are considered excellent but non-obligatory, I will conclude that Scheffler's kind of supererogation is a different one than the commonsensical one. Therefore, not only his version of consequentialism is modified in trying to make room for supererogation (by introducing the agent-centered prerogative) but also the concept of supererogation is modified in the process. The kind of supererogation emerging from his considerations is not properly describable as "an action going beyond duty". Rather, it would be better described as choosing an action considered best on the impersonal scale in circumstances where the agent is allowed to ignore the impersonal considerations of greater good and tend to her own plans and

interests (e.g. the doctor who would choose to go where she is most needed instead of following her personal preferences).

IV. Conclusion

The strategies used to acclimatize supererogation in a consequentialist environment might seem to be unilateral strategies, proposing a one-way modification: consequentialism (or utilitarianism) needs to adapt itself to commonsensical moral intuitions about what can be required from a moral agent and therefore needs to change the force and the extent of its requirements. As a result, the agent will be morally justified in doing less than the best[18] (by various theoretical devices).

My aim was to show that the proposed modifications are not unilateral, in that they also modify the commonsensical concept of supererogation. As a result, some form of supererogation can be allowed to subsist within various consequentialist/utilitarian frameworks. However, the positions discussed here make room only for a modified (usually restricted) concept of supererogation, not for the full commonsensical concept of supererogation. Therefore, the problem if they truly accommodate supererogation-as-we-know-it (instead of changing supererogation to fit their theoretical needs) remains an open question.

One might argue that the commonsensical concept of supererogation is a vague, incomplete one. Various theories may fill in the blanks according to their specificity without actually changing the main contours of the commonsensical picture. However, I believe this cannot be said, at least about the theories presented here. Slote and Scheffler not only add various features to the commonsensical picture of supererogation, but they also subtract other important features, which change amounts to a more visible and clearer transformation of supererogation.

[18] That is, the agent will be justified in acting such that the outcome of her actions will not be the best outcome, impersonally considered.

Slote's proposal is to drop the optimizing requirement and to adopt satisficing as a reasonable threshold of what can be morally required from an agent. This is an important change in a consequentialist theory. It is such an important change that other authors[19] have doubted that without a maximization/optimization requirement a theory may be said to still be utilitarian/ consequentialist. Nevertheless, this is not the only change taking place: the concept of supererogation allowed by satisficing has a narrower domain than the commonsensical one because some actions declared obligatory by the satisficing theory will still be merely permitted according to common sense.

Scheffler's proposal is to allow agents to sometimes not measure the outcome of their actions according to the impersonal scale. In this way an agent-centered prerogative is created, a prerogative that allows the agent to choose between the impersonal overall good and the personal private good (under certain conditions). If the agent chooses to act in accordance with the impersonal scale, then that is supererogatory action, according to Scheffler. However, the resulting concept of supererogation does not have a threshold of actions that are required (which is an important feature of the commonsensical notion of supererogation). There are two ways of measuring the outcome and the agent may choose one or the other, but no *one* threshold of what is required.

To conclude, I believe that the various strategies used to accommodate supererogation within consequentialism and utilitarianism sometimes change the concept of supererogation itself. Not only does ordinary supererogation appear to be rather difficult to incorporate in an utilitarian setting, but also, the question remains open if these theories have indeed incorporated the commonsensical moral intuitions about what is permitted and what can be required.

[19] New (1974), Vessel(2010), Zimmerman (1993)

REFERENCES

Dreier, James (2004). Why Ethical Satisficing Makes Sense and Rational Satisficing Doesn't. In *Satisficing and Maximizing*, edited by Michael Byron, 131-154. Cambridge: Cambridge University Press.

Hampton, Jean (1992). "Selflessness and the Loss of Self." *Social Philosophy and Policy* 10: 135-165.

Heyd, David (1982). *Supererogation: Its Status in Ethical Theory.* Cambridge: Cambridge University Press.

Hurka, Thomas (1990). "Two Kinds of Satisficing." *Philosophical Studies* 59: 107-111.

Hurka, Thomas, and Schubert, Esther (2012). "Permissions to Do Less Than the Best: A Moving Band." *Oxford Studies in Normative Ethics* 2: 1–27.

Jenkins, Carrie S., and Nolan, Daniel (2010). "Maximising, Satisficing and Context." *Noûs* 44: 451-468.

Mellema, Gregory (1991). *Beyond the Call of Duty: Supererogation, Obligation, and Offence.* State University of New York Press.

New, Christopher (1974). "Saints, Heroes and Utilitarians." *Philosophy* 49: 179-189.

Portmore, Douglas W. (2003). "Position-Relative Consequentialism, Agent-Centered Options, and Supererogation." *Ethics* 113: 303-332.

Scheffler, Samuel (1994). *The Rejection of Consequentialism: A Philosophical Investigation of the Considerations Underlying Rival Moral Conceptions.* Oxford: Oxford University Press.

Slote, Michael (1984). "Morality and Self-Other Asymmetry." *Journal of Philosophy* 81: 179-192.

Vessel, Jean-Paul (2010). "Supererogation for Utilitarianism." *American Philosophical Quarterly* 47: 299-319.

Williams, Bernard A.O. (1981). *Moral Luck: Philosophical Papers, 1973-1980.* Cambridge: Cambridge University Press.

Zimmerman, Michael J. (1993). "Supererogation and Doing the Best One Can." *American Philosophical Quarterly* 30: 373-380.