

Assimilating Supererogation

D.K. LEVY

1. The interest in supererogation and supererogatory actions derives from the perception that there is something problematic about them. I shall argue that there is nothing problematic about them. The perception to the contrary arises from preconceptions common in ethical theory. When these are relaxed or dismissed, supererogatory actions are easily assimilated as well-motivated responses to moral situations. Assimilating, rather than denying, them is important for a sound moral philosophy.

The idea of the supererogatory predates Urmson's well-known article.¹ However, I shall treat Urmson's discussion as foundational in what follows. Supererogatory actions, I shall say, are actions that are morally good but not required by duty nor obligation. Specifically, a consequence of supererogatory actions' not being required is that one cannot be blamed morally for failing to do what is supererogatory. Urmson notes that supererogatory actions are not readily assimilated to ethical theories, because they do not seem to fall under a traditional division into three kinds of action, viz. obligatory; permissible but not obligatory; and impermissible.² The supposed problems arising from this incompatibility are the subjects of sections §5–6. I shall argue that these problems and the puzzles regarding the supererogatory can be bypassed without denying the phenomenon that gave rise to the category (§6–7). Before addressing these I give two examples of supererogatory actions, varying by the immediacy of their responses. In §2, the way in which supererogatory actions appear as necessary is elaborated prior to describing (§3) this appearance in terms of holding someone responsible for actions. Refining this characterisation (§4), a supererogatory action is described as one in which solely the actor can hold himself responsible for the action.

Consider two examples of supererogatory actions. The first is from Urmson. "We may imagine a squad of soldiers to be practicing the throwing of live hand grenades; a grenade slips from the hand of one of them and rolls on the ground near the squad; one of them

¹ J. Urmson, "Saints and Heroes," in *Essays in Moral Philosophy*, A. Melden (ed.), Seattle: University of Washington Press, 1958.

² Urmson, 198–9.

sacrifices his life by throwing himself on the grenade and protecting his comrades with his own body.”³ It is notable that it is not essential to this example that the soldier actually sacrifice his life. For if the grenade is a dud or fails to kill, that action is no less exemplary. For this reason, we can consider what the soldier might say after he has thrown himself on the grenade. I note also that this example is one in which there is no time for reflection and any deliberation is of necessity immediate.

The second example is one in which the supererogatory action arrives after considered reflection and deliberation. Imagine someone who has suffered a grievous wrong who considers whether to forgive the wrongdoer. Suppose after much consideration, reflection and deliberation over a period of weeks or longer, she decides to forgive the wrong and does so in a complete and sincere way. We can imagine many considerations that might weigh on her and many variations in what occurs during deliberation, e.g. the wrongdoer apologises, seeks forgiveness, shows remorse, etc. No doubt any of these may explain a forgiving, but none is an essential condition on so doing. It is widely held – though not undisputed – that forgiveness is neither a duty nor obligatory, no matter the circumstances nor the reparations proffered by the wrongdoer. It is also widely held that forgiveness can be morally good when done for the sake of the forgiven. I suggest therefore that forgiveness too, when the circumstances are right, is supererogatory.⁴

2. Both of these examples have in common that each person is motivated by the thought that what they do is good and neither could be blamed morally for failing to do what each does. This way of characterising the examples risks mistaking the import of the motivations for acting. For when we accept that neither could be blamed morally by someone else for failing to act, we might infer that the actions appeared to the actors as *optional* because blameless whether done or not. This sense of ‘optional’ is not self-evident. It includes the sense in which the action is dispensable and could be skipped. It also includes senses similar to preference or whim, like the choice of one of two equidistant routes to a destination or Béarnaise sauce rather than hollandaise. However, none of these senses of ‘optional’ is apt for supererogatory actions. Supererogatory actions are motivated by the thought that their performance is good. If they were not, they would not be morally noteworthy.

³ Urmson, 202.

⁴ For a relevant defense of this idea, see Espen Gamlund, “Supererogatory Forgiveness,” *Inquiry* 53 (6):540–564, 2010.

Assimilating Supererogation

Insofar as an action is recognised as good by a person, it is not dispensable to her; neither is it the expression of whimsy or a mere preference of the moment. The soldier could not credibly describe his action as a whim. The extent to which forgiveness is given on a basis similar to choosing a food's sauce is the extent to which it is not genuine. So the sense in which a supererogatory action is optional should be understood to reflect how *others* are restricted from asserting that the supererogatory action is obligatory or (morally) blaming the actor for non-performance.

Notwithstanding restrictions on others, Urmson notes that the supererogatory action *appears* to the actor as necessary, as something he is obliged to do.⁵ He responds to the situation under this modality, the appearance of necessity, perhaps saying, "I did what I had to do." For the soldier, aware of the likelihood of death, the protection of his comrades must have presented itself as what he had to do, by diving on the grenade. This might be variously expressed as the best thing to do, the right thing to do, the good thing to do, and so on. He might modestly and mistakenly say that anyone would have done the same or that it was his duty. The individual expressions of an actions' appearing necessary are many, but I will describe its appearing necessary as *consenting to his responsibility* for acting.

Something similar can, I suggest, be said for the person who arrives at the willingness or decision to forgive. She might express her willingness or decision in many ways, e.g. as its being the right thing to do, the best thing to do, the best outcome, a good thing to do, etc.⁶ No doubt, while deliberating or reflecting on whether to forgive, forgiving may not have appeared necessary. Once it does appear necessary, as the person who can forgive, who is willing to forgive, the forgiving appears as something to which she is now obliged. She might express this modality by saying, "I could not live with myself if I did not..." Indeed, she may now think it was always necessary, though she had not yet realised it. We need not settle here the question of whether so thinking is a mistake. This idea is similar to Urmson's observation that St. Francis upon recognising his duty to all of God's creatures including birds, reproached himself for having not recognised his

⁵ Urmson, 203.

⁶ There are of course many motivations for saying that one forgives. Many counterfeit an act of forgiveness, e.g. making it mere reconciliation or self-serving. This example depends on the idea that forgiving is a moral act motivated by the good of forgiving. If this is denied, an alternate example of a reflective, deliberative supererogatory act will serve.

duty earlier.⁷ Again, I should like to describe the appearance of an act of forgiving as necessary as someone's consenting to her responsibility for forgiving.

3. We have arrived at a refined elaboration of supererogatory acts. These acts are morally good but not required by duty or obligation, where this is understood to mean that no one could be blamed by others for not undertaking the acts. The acts are motivated by the actor's recognition that they are good. This same recognition is part of the act's appearing to the actor as necessary. This appearing I have described as consenting to a responsibility for the act. The point of this turn of phrase is that it allows me to describe supererogatory acts as those in which the subject who acts can hold himself responsible, when no one else could. This asymmetry in who can hold whom responsible is essential to supererogatory acts, while focusing on responsibility is my description of this essential character.

As I will discuss below, the supposed problematic aspects of supererogatory actions arise from theoretical claims about the nature of action generally. My focus on responsibility bypasses these claims by offering an alternative view of action in which actions are individuated by who can be held responsible by whom. By making the idea of "holding responsible" foundational in moral philosophy with regard to action, among other things I show that we can assimilate supererogatory actions with no difficulty. By 'foundational', I mean, for example, that "holding responsible" is not answerable to an account of responsibility that is causal, in the sense of "this caused that." By 'not answerable', I mean at least that divisions in responsibility between not-responsible and responsible in a causal account could differ from those divisions in a holding responsible account, without one account having priority over the other. Similarly, being foundational, a holding responsible account is not answerable to a metaphysical theory of action (i.e. an account of the ontology of action). That said, logical considerations from a metaphysical theory of action could still apply as one basis for holding someone responsible, as I indicate below when speaking of the variety of bases for holding responsible.

'Responsible' in the sense of 'responsible for' is a notoriously polyvalent word in philosophy – similar to *αἰτία*⁸ – whose senses span the philosophical distinctions in question here. I call my account

⁷ Urmson 203–4.

⁸ Michael Frede, 'The Original Notion of Cause', in: *Essays in Ancient Philosophy* (Minneapolis: University of Minnesota Press, 1987), pp.125–150.

“holding responsible” in part to distinguish it from familiar senses of ‘responsible’. My primary motive though is to emphasise that it is people who hold other people responsible. In holding someone responsible, someone is not simply reporting a fact about the aetiology of an event. Rather he is demanding attention to that activity, that it be given a regard consistent with the actor’s responsibility for it.

A crucial aspect of holding responsible is the variety of bases for doing so.⁹ For this reason, the plural, “holdings responsible,” is apt for understanding action on this approach, because someone may be held responsible on differing bases by different people. When I speak of the varieties of holdings bases responsible, I have in mind at least two divisions. First, the norms governing holding(s) responsible vary by domain. For example, the norms for holding responsible vary in the law, e.g. between civil and criminal law; or between *mens rea* and *actus reus*. Similarly, norms or standards for holding responsible vary in professional ethics, morality, sports, etc. The effect is that one may be held responsible by one standard and not another. The practical import is that every holding responsible is based on a norm for doing so, what we could call part of its normative basis. It is this basis that governs the correctness of a holding responsible, namely whether a person is correctly or rightly held responsible for an action.

The second division in holdings responsible concerns who may hold (whom) responsible. Specifically, not anyone can hold another person responsible. They must be in the right relation to them. For example, someone who has promised can be held responsible by the person to whom the promise was made. A judge can hold someone responsible to the law, when in court. A collective can hold someone responsible in ways others could not. For example, fellow Americans can hold a countryman responsible as Britons could not; or members of an orchestra could hold a fellow responsible in ways others could not. The effect is that one could be held responsible by one person, while someone else is unable to hold the same person responsible for the (seemingly) same action. The practical import is that every holding responsible depends on the propriety of someone holding another responsible on the basis of a relation between them. The propriety of that relation’s making it licit for

⁹ The variety of bases for holding responsible is similar in inspiration to the idea of the variety of bases by which one might be excused responsibility, which are canvassed in J. L. Austin, ‘A Plea for Excuses’, in: J. O. Urmson and G. J. Warnock, editors, *Philosophical Papers* (Oxford: Clarendon Press, 1961), pp. 123–152.

one to hold the other responsible is the other part of the normative basis for holding responsible. Again, it is this basis, *in toto*, that governs the correctness of holding responsible.

The *practice* of holding each other responsible for our actions is familiar and uncontentious. The claim that the basis by which we do so depends on varying norms and the relations between those holding and those held responsible is also not revisionary of common (or philosophical) sense. It is less obvious though that this approach to individuating actions could be foundational in the sense I have specified. However, I suggest that reflection on the idea of individuating actions – as in a theory of action – reveals that there is nothing radical or revolutionary in proposing this individuation as foundational. The foundation of every account of action is the distinguishing of bodily movements from actions – something parallel might be said for “movements” of the mind as mental actions.¹⁰ It is a philosophical commonplace that not every bodily movement is an action. For example, none of autonomic movements, movements by reflex, or movements from an external cause would ordinarily be considered actions. Moreover, for every bodily movement, there is not solely one action with which it is identical, because one movement can be several actions, e.g. when the same movement is waving off and taunting. Reflection on grammar also shows that ‘holding responsible’ is a suitable basis for distinguishing bodily movements from actions. For any bodily movement, it makes sense to ask whether someone is responsible for it, for there is someone whose bodily movement it is. Of course, the same bodily movement may prove to be identical with different actions under different standards for holding responsible. By the standard of voluntariness, we may hold someone responsible by asserting the action (i.e. the movements) was voluntary. By a standard of culpability, we may not hold someone responsible for that action because it was unintentional. Someone who misunderstands that a switch is not for the lights, but for the oven, voluntarily flips the switch, but does not intentionally switch off the oven. None of these observations about individuating actions is meant as a revelation. I have highlighted them to underwrite my taking it as unproblematic that we can use the idea of ‘holding responsible’ as a foundation for individuating actions.

4. The focal case in our discussion is a supererogatory action, for which we can now give a good account, i.e. one that is faithful to how the action appears to actor and spectator alike. Central to a

¹⁰ Cf. Peter Geach, *Mental Acts*, Routledge and Kegan Paul, 1957.

Assimilating Supererogation

supererogatory action is its being one in which, for good and intelligible reasons, someone can hold themselves responsible when no one else could. Holding themselves responsible depends on consenting to responsibility for the action, which I have suggested is of a piece with the action coming to appear necessary, because good. However, the responsibility must originate in the actor – what I have glossed as consent – since it could not originate elsewhere because no one else could hold him responsible. This follows from the blameless nature of a supererogatory action from the perspective of the spectator.

The question of blame, specifically moral blame, is apt because for actor and spectator alike the supererogatory action is one that is good, that is undertaken because it is good. Again, the asymmetry of supererogatory actions shows itself, this time in the area of performance and non-performance. Specifically, as spectators we cannot hold someone responsible for the non-performance or failure of a supererogatory action. By contrast, we can hold someone responsible for her performance of a supererogatory action by the moral praise we think appropriate.

Consideration of the case of failure is more illuminating, because the failed actor can hold himself responsible. A soldier who did not jump on the grenade could hold himself responsible for not doing so, though no one else could. He might do so with many of the hallmarks of holding himself morally responsible such as shame, remorse, atonement, etc. His explanations for his non-performance might be those familiar to moral cases of failure, e.g. failure of nerve or cowardice. No doubt he may be mistaken in some aspects of his holding himself responsible. Holding responsible is not self-authenticating even in the reflexive case. Norms of intelligibility, especially from the moral domain, will be relevant – as will the normative basis for holding responsible described above. For example, if the grenade were too far away for him to have jumped on it, his holding himself responsible would be unsound. However, at the limit, when the norms are not traduced, we should defer to the moral authority of the subject to describe how he has failed morally. It is not decisive evidence of misunderstanding or mistake that someone is not consoled by the truth of the assertion that no one else does or could hold the failed actor responsible. I return to personal moral authority in §7.

I have elaborated how supererogatory actions can be assimilated to a 'holding responsible' account of action to show that they are unproblematic. For all the considerations I have given above about how one may hold oneself responsible morally are readily seen to parallel the ordinary case in which one person holds another responsible

morally for their actions. There is at least this difference with the ordinary case: no one is *wronged* by the non-performance of a supererogatory action. This difference is not however so distant from familiar deviations from the ordinary case. For example, when one fails in an imperfect duty, no one is wronged, but the failure is a moral failure, a likely symptom of vice. There is also the difference that a successful supererogatory action always does good that could not have been demanded by others. This difference is not a concern, as I will argue below (§6).

5. The 'holding responsible' approach has a dual role in assimilating supererogatory actions. First, it supplants a picture of the nature of action. Second, it secures the normative considerations associated with blame, which are allowed to be asymmetric. I move now to describe why supererogatory actions have seemed puzzling. In §5 I locate the difficulties in a prevailing picture of action. In §6, I connect these with difficulties in making focal obligation as the basis for the moral worth of an action, and thus of moral blame. In both cases, I will show that the problematic character of supererogatory actions derives from pre-conceptions about action and moral worth.

Supererogatory actions are puzzling because it seems odd that someone could have motives for acting that when acted upon confer merit, but when ignored do not confer demerit. How could motives operate in such a one-sided or asymmetric way? More or less structurally analogous to this puzzle, now with a moral aspect, is the idea that there could be a good at which an action aimed that—even when recognised as good—could yet be disregarded without moral demerit.

I suggest that the central confusion that gives rise to this puzzle about supererogation occurs because a causal picture of action has infected a normative explanation of action. The pre-conceived causal picture of action is oriented around the idea of motive forces combining to push a body into motion, rather like a mechanism moves when a sufficient accumulation of pneumatic pressure overpowers the mechanism's inertia. On this view, we are moved into action by that behind which the most motive force has accumulated. Each motive is conceived as having force oriented to an end, like a vector with a magnitude and direction. Roughly, one is moved to the end behind which is the strongest motive, i.e. for which there is the highest vector (perturbed perhaps by other vectors). The *form* of this explanation is applicable to physical, psychological, and normative explanations. They differ in the nature of the forces attributed. In one they are physical. In another they are psychological, e.g. pathological like fear. In the third they are normative, e.g. rational like a reason.

Assimilating Supererogation

With this picture in mind, the puzzle raised by supererogation is how one could be moved to do that for which one does not have the most (e.g. normative) force? Following the picture, we suppose that if a subject acted in a supererogatory act he must have had the most (normative) force (e.g. reasons) behind doing so. *Yet* had he *not* acted he would not have been irrational or inexplicable (because he would be blameless), which is tantamount to saying that he had not had the most force behind that end after all. On the causal model, that cannot be. It is as if he had the most force (*causes*) to act, but when he does not act (i.e. the act does not happen), we need not revise our description of his situation (i.e. of the causes in play). That seems crazy in a case of physical causation and troubling in a normative case.

The puzzle can be inverted. If the actor will be blameless (i.e. not irrational or inexplicable) when he does *not* act, we can infer on the causal picture that he must *not* have the greatest accumulation of motives (force) behind that action. If that is so, when he makes the supererogatory act the puzzle is how can someone have been moved to act toward that for which he had not had the most motive (force, cause, reason)? Beguiled by the causal picture, we are troubled when we seem forced to say that the soldier would have been blameless because he had the most reason *not* to dive on the grenade; but did so anyway; his so doing was therefore irrational; but is morally estimable nonetheless for having done so. The puzzle might be described with more finesse so as to irritate our sensibilities less, but this essentially problematic character will not be obscured.

Wearing the clothes of ethical theory, the supererogatory again confounds. For an ethical rationalist, the difficulties are parallel to those sketched above, viz. that the supererogatory is irrational, yet morally estimable. That conclusion is difficult to endorse, for how could what is irrational also be estimable, morally or otherwise? The earlier form of the puzzle, in which supererogation was inexplicable because motivationally deviant or insufficient, also raises the concern that there are more reasons than are decisive or that reasons do not apply uniformly as motives. In this case, an accommodation would have to be found by distinguishing types of reasons; or distinguishing reasons that explain (non-moral) from those that justify or confer merit (moral).¹¹ This is unappealing because it complicates

¹¹ Raz attempts the first accommodation while Portmore attempts the second. J. Raz, "Permissions and Supererogation", *American Philosophical Quarterly*, 12: 161–168, 1975; D. Portmore, "Are Moral Reasons Morally Overriding?", *Ethical Theory and Moral Practice*, 11: 369–388, 2008.

an idea of reasons as causes in a univocal sense, each a force applied to the mechanism. For a generic consequentialist, one is blameless if one does that which produces the best consequences. If an actor would be blameless if he did not perform his supererogatory act, his non-performance seems to produce the best consequences. It is puzzling then that the performance of his supererogatory act could be of moral merit, since *prima facie* it does not contribute to the best consequences. It is as if the consequentialist must allow that there are consequences that are better still than the best and it is to these that the supererogatory act contributes.¹² However ‘better than best’ is on the face of it a logical and grammatical nonsense, so an accommodation is needed for consequentialism to assimilate supererogation.¹³

6. I have sought to show that one source of the problematic character of supererogation arises from a “pneumatic” picture of action modeled on causal forces. The problematic character shows itself in an unsatisfactory explanation of motives for action that are acknowledged when acted upon and disregarded when not. The problematic character varies its guise by ethical theory, but the problem remains essentially motivational, in part owing to the blameless character of the supererogatory. A response then is to challenge the blameless status of supererogatory actions, by insisting that all good actions are obligatory.¹⁴ Deontic ethical theorists might aim for a similar result by suggesting that all supererogatory actions are covertly the discharge of a duty or obligation. These approaches have in common that they deny the phenomena of supererogatory actions. Specifically, they deny the propriety of the apparent inability or unwillingness to state in advance that someone has a moral duty or obligation to act (e.g. jump on the grenade or forgive), nor to hold someone morally accountable for the act’s non-performance. Instead, in every case there is some obligation or duty to which we could advert if only it were understood which applied. Thus, supererogatory actions are only seemingly morally good yet blamelessly disregarded.

¹² A parallel version of this complication for value-monistic consequentialism could be constructed using value that seems to be (intermittently) more valuable than ordinary value, i.e. “especially valuable.”

¹³ Urmson makes a tentative suggestion of how to accommodate supererogation in a utilitarian theory, pp. 208–15; as does Portmore in “Position-Relative Consequentialism, Agent-Centered Options, and Supererogation,” *Ethics*, Vol. 113, No. 2, 2003 pp. 303–332.

¹⁴ See Susan Hale, “Against Supererogation,” *American Philosophical Quarterly*, Vol. 28, No. 4 (Oct., 1991), pp. 273–285.

The status of moral phenomena is certainly open to dispute. Its being so is one of the sources of the difficulty of moral philosophy. Indeed, one could think that clarity about moral phenomena was the subject of moral philosophy. Therefore, when a putatively moral phenomenon is denied, the dialectical import is not easily established. Rather than insist on the moral phenomena in the face of these denials, I would ask after the motive for the denials. What motive is frustrated by admitting supererogatory actions? I suggest the underlying motive is a desire to preserve obligation as the focal *moral* concept. I call it 'focal' because it is the root of moral worth, i.e. the goodness of actions. Using it, we can specify the familiar trio of deontic categories, viz. actions that are obligatory to do, obligatory not to do, not obligatory either way. The last of these is of no moral worth for a deontic ethical theorist, since it does not discharge an obligation (e.g. an incurred obligation like a promise; or one not incurred like a duty). Supererogatory actions frustrate this moral exaltation of the obligatory by being morally good actions that do not derive from obligations, for the actions appear not obligatory in any way. If there are supererogatory actions, then there must be *more* sources of the goodness or moral worth of an action than obligation. Countenancing more sources is the problem supererogatory actions pose for deontic moral theorists, which problem motivates denying the existence of such actions.

The deontic theorist does not deny that supererogatory actions are morally good, she only denies their apparent deontic status, i.e. the source of the goodness. Someone can insist, following Kant, that it is self-evident that the idea of morally worthy action derives solely from duty – a special case of obligation, itself derived from the idea of a moral law.¹⁵ Dialectically, this insistence functions as a pre-conception. It is a pre-conception, I suggest, that presents a distorted picture of morality and moral responses. The considerations I give below are reasons to reconsider whether the pre-conception is sound. If the pre-conception were given up, the supererogatory would no longer prove problematic in this regard.

Without doubt, the idea that there are actions that are better to do irrespective of inclinations to the contrary lies close to the source of the very idea of morality. One can accept that idea while resisting the idea that the motive of all actions contrary to inclination arises from the recognition of an obligation. Put another way, we can recognise more motivations and modalities than obligation that are sufficient

¹⁵ Kant, *Groundwork of the Metaphysics of Morals*, 4:389.

to turn aside inclination.¹⁶ The unwillingness to admit alternative modalities whose recognition motivate our responses forecloses the variety of moral responses. The foreclosure produces a shriveled, foreshortened picture of morality.

It is obvious that the supererogatory, if it were admitted unabridged, is an example in which a subject acts for the good, because it appears necessary, without recognising any obligation. If admitted, this is a motivating modality other than obligation. Another example is a response motivated by love that has no essential dependence on the idea of obligation (unless it is supposed all love must be obliging, which is question-begging). Nussbaum develops an example she characterises as a moral challenge faced by father and daughter in which neither acts from obligation, but each acts to preserve the love between them.¹⁷ Some morally fine responses preclude acting to discharge an obligation because they are non-deliberative and non-teleological, e.g. responding with pity. Insofar as the responses are non-deliberative—i.e. because a response is immediate—discharging an obligation does not enter as a motive. Insofar as the responses are non-teleological, they do not have ends, including discharging an obligation. Consider for example the person brought low by catastrophic hubris or maimed by random evil, like terrorism. Pity in response to these examples can be immediate on recognising the suffering of another person and it serves no end but the recognition of that suffering. In these cases of pity, the moral aspect of the response depends not on brute sympathy, but in the recognition that what is suffered arises from an antagonism of good, i.e. vice and evil.

More generally, as these examples intimate, we should allow that the value or *allure* of the good is sometimes modal (i.e. appears as necessary) and motivating but not as an obligation or a reason to be weighed, but in the recognition that it is good. The moral worth of an action can arise, I suggest, by its being oriented to or motivated by the good. The supererogatory is clearly this and our esteem for it flows from this recognition. So while the idea of action contrary to inclination – sometimes rightly described as motivated by obligation – lies close to the source of the very idea of morality, equally close to that source must lie the good. Indeed, one could think that it was good (and evil) that was the source of the very idea of morality. This is

¹⁶ A similar claim is made in M. Stocker, “The Schizophrenia of Modern Ethical Theories,” *Journal of Philosophy* 73 (14):453–466, 1976.

¹⁷ Martha Nussbaum, “‘Finely Aware and Richly Responsible’: Literature and the Moral Imagination” in *Love’s Knowledge*, Oxford University Press, 1990.

no more radical than the idea that moral philosophy must begin by establishing the *summum bonum*.

I do not imagine to have canvassed the full range of motivations or modalities the acting under which confers moral worth on an action. The variety is many and of necessity cross-classified in various ways. Some will appear under the aspect of the voluntary, the involuntary, and the modal. Others will appear under an aspect of the purposive/teleological, some responsive, some by allure or attraction. We already recognise moral motivations that are not obligatory in an immediate sense by imperfect duties, but to this we could add doing good, being good, evincing good and others. In so doing, we would re-balance moral philosophy from an excess concern with avoiding wrong-doing rather than doing good. By setting out these sources alternative to obligation for the moral worth of an action, I have sought to undermine the dialectical probity of insisting that a single deontic concept – obligation – should be focal in moral philosophy. If the preconception that it is so can be dropped, supererogatory actions will not prove problematic in this area.

7. As advertised at the outset of §5, I have shown that supererogatory actions appear problematic because of two pre-conceptions: a causal account of action; and obligation as the focal concept in morality and thus moral worth. I could not hope to have overthrown pre-conceptions as well-established as these by the short considerations I adduced above. My dialectical goal was to vindicate supererogatory actions as ones that can and should be assimilated in moral philosophy by drawing out the pre-conceptions that create their problematic appearance. The considerations advanced weaken the pre-conceptions, thereby bolstering the account of supererogatory actions I developed in §§2–4.

There is a further lesson for moral philosophy that I believe can be learned from this consideration of the phenomenon of the supererogatory. Specifically, supererogatory actions – also heroes, saints, and their intermittent like – force a confrontation with a *personal* dimension in morality. The resistance to the supererogatory stems from the urge to deny a personal dimension in morality in favour of the universal. At the core of the phenomenon of a supererogatory action is a person's consenting to responsibility for an action. (That is one way to describe it among many – I do not insist on my description, only the idea.) In consenting to responsibility, the supererogatory actor gains the motivation to act by, we could say, placing the burden to do so on himself. This burden cannot be placed by anyone else. Insofar as the burden is placed on himself as a moral burden, it is done on his own (personal) moral authority. No one

else has the authority to do so. Therefore, in part, the challenge of assimilating supererogatory actions is accepting the elevation of an individual's moral authority to where he can determine his moral perspective and which actions will be necessary within it.¹⁸

By contrast, the resistance to the supererogatory stems from wishing to ensure that morality applies *universally*, because this secures for morality a supreme authority. By 'supreme authority', I mean an authority that surpasses all others, including personal moral authority, however conceived. Duty is a paradigm of universal moral authority. By trying to make an action into the discharge of a duty, we are trying to make it into something that *anyone* can demand from someone, i.e. hold responsible. Duty, as it were, has a universal authority that anyone can use. By trying to impugn the rationality of supererogation, one is trying to restore the universality of reasons to which *anyone* should conform, i.e. be held responsible. Rationality too has, as it were, a kind of universal authority. By asserting that the good is always obligatory, one is insisting that *anyone* is obliged to pursue it. It is as if we are concerned to keep apart the moral worth of an action from the moral character of the actor, so that *anyone*, irrespective of moral character, can be held responsible for the act. A supererogatory act undermines this separation by mixing the personal authority of the actor – as a proxy for character – with the act to give it moral worth. The tension between the personal nature of morality, of which moral authority is one manifestation, and the universal nature of morality is recurrent in moral philosophy. The question of how to assimilate supererogatory acts puts the tension into focus.

I suggest the desire for a supreme authority expresses a lack of confidence. This is most evident in the pre-conceptions about action on a causal model. The confidence lacking is in taking it upon oneself to hold someone responsible, i.e. holding someone responsible on one's own authority. Instead, we seek the authority of a causal (or metaphysical) account of action in which it is presented as fact whether someone is responsible. Fact is the basis for holding someone responsible instead of the exercise of one's own authority in determining which is an action, under which description, with which responsibilities, etc. Borrowing the authority of fact, and its foundations, avoids testing our confidence in the use of one's own moral authority. A parallel case can be made for obligations. Incurred obligations are incurred

¹⁸ This contest could perhaps be re-staged around obligation and the question whether individual moral authority is ever sufficient to produce an obligation with the same modality as duty, i.e. an obligation not incurred.

on the basis of personal authority, so the tension does not arise unless it is asked, e.g. why promises made must be kept? However, obligations that are not incurred, such as duties, require a supreme authority if they are to be pursued against someone. For on whose *personal* authority could someone pursue an obligation not incurred by a subject against that subject? As Anscombe noted, this pre-conception about the supreme authority of obligation is a residue from a past time when supreme authority emanated from a supreme being.¹⁹

Once the desire for universal authority is seen for what it is, viz. a lack of confidence in one's own moral authority, the appeal of the pre-conceptions will lessen. Indeed, once we see holding responsible for what it is, viz. the exercise of one's moral authority, engaging in the practice will tend to increase our confidence.

8. My point in highlighting the tension concerning authority and assimilating supererogatory acts overall is to affirm the validity of a demand we place on ourselves, sometimes. That is, sometimes, we seek recognition and deference for our moral perspective, in which we have consented to responsibility for an act that could not have been demanded of us by anyone else. This is an exercise of personal moral authority. Notwithstanding my claim that a supererogatory act is an expression of personal moral authority, the scope of my claim is limited. Holding oneself responsible is not self-authenticating, as I argued. The scope of the exercise of personal moral authority is limited to the self, i.e. when I consent to responsibility for the act I do not create an obligation for anyone else, nor make them responsible. There is nothing problematic about *explaining* supererogation if we give up preconceptions about ethical theory. These preconceptions arise not from the phenomena of ethics, but from the goal of ethical theorising, viz. explaining the springs of action and endorsing some, but not others, as correct (in deontic or rational senses). That is, ethical theory aims to make systematic the characterising of actions. My approach implies that actions are not entities amenable to a complete systematisation, but have their being by what we can distinguish – something I argued was the first step in any theory of action, viz. distinguishing action from bodily movement. A supererogatory action is unproblematic to explain with regard to the motive that moves the actor, because the motive is integral to consenting to responsibility for the act. It is also unproblematic to rationalise because while the actor does not oblige herself to act, she acts in

¹⁹ E. Anscombe, "Modern Moral Philosophy," *Philosophy*, 33, 1958, pp. 1–19.

D.K. Levy

pursuit of the good whose recognition moved her to consent to responsibility in the first place.

In summary, there is nothing problematic about supererogatory actions that requires a special deontic category or motivation if we (i) are free of a causal picture of action; (ii) accept asymmetries in who may hold whom responsible; (iii) break the conceptual interdependence between obligation and morally praiseworthy action. In particular, the value of this approach to assimilating supererogation is that it permits an improved place in moral philosophy for goodness and the motivation to do good, rather than a present bias toward deontic control and the avoidance of wrong-doing. When the motivation to do good is on par with avoiding wrong-doing, the supererogatory will remain impressive and praiseworthy but should cease to be puzzling to ethical theorists.

University of Edinburgh
david.levy@ed.ac.uk