Can Virtue Ethics Account for Supererogation?

DAVID HEYD

In his classical article, 'Saints and Heroes', James Urmson single-handedly revived the idea of supererogation from it astonishingly long post-Reformation slumber.[1] During the first two decades after its publication, Urmson's challenge was taken up almost exclusively by either utilitarians or deontologists of some sort.[2] On the face of it, neither classical utilitarianism nor Kant's categorical imperative makes room for action which is better than the maximizing requirement, on the one hand, or beyond the requirement of duty, on the other. Nevertheless, both utilitarians and Kantians, as well as deontic logicians, offered more flexible and sophisticated versions of their respective theories which could accommodate supererogatory action. In my 1982 book on supererogation I tried to address the question whether virtue ethics could capture that new category of actions which are praiseworthy though not strictly required. But the focus of my discussion was mostly Aristotle (and Seneca) and accordingly more interpretive in nature.[3] However, that was just before the tremendous surge of interest in virtue ethics and the vast literature debating the merits of agent-based vs. action-based approaches in moral theory. It turned out that fitting supererogation into virtue-based moral theory proved to be a more difficult task than doing so in consequentialist and deontological theories. Some argued that supererogation could nevertheless be accounted for in aretaic terms; others held that it could not and that this fact attested to either a theoretical weakness – even if not a refutation – of virtue-based ethics, or to the incoherence of the concept of supererogation.

Unlike normative judgements (including those relating to supererogatory action), moral theories are not true or false but only stronger or weaker (good or bad) in explaining and justifying normative

[1] James Urmson, 'Saints and Heroes', in A. I. Melden (ed.), *Essays in Moral Philosophy* (Seattle: University of Washington Press, 1958), 198–216.

[2] A few notable examples of this 'first generation' of philosophers writing on supererogation are Joel Feinberg, Roderick Chisholm, John Rawls, Joseph Raz, David Richards, Thomas Hill and Michael Stocker.

[3] David Heyd, *Supererogation: Its Status in Ethical Theory* (Cambridge: Cambridge University Press, 1982), chap. 2.

doi:10.1017/S1358246115000193 ©The Royal Institute of Philosophy and the contributors 2015

https://doi.org/10.1017/S1358246115000193 Published online by Cambridge University Press

judgements. The theoretical issue which is discussed in this article relates to alternative *conceptual mappings* of moral phenomena. The argument made here is that a deontological framework is much superior to that of 'virtue ethics' in both recognizing the special place of supererogation in moral behaviour (which the more radical virtue ethicists such as Roger Crisp refuse to do) and accounting for supererogatory action once its importance is acknowledged (which is the more common line taken in current virtue ethics).

In the first section of the paper a few general comments are made about the indispensability of deontic concepts in virtue theory. In the second section the crucial role of the distinction between duty and supererogation within deontologically based morality is presented. The third section tries to show why virtue ethics cannot provide a good account of the concept supererogatory action. Then, in the fourth section, the possibility of denying supererogation altogether due to its incompatibility with virtue ethics is addressed and rejected. Finally, the last section is devoted to meta-ethical or methodological issues that arise in the debates about virtue and supererogation – the alleged parochiality of the concept of supererogation, the act-agent distinction (which is parallel to the distinction between the deontic status of the act and the praiseworthiness of the agent), and the general question of alternative conceptual mappings of ethical judgements and evaluations.

1. The unavoidability of deontic terms in virtue ethics

In a famous passage in the *Nicomachean Ethics* (1105a)[4], Aristotle brings to our awareness the threat of circularity in the definition of just actions as the actions performed by the just person. We can become just only by doing just actions, but the genuinely just actions are only those performed by a just person, that is to say out of a certain character. This statement reminds us that Aristotle is on the one hand opposed to a purely agent-based ethics, yet on the other hand is committed to the primacy of character over action, at least in the moral sphere (the arts having a different conceptual structure). So, on the one hand, action is not only necessary for the acquisition of virtue but also for *eudaimonia*, which is achieved through activity rather than through character alone (since character is in

[4] All quotations are taken from the W. D. Ross translation in Richard McKeon (ed.), *The Basic Works of Aristotle* (New York: Random House, 1941).

https://doi.org/10.1017/S1358246115000193 Published online by Cambridge University Press

itself only disposition, i.e. mere potentiality). On the other hand, virtue is the ultimate standard or test of ethical behaviour and in that sense has priority over action in the circle mentioned above.

But can virtuous character or disposition serve as the sole criterion for right action? Plato might be a candidate for holding such a radical version of virtue theory: virtues are signs of the harmony of the soul, of an inner psychological balance, of mental health. Their basis is natural or metaphysical. Thus, *whatever* the virtuous person does, it would be morally right. But even in this conception virtue must relate to the *functioning* of human beings in the real world and in society and hence cannot be identified independently of action and social behaviour. In other words, virtue ethics cannot dispense with rules of actions and standards of what is considered a right action. Beyond the *strength* of character, the psychologically balanced personality, Plato recognizes that the virtuous person must be wise, or prudent (*phronimos*), namely know (in the standard cognitive sense) that a particular course of action in specific circumstances is called for. Virtue might be the basis for agent-oriented morality, but it must be complemented with criteria of right action. As Kant famously argues in the opening of the *Groundwork*, virtue can be abused when not guided by "the moral law": a courageous criminal is worse than a coward. Justice cannot be described in purely psychological terms of the relations between the various powers of the soul and the Platonic theory of justice must supplement such description with certain distributive or retributive criteria (that the good person *knows*). I follow here Michael Slote, who has compellingly argued that pure versions of agent-based ethics must be regarded as 'autistic', detached from the world and from other people and their circumstances.[5] The virtuous person does what is noble because it is noble. It is not the case that the act is noble because the virtuous person did it.[6] I would add to that the extreme example of Rousseau's obsessive concern with being 'true to oneself', or in our

[5] Michael Slote, 'Agent-Based Virtue Ethics', in Roger Crisp and Michael Slote (eds.), *Virtue Ethics* (Oxford: Oxford University Press, 1997), 244–5. See also Ramon Das, 'Virtue Ethics and Right Action', *Australasian Journal of Philosophy* 81 (2003), 324–339. Das talks about the risk of the 'insularity' of pure virtue ethics that strives to break the circularity of Aristotle's description by detaching virtue from right action.

[6] Slote, p. 240. "[I]t is once again not simply true that agent-based theories inevitably treat human actions as subject to no moral standards or requirements. Rather, those requirements and standards operate and bind, as it were, *from within*." p. 244 (emphasis in the original).

https://doi.org/10.1017/S1358246115000193 Published online by Cambridge University Press

language of 'authenticity', which illustrates the vacuity of such exclusive concern with the self as a guiding ethical ideal.

Rosalind Hursthouse has suggested this often quoted definition: 'An action is right *iff* it is what a virtuous agent would characteristically (i.e. acting in character) do in the circumstances'.[7] This definition does not fall into the trap of the 'insularity' of character (from the world) because it refers, like Aristotle himself, to 'the circumstances' in which virtue is exercised in a particular action in the world. But we still need some test of what makes a particular action in particular circumstances 'right' (or wrong), a test which would be independent of the character of the agent and would reflect some normative judgement about the objective situation. Without such a test the question of what we ought to do in a given situation would remain at least underdetermined.[8] For the general disposition to courage cannot be sufficiently fine-tuned so as to define what an act of courage is in any specific situation. But adding such a test would introduce a typically deontic or consequentialist dimension into virtue ethics. Thus, a courageous person would perform courageous acts, but those acts are defined as courageous according to their potential service to others in the circumstances, their chance of success, and other features of the situation which make some sort of rational requirement upon the agent (and which the *phronimos*, the wise or prudent agent, would be able to track).

Sometimes these requirements are referred to by Aristotle as duties (typically in cases of justice and contracts) and sometimes as what is 'fitting' for an agent to do; but in either case, an external deontic source of normativity is added to the purely agent-based characterization of a psychological disposition or character. Accordingly, I follow Roger Crisp's reading of Aristotle that we have a duty to *do* what is virtuous.[9] It might also be the case that beyond our duty to do what is just, we have a duty to *be* virtuous, to develop in us (and of course in our children or citizens) the virtue of justice, that is to

[7] Rosalind Hursthouse, *On Virtue Ethics* (New York: Oxford University Press, 1999), 28.

[8] Robert N. Johnson, 'Virtues and Rights', *Ethics* 113 (2003), 810–834. Hursthouse also speaks about helping a wounded stranger lying by the roadside as 'absolutely required', and although she says that it is a requirement of charity rather than of justice, it is clearly a typically deontic category. *On Virtue Ethics*, 6.

[9] Cf. Roger Crisp, 'Supererogation and Virtue', in Mark Timmons (ed.), *Oxford Studies in Normative Ethics*, vol. 3 (Oxford: Oxford University Press, 2013), 13–34.

https://doi.org/10.1017/S1358246115000193 Published online by Cambridge University Press

say, to do just actions out of the disposition of justice.[10] But it seems that virtue ethics cannot explain the requirement to be or become virtuous, since, if we accept the standard Hursthouse analysis, we can only say that to be or become virtuous is what a virtuous person would characteristically strive to do. Yet, why should the virtuous person serve as a model for imitation if there is no independent duty to become one?

Interpreters of Aristotle's doctrine of the mean emphasize that virtue judgements must be according to Aristotle himself *relativized* to the particular circumstances in which the action or the feeling of the agent is assessed: the virtuous person feels anger or pity at the right time, with reference to the right people and about the right objects, etc.[11] But how can this crucial *rightness* of feelings and actions be determined in virtue ethical terms when the whole point of the relativization is to regard the circumstances in the world as the object of judgement? Again, we must have prior, independent standards for who should be blamed for an offensive act or when such act should be tolerated or forgiven in order to judge the 'rightness' of our moral feeling of anger towards her.

So it seems that neither in Aristotle's view of virtue nor in modern virtue ethics can deontological considerations of justice or consequentialist considerations of overall utility be sidestepped. This, as we shall see in the next two sections, raises for virtue ethics the challenge of supererogation, since there might be actions that go even beyond what the virtuous person would characteristically do in a given situation.

2. The deontic distinction between duty and supererogation

Both Kantianism and utilitarianism have failed to give an account of supererogation (as noted by Urmson), although unlike virtue-based ethics they do not completely ignore it. In both Kant and Sidgwick

[10] Gregory F. Mellema, 'Moral Ideals and Virtue Ethics', *Journal of Ethics* 14 (2010), 173–180. Mellema speaks about *trying* to realize moral ideals, which is sometimes obligatory and sometimes supererogatory, like in the case of generosity (in which we would blame a person for *never* giving any free donation). But it is not clear that we can make sense of a duty to attempt 'from time to time' to perform courageous acts or to display temperance in particularly difficult actions that could be described as supererogatory.

[11] *Nicomachean Ethics*, 1106b.

https://doi.org/10.1017/S1358246115000193 Published online by Cambridge University Press

we find attempts to somehow address the category of action beyond the call of duty.[12] These attempts attest to at least an implicit recognition (which does not exist in much of virtue ethics) that supererogation is a real challenge to any duty-based moral theory and hence must be tackled. My contention is that not only is the acknowledgement of a distinction between duty and supererogation imposed on deontological (and utilitarian) theories, but that the distinction finds its most natural explanation in terms of these theories.

Virtue ethicists have often accused the deontological conception of morality as being too 'juridical' and overly preoccupied with justice, contracts, promises, and rights. In other words, it is too minimal in its scope of norms and gives priority to prohibitions of causing harm over positive duties to do good. Indeed it is true that social morality as it is often described in the last few decades is concerned with the minimal conditions of social co-existence, in the solution of game theoretic dilemmas, in the problem of trust in interpersonal interaction, in issues of cooperation and coordination, in the protection of rights and in contract-based political morality. Utilitarianism has also replaced its classical maximizing principle with a more minimal satisficing one. However, this image of morality is just apparent and it would be misleading to conflate deontological ethics with the law. Moral duty has always been understood in a much wider, positive and demanding way, which includes norms of care, the inculcation of justice, and assistance to the needy (even when they have no correlative right to it). The so-called 'imperfect duties' are of this non-juridical kind as has been most famously articulated by Kant.

Yet, even within this non-juridical conception of deontological ethics, despite its informal and open-ended nature calling for interpretation and discretion, there is an inbuilt *limit* to the scope of duty. There may be a normative disagreement about the dividing line between what is required of us and what is purely optional, but there is no doubt that Dr. Rieux in Albert Camus' *The Plague*, or the volunteers in 'Doctors without Borders', do more than is required of them.[13]

[12] For a discussion of such attempts, see my *Supererogation* chaps. 3 and 4.

[13] One should, however, note that although signing up as a member of the organization is optional, once serving patients in a war zone the doctor is under an obligation to continue treating them.

https://doi.org/10.1017/S1358246115000193 Published online by Cambridge University Press

These people are Urmson's saints and heroes, although I want to argue that common daily actions such as acts of forgiveness or unexpected, unrequired gifts are equally supererogatory. Some disagree, arguing that forgiveness and gifts are duties; but then, what about promises? It is surprising that the literature on supererogation has not noticed the supererogatory nature of promising. Indeed, fulfilling promises is the standard example of an obligation, but *making* a promise can never be an obligation or a duty.[14] If one is under a moral duty to Φ, then promising to Φ is redundant (except for cases in which one wants to convey to the receiving party the message that despite being a burdensome duty, one is resolved to do what is required of one); and if Φ-ing is prohibited so is promising to Φ. But promising is often a good, an altruistic commitment to another who benefits from being able to rely on the promise, and it involves some price to the promisor in terms of the self-limitation of her freedom of future action. Hence promising is always supererogatory although usually involving no heroism or extreme self-sacrifice.

Accordingly, I want to suggest that supererogation and duty are mutually dependent. They are correlative concepts. Supererogation cannot be *conceptually* articulated without reference to duty and duty cannot be *normatively* justified without reference to what lies beyond duty. In that respect, supererogation is *essentially* a deontic concept or a phenomenon which is naturally accounted for in deontological theory. Furthermore, supererogatory acts are of the same kind as required or obligatory acts in the sense that they promote the same kind of value, namely what is morally good, only to a further extent. Onora O'Neill puts it in terms that are similar to what I have called the conditions of correlativity and continuity: 'supererogatory action… must be understood by *reference* to duty', and 'in supererogatory action the ordinary *measures* of duty

[14] There is some debate in the literature on the possibility of promising to supererogate and the question whether fulfilling such a promise (e.g. offering you a ride to the airport in the middle of the night) is an obligation or a supererogatory act. I find this debate odd, since it seems clear that making a promise is always supererogatory while fulfilling a promise always, *prima facie*, obligatory despite the supererogatory nature of the promising act itself. See, J. Kawall, 'Promising and Supererogation', *Philosophia* 32 (2005), 389–398; David Heyd, 'A Comment on Kawall's "Promising and Supererogation"', *Philosophia* 32 (2005), 399–403; and Claire Benn, 'What is Wrong with Promising to Supererogate?', *Philosophia* 42 (2014), 55–61."

https://doi.org/10.1017/S1358246115000193 Published online by Cambridge University Press

rather than the category of duty are exceeded'.[15] That is to say, supererogatory action is not only transcending the 'juridical' kind of social obligations but also the imperfect duties often referred to as 'duties of virtue'. If that is the case, supererogatory action seems to go beyond what the virtuous person would characteristically do (his or her duties of virtue), which brings us back to the difficulty of accounting for supererogation in virtue ethical terms.

3. The categorical distinction between virtue and supererogation

One superficially tempting way to analyse supererogation in virtue-ethical terms would be to extend Hursthouse's original definition: an act is supererogatory *iff* a virtuous agent would characteristically (i.e. acting in character) do it in the circumstances. But as has already been shown,[16] this definition will not do because it conflates duty and supererogation rather than distinguish between them. It is indeed true that we speak of the virtuous agent as someone who does not *just* act according to duty but does so out of a particular disposition or character. But that does not mean that a virtuous person does necessarily *more* than her duty. If the virtuous person has only one moral option in any particular situation, then the question of whether this choice is obligatory or supererogatory does not make sense. It is simply the only right choice. If, on the other hand, virtuous agents can choose between doing what is their duty and doing what goes beyond it, then there must be two kinds of virtuous people or virtuous choices.

And indeed there are suggestions that we should distinguish between the 'maximally virtuous' and the 'minimally virtuous' person, between 'supreme' or 'saintly' virtue and ordinary or 'perfectionist' virtue.[17] But such categorical distinctions between *degrees* of virtue are not easy to explain. People can obviously be more or less

[15] Onora O'Neill, *Towards Justice and Virtue* (Cambridge: Cambridge University Press, 1996), 207–8.

[16] J. Brännmark, 'From Virtue to Decency', *Metaphilosophy* 37 (2006), 589–604; Jason Kawall, 'Virtue Theory, Ideal Observers, and the Supererogatory', *Philosophical Studies* 146 (2009), 179–196.

[17] C. Swanton, 'Satisficing and Perfectionism in Virtue Ethics', in M. Byron (ed.), *Satisficing and Maximizing* (New York: Cambridge University Press, 2004), 187. For a critical response, see Kawall, 'Virtue Theory, Ideal Observers and the Supererogatory', 182–4.

https://doi.org/10.1017/S1358246115000193 Published online by Cambridge University Press

virtuous, but those who are less virtuous are constantly taking *the* virtuous person as their model and cannot themselves serve as model for the non-virtuous individual trying to become a good person. The theoretically idealized person of virtue is the common exemplar for both the minimally and the maximally virtuous human beings. The question left is whether 'saintly' dispositions of the kind exemplified by Mother Theresa or St. Francis are part of the character of that ideal model of virtue. This is a substantive issue to which different normative answers are given. But once we decide that they are exemplars of virtuous character, they cannot be set aside as lying beyond the scope of the ideal of virtue for *all* moral agents (namely, that which can be expected also of ordinary human beings). But even if we could maintain a two-level description of the virtuous person, what could be the criterion of the distinction other than deontic. What makes a person not just simply virtuous but virtuous in a saintly and heroic way are the kinds of *acts* she performs rather than any particular stable disposition. For is there a more courageous profile of character than that of the typically courageous person? And if so, why should not the typically courageous emulate the supreme form of courage?

Aristotle's doctrine of the mean seems to suggest that indeed virtue allows for no such categorical degrees, and that although relativized to particular circumstances there is still only one virtue of courage which should explain and justify both the act of jumping from an airplane with a parachute and throwing oneself on a hand grenade before it explodes and kills one's companions (Urmson's proverbial and extreme example). And accordingly we have to decide whether the latter action is either a straightforward duty (although hard to fulfill) or a foolish self-sacrifice which goes against the ultimate self-serving goal of the virtuous life (namely achieving *eudaimonia*) and hence a kind of excess or plain vice. Again, although Aristotle insists that moral judgement be relativized to the particular conditions of a concrete action, he equally insists that there is always one *objective* answer to the question of what is the right thing to do and this is what the *phronimos* or virtuous person would do in those circumstances (even if it is 'grasped by perception rather than by reasoning'). In that sense Aristotle's theory leaves no room for supererogation.[18]

[18] We do not know what Aristotle would have thought on the soldier throwing himself on the hand grenade, but we do witness some of the tension between what is simply virtuous and what is 'grandly' so in his discussion of magnificence (1122a–1123a). On the one hand, magnificence is

https://doi.org/10.1017/S1358246115000193 Published online by Cambridge University Press

Although supererogatory acts often display a virtuous disposition (like courage, piety, generosity, sympathy etc.), they do not arise out of any separate distinct virtue. It is doubtful whether there is a particular way to inculcate in children (or in ourselves) a disposition to act supererogatorily and hence that there is such a thing as 'supererogatory character'. Indeed, we can speak of a giving, forgiving and tolerant character – which typically is expressed in supererogatory action, and saints are (by definition) those who choose a life which is based on supererogatory resignation (as exemplified by the three tranditional Catholic "counsels" – poverty, celibacy and obedience). However, actions beyond the call of duty are very often performed on an impulse, with no prior deliberation, and are not necessarily the expression of a stable inclination, firm character, a consistent life plan, or a behavioural policy. There is some arbitrariness in the choice to go beyond duty and neither others nor the agent herself can usually predict it. This makes supererogatory behaviour different and even incompatible with the ethics of virtue. Based on individual discretion, supererogatory action has an ad hoc nature and is not universalisable. The absence of any typology of the personality of people who saved Jews during the Third Reich, or any good generalization about their motives, empirically proves the absence of underlying character traits of these moral heroes.[19] Furthermore, in contrast to

identified as a separate virtue from that of liberality, magnificence being a virtue of only those who have a lot of money and are spending it in style. In that respect it is a sub-category of liberality, of the right disposition in behaviour related to money spending. But I am not sure that magnificence has a parallel sub-category in the field of actions related to danger, namely courage, and that there is an extra sub-category of courage that relates to people who have a particular natural or circumstantial property of being able to sacrifice their lives by throwing themselves on hand grenades. The reason is that unlike magnificence, the resources required for acting courageously are not external (like money) but internal (an acquired disposition) and hence an attainable (or morally required) goal for any human being.

[19] Research about the "Righteous among the Nations", as non-Jews who saved the lives of Jews during the Holocaust are called, demonstrates the heterogeneity of motives (not to speak of social class, gender, religious belief and education) of these people. They could be religious, humanitarian, ideological, or personal acquaintance. It is nevertheless a fairly good generalization that the supererogatory saviors share one character trait-modesty. "We did what we were supposed to do", or "what else could we do" are typical responses to the question "why did you take that risk of death to you and to your family?".

https://doi.org/10.1017/S1358246115000193 Published online by Cambridge University Press

Aristotle's insistence that virtuous acts are done with ease, even with style, acts of supererogation may be performed with much effort, pain or even reluctance – that is to say, not 'naturally'.

The attempt to accommodate supererogation within virtue ethics sometimes takes the form of the original deontic formula suggested in the 1960's by Roderick Chisholm: supererogatory acts are those which it is right (or good) to do but not wrong (or bad) not to do.[20] Michael Slote, for example, who represents the 'sentimentalist' theory of virtue, proposes on these lines that supererogatory acts are those which show 'an unusually high degree of empathy' but failure to do them does not 'evince an absence of normally or fully developed human empathy'.[21] Slote is introducing here an implicit notion of what can be expected of people ('what people can be led to develop') and by that creates a two-tier morality – that for highly empathetic people and that for ordinary people with normal capacities for empathy. But the question then arises: is this distinction into two categories of people a natural description, a psychological typology, or is it a normative principle of what can be expected of people? If it is the former, then it is not clear why for people with naturally given high degree of empathy the allegedly supererogatory acts of self-sacrifice are not obligatory. And if the distinction between the two kinds of people is of a normative nature it is not clear on what basis are the 'ordinary people' exempted from the duty to do such acts of supreme empathy (or virtue)? A person of virtue would definitely be expected to help a victim of an accident in the street, but would that empathetic care also include visiting him later in the hospital? Can the virtue of empathy distinguish between the two acts as hinted at by the Biblical story of the Good Samaritan (who both carries the wounded man to an inn and pays for all the costs of his care later on)?

I believe that there are such principles of 'exemption' from highly demanding standards of action but that they cannot be formulated in virtue ethical terms but only through drawing the line limiting duty as a deontic concept. Thus, failure to jump on an exploding hand grenade may often reflect exactly the lack of empathy (or better, for our analysis, the lack of courage) and be motivated by sheer instinctive self-regarding concern for physical survival (or fear) and hence

[20] Roderick M. Chisholm, 'The Ethics of Requirement', *American Philosophical Quarterly* 1 (1964), 147–153.

[21] Michael Slote, 'Famine, Affluence, and Virtue', in Rebecca C. Walker and Philip J. Ivanhoe (eds.), *Working Virtue* (Oxford: Clarendon Press, 2007), 291.

https://doi.org/10.1017/S1358246115000193 Published online by Cambridge University Press

cannot be said not to reflect the absence of a sentiment or a virtue. Empirically, it is the same limit of empathy (or courage) that causes people to avoid jumping on a grenade and jumping from a plane with a safe parachute. That is to say, lack of empathy or virtue cannot differentiate between supererogatory and obligatory action. The difference between supererogatory self-sacrifice and obligatory self-sacrifice is not marked by the aretaic language of psychological dispositions but by the deontic terms of what can be justifiably demanded of individuals.

Virtue ethicists often appeal to the notion of 'duties of virtue' (or imperfect duties) as the way in which supererogatory action can be integrated in virtue theory. But the concept of a duty of virtue is unfortunate and misleading since it confuses deontic and aretaic terms. Duties of virtue are contrasted in Kantian theory to duties of justice, but from an Aristotelian point of view, the virtuous person is equally expected to act according to the virtue of justice and according to what other virtues (such as generosity and courage) require. But beyond that, typical supererogatory action goes beyond *both* justice and virtue, beyond both perfect and imperfect duties. One may be more charitable than what the imperfect duty of charity demands and more generous than what is expected of a generous person. It may be foolish to do so or a sign of excess in one's character, but this does not undermine the value of the action as morally praiseworthy in its altruistic intention to promote the good.[22]

Finally, Jason Kawall has suggested accounting for supererogation in terms of virtuous ideal observers: 'An action is morally *supererogatory* for an agent in a given set of circumstances to the extent that fully-informed, unimpaired, virtuous observers would deem the action to be supererogatory.'[23] This attempt at explaining the distinction between supererogation and obligation is in my view highly problematic. First, it seems to confuse the normative and the meta-normative levels of explanation. On the normative level we are searching for principles for distinguishing between duty and what lies beyond duty. On the meta-level we are trying to

[22] There is rich literature (including, among others, Thomas Hill, Marcia Baron and myself) on the interpretation of Kant on the subject of supererogation but I will not enter this debate here.

[23] Kawall, 'Virtue Theory, Ideal Observers and the Supererogatory', 187 (and see 187–195). I am in agreement with Kawall's critique of virtue theories that fail to accommodate supererogation, but not with his positive proposal for an alternative explanation.

https://doi.org/10.1017/S1358246115000193 Published online by Cambridge University Press

identify those who can reliably make such distinctions. But it is difficult to see how the selection of such ideal judges of the supererogatory can be undertaken independently of the principles on which such selection is made. Now it seems that Kawall is inclined to an intuitionist view of moral judgement since he emphasizes that the ideal observer focuses on 'the actions themselves' rather than on any general criteria or principles. But when reflecting on the 'significant question' concerning the *basis* on which the judgements of ideal observers are made, he declares the question as lying beyond the scope of his discussion. However, this question cannot be just brushed aside. For *whatever* the basis for the distinction between supererogation and duty is, *that* is the source of the validity of the description of an action as supererogatory rather than the authority of the ideal observer (although of course we might have second-order reasons to sometimes rely on the way some particular individuals can be trusted for making such normative judgements more successfully than us).

The second problem in Kawall's account is that it is far from obvious that only *virtuous* observers can be ideal judges of what is supererogatory. We actually know from our experience that we can easily judge the act of throwing oneself on a hand grenade as supererogatory without having the least dispositional capacity (virtue) to do so ourselves. You don't have to be a Mother Theresa to be capable of judging her behaviour as typically supererogatory and most people who are entrusted with the job of canonizing saints in the Catholic Church are not, and will never be, saints. Even if good moral judges are usually morally decent people, the perception of the limits of duty and the detection of actions that go beyond duty may better be entrusted to the decent rather than to the saints and heroes. The latter, just because of their extraordinary virtuous character, may turn out to be bad judges of supererogation since they are typically modest and describe their own supererogatory action as duty! ('I just did what I felt I ought to do'). Kawall's suggestion conflates standards of judgement with standards of action.

So despite being sensitive to it, Kawall seems unable to remove the threat of circularity characteristic of all ideal observer theories. The derivation of the criteria for supererogation from the actual judgements of virtuous observers regarding supererogation leads Kawall to the counter-intuitive proposition that supererogation is a matter of *degree*: the more ideal observers judge an act to be supererogatory, the more supererogatory it is. And only the action which *all* ideal observers judge as supererogatory would actually be 'universally

https://doi.org/10.1017/S1358246115000193 Published online by Cambridge University Press

morally supererogatory'.[24] This conclusion is no less bizarre than deciding the truth of a scientific theory by a vote of members of the American Academy of Arts and Sciences. Kawall is aware of the 'awkwardness' of his proposal[25] but nevertheless maintains that supererogation is often a matter of degree. However, we should keep apart the normative disagreement about the boundaries of duty with the conceptual distinction between supererogation and duty. This distinction allows for no degrees in exactly the same way as the question whether an act is obligatory is not a matter of degree: it either is a duty or it is not – no matter how many people agree that it is (or is not). A woman is either pregnant or not, even when it is not easy do diagnose; and the number of professional obstetricians diagnosing her as pregnant does not make her more or less pregnant.

Indeed, failure to act supererogatorily may sometimes cause a sense of guilt or shame in the agent, but guilt is not in place is such cases since supererogatory action is defined as that action the omission of which is not morally wrong. Shame or a sense of disappointment with oneself may be rational in such omissions because they often are regarded by the agent as a failure to live up to certain moral ideals (or what is called, misleadingly, self-imposed duties). Kawall may be right that a person who never strives to attain moral ideals can be characterized as 'morally mediocre', but that does not imply any guilt in the strict sense, that which is associated with doing something that is morally wrong.

4. Denying supererogation as incompatible with virtue ethics

Rather than trying to revise virtue ethics so as to accommodate supererogatory action, a more radical solution to the tension between the deontic category of supererogation and the terms of virtue ethics is to deny the possibility of supererogation altogether. This is the line taken by Roger Crisp which will be examined in this section.[26]

Crisp, as we have already noted above, reads Aristotle as subjecting the concept of virtue to that of duty: the virtuous person will get

[24] Kawall, 189.
[25] *Ibid*.
[26] Roger Crisp, 'Supererogation and Virtue', in Mark Timmons (ed.), *Oxford Studies in Normative Ethics*, Vol. 3 (Oxford: Oxford University Press, 2013), 13–34.

https://doi.org/10.1017/S1358246115000193 Published online by Cambridge University Press

angry when it is his duty to do so. Aristotle, adds Crisp, is in that respect closer to Protestantism (which strongly denies the possibility of supererogation) than to Catholicism (in which supererogation has its historical foundations).[27] And it is not only the soldier's duty to die for his country, as Aristotle says, but also to throw himself on the hand grenade (which Aristotle does not explicitly say). There is no way to go beyond duty because 'virtue is itself an excess of a kind'.[28] Since our duty is to act according to what is 'fitting' in the circumstances and since what is fitting is just *one* way of acting, going beyond duty is going against duty and hence a wrong action, even a vice.

I am not in a position to judge whether Crisp's careful reading of Aristotle is correct although there are philosophers who read his virtue theory as allowing for *various* ways of exhibiting virtue and who, as we have seen, try to avoid the concept of duty in the way they understand virtue. It is also difficult to understand how Aristotle could accept the idea that virtue is a kind of *excess* since it does not lie in harmony with the doctrine of the mean (even when it is relativized to the particular agent and her circumstances). But Crisp's deeper point, which is quite compelling, is that for Aristotle the concept of virtue is conceptually tied to the subject of the virtue, the agent. All virtue is ultimately aimed not only at human flourishing in the abstract, but at that of the agent of the virtuous act, *her eudaimonia* (rather than the Benthamite overall good of human beings and other sentient beings).

But if we take seriously this self-centered nature of virtue ethics we get entangled in a dilemma: heroic acts of altruism, like the grenade case, are either plainly wrong (since they involve the sacrifice of the self), or a clear moral duty. Crisp suggests that such acts may be duties since they are *noble* and consequently bring glory to the agent – even if posthumous. This seems to solve the problem of supererogation, but at a very high price. All actions that we call 'supererogatory' are declared 'noble' and accordingly a duty. This

[27] Crisp mentions the three "Evangelical Counsels" (referring to them as "recommendations") of poverty, chastity and obedience as allowing for actions which go beyond the call of duty according to the Catholic doctrine of the Church Fathers. These were strongly rejected as absurd by Luther and Calvin, who argued that every religiously valuable act or way of life must be *obligatory*. And since Crisp reads Aristotle's virtue ethics as subjected to the concept of duty ("the duty to be angry"), he finds it closer to the Protestant world view than to the deontologically open-ended view of the Catholic church.

[28] Crisp, 20.

https://doi.org/10.1017/S1358246115000193 Published online by Cambridge University Press

makes virtue ethics either astonishingly demanding or excessively self-centered. But as Crisp shows, this is indeed what Aristotle seems to be claiming: the virtuous man prefers 'one great noble action to many trivial ones' and hence 'those who die for others doubtless attain this result', namely 'gaining for [themselves] nobility'. So from that point of view, altruistic moral heroism seems to ultimately be a self-serving moral duty. And it is the ultimate moral ideal, for 'if *all* were to strive towards what is noble and strain every nerve to do the noblest deeds, everything would as it should be for the common weal, and every one would secure for himself the goods that are greatest, since virtue is the greatest of goods' (1169a). This sounds indeed a noble ideal for human society but it is doubtful whether it is coherent, for if everybody was willing to sacrifice one's life for the sake of others, there would be no 'others' left (as is clear in the company of soldiers standing beside an exploding hand grenade). And the same would apply to universal charity.

To that Crisp could answer (in the name of Aristotle) that some altruistic acts are not morally valuable and hence should not be considered noble or be a duty. Generosity, for example, has limits, and what lies beyond those limits would be regarded as 'wastefulness'.[29] But Crisp is not clear about the possible spectrum of acts that lie beyond duty but are not prohibited or considered vicious, for he admits that Aristotle would (or at least could) allow for individual discretion regarding the amount donated to others by a generous person: 'She might of course choose to give more than 5% [of her income], and though this would be permitted it would not be especially morally valuable'. Let us assume that 5% is obligatory and that 75% is 'wasteful' or contrary to self-regarding considerations regarding the agent's flourishing (or duties to one's children) – how should we regard a 30% donation? I find it hard to judge such a donation as either obligatory or as lacking 'special moral value'. It is exactly what we naturally refer to as supererogatory, as not obligatory yet carrying moral value to some (significant) degree – both in the sheer amount and value for the beneficiary and in the very choice of the agent to act beyond her duty. Leaving the agent with personal autonomy in exercising her virtue necessarily leaves room for supererogation. Virtue ethics may reject such autonomy and consequently deny supererogation, as Crisp seems to be willing to do. This would not be incoherent or inconsistent, but such ethics would be, at least from our normative point of view, impoverished. Generous

[29] Crisp, 28

https://doi.org/10.1017/S1358246115000193 Published online by Cambridge University Press

acts of giving would be *either* patently obligatory *or* simply wrong and even vicious (in the sense of displaying a vice). This would make the virtue of generosity either over-demanding (in the name of the value of making the agent noble) or excessively self-centred (judging any deviation from what promotes the flourishing of the self as ignoble).

Crisp also makes the claim that, contrary to my defence of supererogation in terms of moral autonomy, a 'limited morality of duty' allows even for more autonomy than a morality of supererogation. For people are permitted to act unconstrained by moral considerations (such as being able and free to do a morally better act than they actually do).[30] But I think this is a mistaken understanding of autonomy, since choosing between two options that are morally neutral might indeed be unconstrained by moral considerations but such choice is also devoid of any moral meaning or value; whereas choosing between a morally neutral (permitted) action and a supererogatory one is of high moral value which lends autonomy a particular moral meaning.[31]

5. Metaethical methodological considerations

As is often the case in genuine philosophical problems, there is no knockout argument in the debate about the very existence of a separate category of supererogatory action. Intuition about the issue is only a partial guide. The question ultimately concerns conceptual mapping rather than moral truth, for, as we have seen, duty and 'ought' are also theoretical terms which serve in different roles in different theories like deontology and virtue ethics. The case of supererogation is interesting since it is a concept which, at least until recent times, was uniquely Christian, or rather Roman-Catholic. But since it has gained some recognition in the last fifty years, there have been attempts to test its possible incorporation in

[30] Crisp, p. 29.

[31] In an appendix to his article, Crisp criticizes my attempt in chapter 2 of my book *Supererogation* to read some passages in Aristotle as containing elements of supererogation. I accept much of that critique although it is partly based on his strict interpretation of virtue as a duty in Aristotle. As Crisp himself notes at the end of his article, I explicitly come to the conclusion that supererogation cannot find a place in virtue ethics due to its rejection of the correlativity condition, that is to say the conceptual dependence of the supererogatory on the obligatory with which it is contrasted.

https://doi.org/10.1017/S1358246115000193 Published online by Cambridge University Press

various theories of ethics – first utilitarianism and Kantianism and then in virtue ethics. Virtue ethics has proven more resistant to the idea than the other two, although there have been quite a few consequentialists and deontologists who have either rejected it completely or given it only a subsidiary role (in terms that I called 'qualified supererogationism').[32]

This paper discussed various attempts to accommodate supererogation within virtue ethics and exposed the difficulties in doing so. Roger Crisp also reaches the conclusion that virtue ethics does not leave room for supererogation and grounds his view in a strong 'duty'-based reading of Aristotle. But here we reach a deadlock: Crisp believes that supererogation is a completely superfluous concept; I tend to come to the conclusion that the strong case for supererogation serves to expose a theoretical weakness of virtue ethics. I am intentionally cautious in the way I put it, since virtue ethics is a major ethical theory, covering most moral phenomena and appealing to a wide range of our moral ideas and intuitions and hence cannot be brushed aside just because it does not easily account for supererogatory action. Supererogation, despite its centrality in our moral experience, is not a full blown theory of ethics and hence cannot in itself be considered as an ultimate test for the acceptability of virtue ethics. But still, despite this asymmetry, both Crisp's analysis and mine are a sort of *reductio*: if supererogation cannot be accounted for in virtue ethical terms, then, for Crisp, it is a reason to deny supererogation; alternatively, if this is the case, then from my point of view it is a strong argument against virtue ethics.[33]

Crisp argues that the concept of supererogation should be tested in terms of virtue ethics rather than the other way round since, he

[32] Qualified supererogationism explains the category of supererogation in terms of satisficing (in contrast to maximizing) or in terms of what one is excused for not doing due to the difficulty of the act or the imperfection of human nature, or in terms of degrees of virtue or special vs. ordinary vocation, or – as suggested by Crisp – the toleration of people who cannot perform difficult obligatory actions.

[33] A similar *reductio* argument is offered by Dreier, who distinguishes between ethical and rational satisficing, the latter being incompatible with supererogation. He consequently draws the conclusion that the idea of supererogation serves to justify ethical satisficing (and reject rational satisficing) rather than the other way round (namely, accepting rational satisficing and rejecting supererogation). James Dreier, 'Why Ethical Satisficing Makes Sense and Rational Satisficing Doesn't', in Michael Byron (ed.), *Satisficing and Maximizing* (Cambridge: Cambridge University Press, 2004), 148–9.

https://doi.org/10.1017/S1358246115000193 Published online by Cambridge University Press

claims, supererogation is a *parochial* concept (implying, of course, that virtue ethics is fundamental).[34] Its sources are in the very particular and historically situated world view of Catholic Christianity and there is no reason why we should make any major change in our ethical theory so as to accommodate it. Despite the historical truth regarding the emergence of the concept of supererogation (which, unlike many fundamental concepts in ethical theory, is indeed a technical term), there is a methodological risk in this argument from parochialism. For parochialism is a matter of degree and to some extent virtue ethics itself can be similarly regarded as 'historically contingent' rather than universal (and indeed it had mostly been abandoned in modern ethics until its partial revival two or three decades ago). But the main doubt about the argument from parochialism is that there is absolutely no valid reason not to appeal to new concepts, even to invent names for normative categories and re-map conceptual schemes, when metaphysical and normative views about the world change. If supererogation is a parochial concept, so are the concepts of human rights, autonomy and toleration – all modern ways of organizing our normative views of the world which are much more recent and even less universal than supererogation. I bring these concepts as illustration for particularly important principles to which we are committed today despite our awareness that they are relatively new and have been adopted only by a small part of human cultures. We are sensitive to the way non-Western cultures often shun these concepts as not fitting their normative views, yet we find it hard and also superfluous to make an effort to relinquish them in favour of older, seemingly more universal concepts. The same, I believe, applies to supererogation.

Crisp is suspicious of such historically contingent concepts because he advises 'parsimony in normative theory'. He urges us to 'begin by assuming an empty world [of normative concepts] and then populate it only so far as is necessary'.[35] And he believes that in such a world, the concept of 'fitting' can do much of the work, as is the case in the proverbial example of what we ought to do when we see a suffering child whose pain we can relieve at no cost. Indeed, this case does not call for any complex theory of duty, moral value, right and wrong – let alone supererogation.[36] However the example is misleading since it is a typically rare case and does not reflect real-life moral

[34] Crisp, p 17.
[35] Crisp, 29.
[36] Although one can convincingly claim that the case already assumes a certain conception of 'cost', not to speak of the 'innocence' of the young

https://doi.org/10.1017/S1358246115000193 Published online by Cambridge University Press

dilemmas and choices in which judgement is not that straightforward and requires a much richer conceptual map. Methodologically, it is absurd to start from an idealized empty world and gradually build from the very thin concepts (such as 'fitting') a whole ethical theory. For 'fitting' is already a concept saturated with thicker concepts such as the duty of someone filling a role or the universal empathy for those who suffer pain. Our moral experience is dense with thick concepts. Parsimony is a good guiding principle but conceptual maps should be rich enough for portraying the ethical terrain.

The problem of supererogation as an ethical category does not lie in its alleged parochiality but in its association with heroism. This is probably due to the religious sources of the concept in which supererogation was concerned with criteria of sainthood. Being conceived in the dramatic terms of saints and heroes, supererogatory action has become closely associated with *virtue*, that is to say with a particularly strong moral character and an admirable disposition of self-sacrifice. However, once the concept is detached from its religious foundations, the door is opened to re-articulate supererogation as an ordinary deontic category, which although referring to necessarily praiseworthy action does often describe non-heroic, common actions such as making small gifts, making promises, forgiving minor offences, or lending assistance to friends or strangers.[37]

Thirty years after writing 'On Saints and Heroes', Urmson, in a critique of R. M. Hare, expresses 'regret' for having used the term 'supererogation' and for having suggested a 'new over-simple tetrachotomy' (replacing the traditional, deontic logical trichotomy).[38] His reason for this striking change of mind is that the adoption of supererogation in our conceptual map 'is to ignore the vast array of actions, having moral significance, which frequently are performed by persons who are far from being moral saints or heroes but which are neither duties nor obligations'. And here Urmson mentions examples such as acts of kindness, considerateness, neighbourliness

---

child, including its exemption from responsibility and other concepts implied in the allegedly irreducible concept of 'fitting'.

[37] Crisp himself notes the difference between 'praiseworthy' as deserving praise and 'praiseworthy' as creating a reason for praising (16). Supererogatory action is necessarily praiseworthy in the former sense but only occasionally and contingently in the second.

[38] J. O. Urmson, 'Intuitive Moral Thinking', in Douglas Seanor and Nick Fotion (eds.), *Hare and Critics: Essays on "Moral Thinking"'* (Oxford: Clarendon Press, 1988), 167–9. I am grateful to Marcia Baron for having turned my attention to this article.

https://doi.org/10.1017/S1358246115000193 Published online by Cambridge University Press

and common everyday gestures of help. Now that is exactly the point I am trying to make regarding the dissociation of supererogation from its traditional Christian connotations of sainthood and heroism, and in that respect I see no reason for Urmson to confess any philosophical regret for his earlier groundbreaking contribution to moral theory. For some reason, though, Urmson resists in his later reflection on the subject the generalization of his examples of non-heroic non-obligatory yet morally significant actions under the theoretical title or concept of supererogation. His resistance has to do with the suspicion that these acts are too heterogeneous and share nothing in common which can merit the special category of the supererogatory. Yet, despite Urmson's claim that this heterogeneity reflects 'the rich complexity of moral life', one cannot avoid the fact that according to Urmson himself all these acts do share two important properties, namely being non-obligatory yet morally significant. From a theoretical point of view of the attempt to draw a map of moral concepts it is hard to deny this property as either trivial, or too abstract or otherwise useless. Even if Urmson is correct in calling for a much more fine-tuned conceptual scheme for our moral language and for much thicker concepts (which, by the way, would go even farther in violating Crisp's parsimony principle discussed above), the thinner concept of supererogation would not be made superfluous. There is something in common between helping a lady across the street, lending a bag of flour to a neighbour and throwing oneself on a hand grenade (all Urmson's examples). Even if it has nothing to do with virtue or special praise accorded to the agent, the common feature lies in their deontic status, in their optionality, agent discretion, and non-universalizable nature (what Urmson refers to as not being subjected to any Hare-like 'principle'). I therefore conclude that Urmson's regret should better be ignored and his original conceptual mapping followed and developed.

Thus, back to the relation between supererogation and virtue, there is nothing particularly 'virtuous' in minor supererogatory action, on the one hand, and there are actions that are 'merely' obligatory that often reflect a particular virtue and justify special praise, on the other. The aretaic judgement should better be kept apart from the deontic status of the judged action, as should the assessment of the value of the consequences of action be from its deontic status.[39]

[39] This is not to deny that in the overall evaluation of a person's life and character we do take into account a *systematic* avoidance of any supererogatory action as a moral defect of virtue, even if not a blameworthy one. See also Nancy Sherman, 'Commonsense and Uncommon Virtue', *Midwest*

https://doi.org/10.1017/S1358246115000193 Published online by Cambridge University Press

Such less parsimonious mapping of ethical concepts does more justice to the fundamental moral intuitions and experience which even if not universal characterizes the way we make moral judgements – at least in our own times and culture.

When I speak of 'our culture' I refer to a world view which is not only different from Aristotle's but equally from that of the Church Fathers who coined the term 'supererogation'. It is a pluralistic world in which we have lost consensus about ethical role models which everyone should strive to imitate. It is accordingly a world in which ethical values cannot be imposed on people unless there is a special reason for doing so, such as securing social cooperation and cohesion, decency and justice. It is also a world in which the right and autonomy of individuals are taken as prior to what is best for all. Society and individuals are expected to tolerate those who strive to lead a life which is not evaluated as good as long as it does not directly harm others. We appreciate the personal choice of individuals beyond its 'objective' value and try to raise children who will ultimately choose their own life plans. These are all background moral beliefs that directly or less directly support the need and fruitfulness of a category of supererogation, a category which leaves room for objective judgements about the moral good but also respects individual personal autonomy.

In that respect, Urmson's revival of the old and unfashionable concept of supererogatory action is a reflection of a *liberal* view in moral life which is an extension of that view in the political sphere in which the right to do wrong (or at least not the maximal good) has been prevalent since the seventeenth century. To virtue ethics such a view does not make sense and hence supererogation cannot be recognized by it, or properly explained in aretaic terms.

The Christian doctrine of supererogation was given a death blow in the times of the Reformation, partly because of its abuse by the Church in the corruptive practice of Indulgences and partly because it was incompatible with the Protestant view of divine grace. The modern, secularized idea of supererogation does not suffer from this bad reputation, and its theological connotations are nowadays limited to the story of the Good Samaritan, which itself is currently understood in a secularized form (in legal language or

---

Studies in Philosophy 13 (1988), 97–114. I would add that in the same way as we recoil from moral saints, so do we feel moral unease with people who fanatically confine their moral choices to their duties and rights, never deviating from them in a supererogatory way.

https://doi.org/10.1017/S1358246115000193 Published online by Cambridge University Press

in social psychology). It can thus serve as a significant challenge to current moral theories, including virtue ethics, and at least as a partial but serious test for their cogency.

The Hebrew University
david.heyd@mail.huji.ac.il

https://doi.org/10.1017/S1358246115000193 Published online by Cambridge University Press