

Supererogation and The Intentions of The Agent

Abstract: It has been claimed, by David Heyd, that in order for an act to count as supererogatory the agent performing the act must possess altruistic intentions (1982 p.115). This requirement, Heyd claims, allows us to make sense of the meritorious nature of acts of supererogation. In this paper I will investigate whether there is good reason to accept that this requirement is a necessary condition of supererogation. I will argue that such a reason can be found in cases where two people act in the same way but with only the person who acted with altruistic intent counting as having performed an act of supererogation. In such cases Heyd's intention requirement plays an important role in ruling out acts that intuitively are not supererogatory. Despite this, I will argue that we should reject Heyd's requirement and replace it with a moral intention requirement. I will then investigate how to formulate this requirement and respond to two objections that might be raised against it.

1. Introduction

In 2007 a man collapsed at a New York Subway station and fell onto the tracks as a train approached. Wesley Autrey, a fellow passenger, jumped down onto the tracks and held the other man down as the train came to a stop above their heads, saving the man's life. Acts like these that go beyond what is morally required are called 'supererogatory'. The basic features of these actions are that they are morally valuable, they go beyond duty and they cannot be demanded of people.

In this paper I will investigate Heyd's claim that in order for an act to count as supererogatory the agent performing the act must possess altruistic intentions (1982 p.115). According to Heyd, acts of supererogation must be performed with the intention of benefitting people other than the agent (Heyd 1982 p.136). I will call this feature of Heyd's definition the Altruistic Intention Requirement (AIR). This claim is important; if true it will limit the range of acts that can rightfully be classed as supererogatory.

This requirement, Heyd claims, allows us to make sense of the meritorious nature of acts of supererogation. In this paper I will argue that Heyd is right to identify the need to account for the meritorious nature of acts of supererogation. However, there is also good reason to think, as Mellema does, that Heyd's requirement is overly restrictive. This means that we must find another way of making sense of the claim that acts of supererogation are meritorious. I will argue that we can do so by replacing Heyd's requirement with a moral intention requirement.

2. Why include AIR in the definition?

Let's now turn our attention to investigate what reasons there are for making AIR a necessary requirement for acts of supererogation. Heyd argues that this requirement is needed to make sense of the claim that acts of supererogation are meritorious (1982 p.136). Heyd argues that without this requirement, an act

could count as supererogatory even if it is completely lacking in moral merit. In this section I will investigate how we should understand merit and explain the importance of merit for understanding supererogation.

Heyd argues that moral merit is an essential feature of acts of supererogation (1982 p.132). By 'merit' Heyd, means something distinct from 'praiseworthiness'. It is unclear exactly how we should understand this distinction. One way that Heyd distinguishes the two is unconvincing. Heyd argues that praiseworthiness can be influenced by a variety of reasons, such as the agent performing the act, the person doing the praising and the consequences that the praise might have while merit is related more strictly to what is deserved (1982 p.139). This, though, seems to be a strange way to think of praiseworthiness. Those that agree with Peter Strawson's claim (1968) that the desirability of praising or blaming someone is the wrong kind of reason to praise or blame them will find it odd to see that this might influence how praiseworthy an action is. Nevertheless I think Heyd is right to think that there is a distinction that can be made here between acts that are praiseworthy and acts that are worthy of some other form of positive appraisal.

A more promising way of making the distinction is suggested by Heyd's claim that the principles of merit, "take into consideration primarily the act itself and not the agent's motives, virtues, strength of will," (1982 p.139). This way of distinguishing between the two makes the merit an agent receives from performing an act distinct from the praiseworthiness. While merit is focused only on the act, praiseworthiness is also concerned with the character of the agent. Merit then is concerned with everything praiseworthiness is concerned with apart from that relating to the character of the agent. Given this we might think that an act is meritorious iff it would be praiseworthy if performed for the right reasons. In other words, the act would be praiseworthy if it is the appropriate features of the situation that motivate the agent to perform the act. However, this would not explain why merit is supposed to be a form of positive appraisal. Someone who accidentally performs an act that would have been praiseworthy if performed for the right reasons is not worthy of any form of positive appraisal. There must, then, be an intentional component to merit. Given this I propose the following definition:

Moral Merit: An act is morally meritorious iff it is an act that is intentional under a description such that it would be praiseworthy if performed for the right reasons.¹

According to this definition, an agent can be said to have earned moral merit if under a legitimate description of the agent's intentions, the agent would be morally praiseworthy for performing the act if motivated by the appropriate features of the situation. This definition allows us to clearly see the relationship

¹ The terminology 'intentional under a description' was introduced by Elizabeth Anscombe, who noted that a single action can have several different descriptions and only be intentional under some of these descriptions (2000 p.11). For example, an act may meet the descriptions 'sawing a plank' and 'making a squeaky noise with a saw,' but only intentional under the first (Anscombe 2000 p.11).

between praiseworthiness and merit. Merit is separate from praiseworthiness, as it is possible for an act to be meritorious without being praiseworthy (when the agent is motivated by the wrong reasons). However, the two are closely related, as whenever an agent performs a meritorious act for the right reasons she will be praiseworthy.

Introducing the idea of merit allows us to associate the performance of a supererogatory act with a special kind of positive appraisal that is distinct from praiseworthiness. It is important to do so, as agents who perform acts of supererogation do seem to be worthy of positive appraisal. When someone performs an act like Autrey's we think it right that he should receive special recognition for his actions.

We are now in a position to see the importance of the Altruistic Intention Requirement. This requirement can, in Heyd's words, "rule out accidental, unconscious, involuntary or self-regarding acts as cases of supererogation," (1982 p.116). AIR prevents acts that are lacking in moral merit from counting as supererogatory. We can see why this is important by considering the following example.

Suppose Julie and Jane, have both made a large donation to charity and, as a result, saved hundreds of people from starvation. Performing an act of this sort seems to be supererogatory. However, it turns out that only Julie intended to give money to charity. Jane meant to transfer her money between two of her own bank accounts but accidentally put the money in the charity's account. Jane would ideally like to retrieve the money but is unable to do so.

Despite the fact that these acts are similar and are just as likely to have morally valuable consequences only one of these acts can be classed as supererogatory. Only Julie's act can be counted as morally meritorious as only her act is intentional under a description (helping the hungry) that would make it praiseworthy if performed for the right reasons. Jane had no intention of benefitting anyone else and only did so accidentally. This example makes clear the benefits of including AIR in the definition of supererogation. By making it a necessary condition of acts of supererogation that the agent performing the act must intend to bring about somebody else's good, AIR prevents acts like Jane's that are lacking in moral merit from being classed as supererogatory.²

² A related issue is the debate between objectivists and subjectivists about moral obligation. For a defence of the subjectivist view see Smith (2010). For a defence of Objectivism see Graham (2010). Objectivists hold that the rightness of performing an act in any given situation is determined by the objective facts of the situation, not the agent's beliefs about those facts. On this understanding of obligation, it does not matter whether or not the agent is aware, or even could be aware, of what the best act would be. Subjectivists, on the other hand, hold that the act that it is right for an agent to perform is determined by her beliefs about the situation. Note that on the objectivist account we can perform a moral obligation without having any awareness that doing so will be the right thing to do. Heyd's point is that whatever we think about moral obligation, the view of the agent plays an important part in assessing whether or not an act can be classed as supererogatory. Heyd's requirement then, has a subjective component to it. In order for an act to count as supererogatory the agent must believe that it will have good consequences for others or at least be performing the act as an attempt to benefit others. Of

3. Why Intentions?

We can see, then, the important role that AIR plays in Heyd's definition of supererogation. However, we might wonder whether Heyd's requirement goes far enough. We might think that it is an agent's motivations that are important rather than her intentions. The difference between requiring altruistic intention rather than altruistic motivation is that for an agent to have an altruistic intention, she must aim to bring about some benefit for others. To have an altruistic motive the agent must be motivated by this appreciation of the benefits the act will have for others. In this section I will consider this challenge to a requirement that looks to the agent's intentions rather than motivations. I will argue, in support of Heyd's position, that it is the intentions not the motivations of the agent that are important for an act to be considered supererogatory.

The claim that it is motivation rather than intention that is important for supererogation is made by Sheldon Peterfreund. He argues that altruistic motivation is a necessary condition of supererogatory action (1978 p.55). To support this claim Peterfreund gives the example of two doctors who decide to go to a plague-ridden city to help the sick residents. One doctor is motivated by a desire to help those affected by the disease while the other is motivated by a sense of adventure. Peterfreund argues that we should only count the first doctor's act as supererogatory.

I think that Peterfreund's point is made stronger if give an example where the agent's motivations are morally bad rather than morally neutral. Suppose a power hungry misanthrope is standing for election as mayor. She wants to be mayor so she can more effectively enact her evil plans. Our election candidate passes a river and hears a drowning child scream for help. She recognizes that saving the child is likely to win her some votes in the forthcoming election. The candidate can be said to be acting with altruistic intent, securing the child's safety is crucial to her plan of securing the townspeople's votes. Her motives though are not altruistic. We might think that it would be strange if someone who knew the candidate's motivating reasons for acting in this way asserted that this act was 'beyond the call of duty'.³

The underlying point in Peterfreund's argument is that altruistic motivation is necessary for acts of supererogation because supererogatory acts are necessarily praiseworthy. This is a point that many who have written on the topic would agree with.⁴ In this section I will be defending the claim that it is merit and intentions that are necessary for supererogation rather than praiseworthiness and motivation.

course, AIR requires more from agents than subjectivists about moral obligation do. As well as performing the act that is right from a subjective point of view the act must also be performed with the intention of benefitting other people.

³ Thanks to an anonymous referee for suggesting an example of this kind.

⁴ McNamara says that this is part of 'The Standard Analysis' (2011 p.203). Mellema says that this is part of 'The Standard Account' (1991 p.17). The view is endorsed by Hale (1991 p.273), Horgan and Timmons (2010 p. 32), Jacobs (1987 p.97), Kawall (2003 p.495), Montague (1989 p.102) and Peterfreund (1978 p.54).

Why then does Heyd opt for intentions rather than motivations? The reason is that Heyd, following John Stuart Mill, thinks that intentions form part of the description of the act while motives do not (Heyd 1982 p.137; Mill 2001 p. 18 Fn.2).⁵ In other words, an act can be performed with different motivations and be the same act but the same cannot be said for intentions. An agent in the same situation but acting with different intentions would have performed a different act. As a result, the motivation that led to the act does not alter the moral evaluation of the act; it only alters our evaluation of the agent. As supererogation is an act evaluation this gives us good reason to think that it is merit and intentions that are necessary rather than praiseworthiness and motives.

Heyd's approach allows for an attractive symmetry between supererogation judgements and other act evaluations. It seems plausible to think that there is no motivation requirement for right and wrong actions. When we want to know whether someone has performed an act that is right we do not need to know what motivated her to act as she did. Someone with an obligation to pay her taxes performs her duty when she pays her taxes intentionally. It does not alter our evaluation of the act if she is motivated to do so by an appreciation of her duty or by a fear of punishment. Knowledge of the agent's motivations may alter our judgement of the agent performing a right action but it will not alter our judgement of the act itself. Given that both supererogation judgements and rightness judgements are evaluations of acts and not agents, it should be considered a benefit of Heyd's account that it fits with this plausible way of thinking about rightness judgements.

If we accept the way that Heyd and Mill choose to individuate acts then we are in a position to respond to the examples. In the doctor example, assuming that both doctors act with altruistic intentions, we can say that the doctors perform the same act. The different moral evaluations that we might think the two doctors are due is not down to the performance of different acts but down to the features of their character that led to the performance of the act. We can accept that the adventuring doctor is less praiseworthy than the other doctor but maintain that his act is supererogatory. Likewise, we can accept that the election candidate is morally blameworthy but maintain that her act is supererogatory.

Of course, the above response to Peterfreund only works if we are happy to individuate acts in the way that Heyd and Mill choose to. While a complete defence of this way of individuating acts is beyond the scope of this paper I will briefly explain two reasons why I think this is the right way to do so.

One reason to individuate acts in this way is given by Heyd.⁶ He claims that it is often very difficult to determine the underlying motives that lead people to act (1982 p.187). Suppose we see someone give money to a homeless person who has asked her to spare some change. It seems hard to pinpoint exactly what would motivate the person to give the money. Of course the recognition that this money will help the person is likely to play some role but there may well be other factors that influence this decision. Perhaps the agent has a desire to

⁵ For an explanation of Mill's position here see Ridge (2003).

⁶ A very similar point is made by Mike Ridge (2002 p.66).

appear generous. Discovering what it was that motivated the agent to act in this way will prove difficult. Discovering intentions, on the other hand, is more straightforward. We can be fairly confident that an agent who gives money to charity is aware that some good shall arise from this. If we need to know that the motivations of the agent are altruistic before we can class an act as supererogatory then it will prove incredibly difficult to class any act as supererogatory.

The second reason why this seems to be the right way to individuate acts is that it allows us to explain a familiar feature of moral duties. By not including a motivation requirement, Heyd's definition allows acts of supererogation to fit easily with how we think about moral obligations and prohibitions. It is a common thought that it is possible to perform the right act for the wrong reasons.⁷ We can also perform a morally wrong act and be in no way blameworthy for doing so. If we individuate acts according to intentions then we can explain how this is possible. Someone can be motivated to perform acts by the wrong kind of considerations. If we individuate acts in terms of motivations, though, then we can no longer make sense of this, as the agent acting with a different motivation will be performing a different act.

Someone sympathetic to Peterfreund's position might respond by claiming that acts of supererogation are special in some way, as only agents with altruistic motivations can perform them. However, given that this would be claiming an important difference between supererogation and other act evaluations, some reason needs to be given to think that supererogation is different in this way. One way to do so would be to show that we would not describe an act as supererogatory unless we judged the agent to have been motivated in the right way.

However, looking to how the term 'supererogation' and its ordinary language equivalent 'beyond the call of duty' are used gives us no reason to think that this is the case. The reports on the actions of Wesley Autrey show this clearly. Those commenting on Autrey's act seemed happy to describe his act as supererogatory without knowing what motivated Autrey to act as he did. Both the Mayor of New York at the time Michael Bloomberg and assemblywoman Audrey I. Pheffer, described Autrey's act as supererogatory without making any reference to Autrey's motivations only his intentions.⁸ What this tells us is that ordinary people seem happy to describe acts as supererogatory (or beyond the call of

⁷ For example, Kant distinguishes acts that are in line with duty from those that possess moral worth. In order to have moral worth, an act must be performed by an agent who is motivated in the right way. (1993 p.11). Similarly, Ross claims that an act can be morally right but not morally good (2003 p.156).

⁸ Both Bloomberg and Pheffer described Autrey's act as "beyond the call of duty". Both mention Autrey's intentions: Pheffer says that Autrey acted, "to save a strangers life", while Bloomberg describes Autrey as acting, "to rescue a complete stranger." For the report on Pheffer's quote see *The Wave* (2007). For a transcription of Bloomberg's speech see (NYC.gov 2007). Philosophers are also happy to describe Autrey's act as supererogatory without explaining what motivated him to perform the act. Both Ferry (Forthcoming) and Weinberg (2011 p.274), describe Autrey's act as a clear case of supererogation without explaining what motivated Autrey to perform the act.

duty) without first discovering what the motives of the agent were. Now, of course, by itself this does not tell us that supererogation does not involve an assessment of the motivations of the agent. However, given that other moral evaluations of acts do not seem to require this, the onus is on those who think that supererogation should be treated as a special case provide some support for their claim. One way of doing so would be to appeal to the way in people use the term supererogation or ordinary language equivalents. My point is that there seems little support for such a move.

Someone sympathetic to Peterfreund's position might respond by claiming that the reason that people are happy to describe Autrey's act as supererogatory is that they are making an implicit assumption about his motives. This assumption might be thought to be a justified one. Perhaps there is good epistemic reason to think that people who perform acts like Autrey's are motivated by altruism. Alternatively, it might be thought that the principle of charity justifies this assumption. Either way, if we can be justified in assuming that Autrey acted with altruistic motivation then it would explain why the Mayor was happy to describe Autrey's act as supererogatory without explaining his view of Autrey's motivations.

We can evaluate this response if we imagine a situation where an agent like Autrey, let's call him Autrey*, acts in a similar way but explains that he was motivated by self-interested reasons. Perhaps he foresaw the rewards and glory that he would gain as a result and it was this that motivated him to act. Maybe he was motivated by a belief that acting in this way would secure him a place in heaven. Nevertheless, Autrey* makes it clear that he did act with altruistic intent and that he knew that he risked his life by doing so. We now have a case where it would be unjustified to assume that the agent was altruistically motivated. According to Peterfreund's requirement we should reject the claim that Autrey* performed an act of supererogation, as it is clear that Autrey* was not altruistically motivated. It seems clear to me that this is a counter intuitive result. If the Mayor made the same comments about Autrey* we would not think that they would be false. Even if Autrey*'s motivations are not praiseworthy we can still correctly describe Autrey*'s act as supererogatory. The response that the mayor must be implicitly assuming that Autrey acted with praiseworthy motivations can be rejected. It would still be correct to describe this act as supererogatory even if the agent made it clear that he was motivated by self-interested reasons.⁹

⁹ Of course, it may well be inappropriate to describe Autrey* as a hero but that is just because describing someone as a 'hero' is a character assessment rather than an act assessment. Although, performing an act of supererogation and being a saint or a hero are often linked they are not equivalent. As McNamara points out (2011 p.219), some heroic acts may be morally required. For example, we might describe a fireman as a hero for running into a burning building whilst also judging that he was morally required to do so. In addition, some supererogatory acts may not be saintly or heroic. Those who accept, as Heyd (1982 p.148) and Mellema (1991 p.7) do, that small favours, such as stopping to help a stranger fix her car, are supererogatory surely do not view all such acts as saintly or heroic. Thanks to an anonymous referee for raising this concern.

In this section, I have given two reasons to think that Heyd is right to focus on intentions rather than motivations. First, it is much harder to determine motives than it is to determine intentions. Second, it is consistent with the possibility of performing the right act for the wrong reasons. Finally I considered the response that acts of supererogation might somehow be different from other moral evaluations of acts. I argued that looking at how the term is used provides little support for this claim.

I accept that some people might still find it odd to think that the election candidate is performing a supererogatory act. However, I do not think that this intuition gives us sufficient reason to add a motivation requirement to our definition of supererogation. My own view, which I defend elsewhere, is that this intuition can be explained in terms of Gricean implicature.¹⁰ Defending this claim here would take this paper too far off topic. Nevertheless, at this stage we can say that although some might have the intuition that it would be odd to describe the election candidates act as supererogatory it is an open question whether this oddness should be explained as part of the truth conditional content of 'supererogatory' or whether there is some other explanation for this. In addition someone seeking to defend this view must respond to the two reasons that I have given in support of Heyd's claim that the agent's motivation are not relevant when assessing whether or not an act is supererogatory. Finally, my primary goal in this paper is to defend the claim that we should replace AIR with a moral intention requirement. I hope that those that are not persuaded by what I have said in this section will be persuaded by what I say in the remainder of the paper and accept that a moral intention requirement is preferable to AIR.

4. Why Altruism is Not Necessary

So far I have argued that AIR plays an important role in ruling out acts that intuitively do not seem to be supererogatory from counting as such. However, in this section I will argue that AIR should be rejected for two reasons. First there seems little theoretical support for restricting the claim to altruistic intentions rather than moral intentions. Second, doing so excludes acts that intuitively should count as supererogatory.

The first reason to replace AIR is that, given the role this requirement plays in the definition, there seems little reason to restrict the claim to altruistic intentions rather moral intentions in general. We have seen that the reason to include AIR is that it explains why acts of supererogation are morally meritorious. However, a moral intention requirement would also do this. There is nothing in what Heyd says about merit, nor in my attempt to give a more precise formulation of the idea, that restricts merit to those with altruistic intentions. There seems to be no theoretical reason then, to restrict the claim to altruistic intentions rather than moral intentions in general.

Nevertheless, we might think that the reason why we should restrict the claim to altruistic intentions is that this is a better fit with our intuitions. Certainly, it does seem to be true that most cases that we think of as supererogatory do involve an

¹⁰ See Archer (Unpublished). For an introduction to implicature see Grice (1991 Chapter 2).

agent acting with altruistic intent. The two examples of supererogation we have looked at so far, those of Audrey and Julie, are both cases where the agent is acting with altruistic intent. Perhaps, then, there is some prima facie support for AIR. Nevertheless, this support will be undermined if we can find examples of acts that are intuitively supererogatory but performed by agents with moral but not altruistic intent.

I will give two examples that I think meet this description, one from Mellema and one of my own. To be clear, what we are looking for are cases where we would happily describe the act as supererogatory even if the agent lacks any altruistic intention. If we find a case of this sort then we will have found a counter example to AIR.

The first counter example is given by Mellema (1991 p.19-20). His example is of a political prisoner being held by terrorists. The prisoner is told by the terrorists to renounce his government and pledge allegiance to their cause. No one will find out about his refusal to cooperate and so doing so will not benefit those he cares about in any way.¹¹ The consequences for him, though, will be terrible, as he will face torture from the terrorists. However, the prisoner is a principled and patriotic man who believes that he should not renounce his government. Out of respect for these principles, and with knowledge of the lack of good consequences for other people that will result from his actions, the prisoner refuses to renounce his government.

Mellema argues that in this example the man performs a supererogatory act in refusing to comply with the demands of the terrorists. The prisoner heroically stands up for what he believes in and does so in a way that exceeds the requirements of duty. However, the act is not performed with altruistic intent. If we accept that the act that Mellema describes is supererogatory then we will have to accept that AIR is not a necessary condition of supererogation.

For the sake of those unconvinced by this example I will give one of my own. Simon Wiesenthal was a Holocaust survivor. After The Second World War, he dedicated his life to tracking down fugitive Nazis so they could be brought to justice. Most famously, Wiesenthal helped to track down Adolf Eichmann, leading to his capture in 1960. The reason that Wiesenthal dedicated his life to this cause was not vengeance but justice.¹² By choosing to spend his life in this way Wiesenthal put himself in a position where he would have to suffer both physical attacks from neo-Nazis and the emotional pain of constantly reliving his traumatic experiences.¹³ It seems to me that Wiesenthal's actions should be classed as supererogatory.

¹¹ In Mellema's description of this example he says that the act will bring about only 'bad consequences' (1991 p.20). I have changed this to 'indifferent consequences' because I believe it makes the example more plausible. This in no way changes the point that Mellema is making.

¹² Hence why the title of one of Wiesenthal's books is *Justice Not Vengeance* (1989). In this book Wiesenthal's friend, Peter Michael Lingens, tells of the important role that the idea of justice had in shaping Wiesenthal's view of the world from an early age (1989 p.3).

¹³ Wiesenthal talks of the emotional pain he put himself through here (1989 p.27) and the physical attacks here (1989 p.401).

It is unclear whether Wieselthal can be said to have intended to benefit others with his act. Perhaps he viewed his acts as safeguarding people against these criminals. Perhaps he viewed his acts as in some way intending to benefit the dead by keeping their memory alive.¹⁴ Importantly for my purpose, we might think that his intention was not to bring about some benefit to anyone; it was just to bring about justice. I do not claim that this is how we should understand Wieselthal's intentions, only that it is conceivable to think of his intentions, or those of someone in a similar position, in this way. For the purposes of this discussion I will be assuming that these were Wieselthal's intentions. The important question we must ask ourselves is whether thinking of Wieselthal's intentions in this way changes our view about whether or not his acts were supererogatory. It seems clear to me that it does not. We would still think of Wieselthal's acts as supererogatory even if we found out that his intentions were not altruistic.

I think that these two examples give us good reason to reject AIR. Before doing so though, it is worth considering a response that might be made to my argument. Someone seeking to defend AIR might argue that an act cannot be valuable unless it benefits people in some way.¹⁵ If we reject the moral value of this act then there is no reason to think that this is an act of supererogation and so no reason to see this as a counter-example to Heyd's definition. For example, a consequentialist may contend that if no positive consequences are expected to arise as a result of the act then according to this view no moral value has been produced. We might think then that we can retain Heyd's requirement by denying the moral value of the prisoner's act. Of course this response depends upon a certain way of understanding moral value, such that it is the consequences of acts that are important for moral value.

If we accept this then in order to be counted as supererogatory an act must benefit people in some way. This leaves us with a dilemma. Either the acts in the previous examples do not benefit people and so have no moral value or they do benefit other people and so are altruistic after all.

However, I think even if we accept this claim about moral value this response does not present my argument with a real problem. Remember, we are interested in whether the intentions of the agent are altruistic. If we accept the consequentialist understanding of moral value then we can view the agent's act as having good consequences for other people, and being morally valuable as a result, without the agent intending to bring these consequences about.¹⁶ The act can be, then, be morally valuable without the intentions of the agent being altruistic.

In this section I have argued that we should reject AIR. However, we must not lose sight of the reason why AIR was seen to be an important part of the definition of supererogation. Without this requirement we would be unable to rule out cases such as that of Jane, the unwitting donator to charity, from

¹⁴ Wieselthal suggests that this was the intention behind many of his acts (1989 p.411).

¹⁵ Thanks to an anonymous referee for pressing me on this point.

¹⁶ A similar point is made by Mellema (1991 p. 20).

counting as supererogatory. This may be viewed as a greater cost than having to rule that the prisoner's act was not supererogatory.

5. Moral Intention Requirement

However, abandoning AIR does not mean we are forced to accept that Jane performs a supererogatory act. Instead we can replace AIR with a similar requirement that allows us to label the prisoner's act as supererogatory but withhold the term from Jane's act. If the requirement looked for moral intentions rather than altruistic intentions then it would be possible to describe the prisoner's act as supererogatory but not Jane's. We should then, replace AIR with a moral intention requirement. Below is a first attempt to formulate this requirement:

Moral Intention Requirement One (MIR1): For an act to be supererogatory the agent must be acting with moral intentions.

Like AIR, MIR1 would allow us to withhold the term 'supererogatory' from Jane's act. As Jane does not intend her act to bring about any moral value it cannot be classed as supererogatory. Julie, on the other hand, does intend for her act to bring about moral value and so can be said to have performed an act of supererogation.

Unlike AIR though, MIR1 would allow the tortured prisoner's act and Wiesenthal's acts (on my interpretation of his intentions) from counting as supererogatory. Not only would MIR1 give us the right result in these two examples it would do so in a way that explains our intuitions. The reason why it seemed implausible to say that Jane was performing an act of supererogation was that any moral value that arose from her action was accidental. Likewise, with both the prisoner and Wiesenthal's acts we want to say that their intentions do not need to be altruistic in order for their acts to count as supererogatory, so long as they are moral. MIR1 allows us to do this.

However, there is a problem with this way of formulating the moral intention requirement. MIR1 requires acts of supererogation to be performed by agents whose intention is an explicitly moral one. There are two problems that might arise from this. The first problem is in cases where an agent intends her act to bring about a state of affairs that seems to be morally valuable but which would not be recognized as such by the agent. For example, suppose Hugo is a nihilist. This does not stop Hugo from performing many morally valuable acts. One day Hugo walks past a burning building and hears shouts for help coming from inside. At great personal risk, Hugo runs into the building to save those trapped inside. If Hugo were asked whether he was aware of a moral reason favouring his act he would deny it. We might imagine him saying the following, "I don't believe in morality. I am just doing this because I feel sympathy for those trapped inside." This seems like an act of supererogation even if Hugo would deny it. This is a problem for any moral intention requirement, as it seems as if Hugo may be said to lack awareness of any moral reason in favour of his action. As a result Hugo's action would fail to meet MIR1 and so could not be classed as supererogatory.

Another example of an agent who intends her act to bring about a state of affairs that seems to be morally valuable but which would not be recognized as such by the agent is the case of Huckleberry Finn.¹⁷ In Mark Twain's *Huckleberry Finn* (1966), Huck fails to turn in his friend Jim, a runaway slave, despite judging that this is what he ought to do.¹⁸ We might think that Huck performs a supererogatory act in this case. Certainly his act seems to be morally valuable and we might think that the risks he is taking would prevent it from being obligatory. Nevertheless, if we require supererogatory acts to be performed by agents with explicitly moral intentions then this act cannot be counted as supererogatory.

The second problem with requiring supererogatory acts to be performed by agents with explicitly moral intentions that we might think that someone whose intentions are explicitly moral would be a slightly strange moral agent. We might think that such an agent has, in the words of Bernard Williams' famous objection to utilitarianism, "one thought too many," (1981 p.18). It would, perhaps, be more appropriate for an agent to be intending to help others rather than intending to perform morally meritorious acts. The problem that could be raised MIR1 is that it requires agents to be acting with the wrong kind of intentions.

However, these problems can be avoided if we formulate the moral intention requirement differently. These criticisms only arise if we take a moral intention requirement to require agents to have explicitly moral intentions. There is though, no need to formulate the requirement in this way. Instead, we can say that in order to be supererogatory the agent's act needs to be intentional under a description in virtue of which it is morally meritorious. The agent need not be aware that her act meets this requirement. We should then give a new version of the moral intention requirement that avoids these problems.

Moral Intention Requirement Two (MIR 2): For an act to be supererogatory the act must be intentional under a description D such that D picks out features in virtue of which the act is morally meritorious.

This formulation shares the advantages that AIR and MIR1 have over having no intention requirement. Like MIR1 and AIR, MIR2 allows us to prevent Jane's act from counting as supererogatory.

MIR2 also shares the advantages that MIR1 has over AIR. Like MIR1, MIR2 allows us to say that both Mellema's prisoner and Simon Wiesenthal's acts (on our interpretation) are supererogatory in a way that explains our intuitions. The reason why it seemed implausible to say that Jane was performing an act of supererogation was that any moral value that arose from her action was accidental. Likewise, in the case of the prisoner, the reason why it may be judged that his act is supererogatory is that he has performed an act that is morally meritorious under the description 'refusing to denounce one's ideals'.

¹⁷ Thanks to an anonymous referee for suggesting that one example might not be enough here.

¹⁸ For discussions of whether or not Huck Finn can be said to be morally praiseworthy see Jonathan Bennett (1974) and Nomy Arpaly (2003).

Unlike MIR1 though, MIR2 avoids the objection that arose from requiring agents to have explicitly moral intentions. The problem with both Hugo and Huck was that they were unaware that their acts were morally meritorious. Nevertheless, they both pass MIR2. Hugo's act is intentional under the description 'Saving those in the building' and to perform an act under this description certainly would be morally praiseworthy when performed for the right reasons. Huck's act is intentional under the description 'Allowing my friend Jim to remain free' and again seems to be praiseworthy under this description, providing of course Huck is performing the act for the right reasons. Similarly, the claim that this would require agents to be acting with the wrong kind of intentions also disappears once we put the point in this way, as the agent does not need to possess an explicitly moral intention.

To conclude this section, including MIR1 or MIR2 in the definition of supererogation is preferable to AIR as these requirements allow acts such as those of Mellema's prisoner and Simon Wiesenthal to count as supererogatory. These requirements are also preferable to a definition lacking in any requirement on the intentions of the agent as it can rule out cases where the agent's act is accidentally morally valuable act. MIR2 is preferable to MIR1, as it does not require agents to possess explicitly moral intentions. As a result, we should replace AIR with MIR2.

6. Problems for Moral Intention Requirement

In this section I will investigate two problems that might be raised against MIR2 and show that neither gives us good reason to reject this requirement.

The first problem that might be raised against MIR2 is given by Heyd. He argues that putting the requirement in terms of moral intentions prevents acts of self-sacrifice that do not increase the overall value of the situation from being counted as supererogatory (1982 p.132). To see this imagine there are exactly two acts available to an agent that will both bring about the same total amount of value. For the sake of clarity I will assign numbers to represent the value of the two acts:

	Value for Agent	Value for Other	Overall Value
Act A:	10	1	11
Act B:	1	10	11

It seems reasonable to think that both acts are permissible here. They will both be quite good for one person and very good for another and neither act will produce more overall value than the other. However, it seems reasonable to think that an agent who performs Act B may have performed an act of supererogation. Heyd's claim is that if the requirement is put in terms of moral

intentions rather than altruistic intentions then this act cannot be deemed supererogatory, as there is no greater moral value in performing this act.¹⁹

However, I think that there is a simple response that can be made to this problem. The fact that performing this act will not increase the overall moral value of the situation does not mean that Act B cannot be said to be more morally meritorious than Act A. Foregoing things that are valuable for you for the sake of those of others' may not bring about anymore value but nevertheless the agent who performs such an act intentionally under the description 'bring about value for others' will be praiseworthy if motivated for the right reasons. As the point of either intention requirement is to capture the idea that supererogatory acts are necessarily morally meritorious, and MIR2 can do so in this case, we need not worry about this problem.

The final problem that might be raised against MIR2 is that it could allow self-regarding acts to count as supererogatory. If we think that self-regarding acts can be morally valuable then we might think that, unlike AIR, MIR2 would allow such acts to count as supererogatory. AIR rules out such acts, as it requires agents to be aiming to bring about some good for other people. If we find it implausible that there can be self-regarding acts of supererogation then this might seem to be an unacceptable consequence of replacing AIR with MIR2. If we think that acts of supererogation cannot be self-regarding, and that AIR would explain this while MIR2 would not, then this puts pressure on the claim that we should replace AIR with MIR2.

There are two ways to respond to this criticism. The first is to accept that acts of supererogation can be self-regarding. Jason Kawall puts forward an argument in defence of this position (2003). He claims that self-regarding acts of supererogation are possible. He supports this claim by giving an example of a waiter who wants to go to university. In order to do so the waiter takes on a second job to save up the money to fund his studies. Kawall argues that this act is supererogatory despite the fact that it only provides benefits to the person performing the act (2003 p.490). If we accept Kawall's argument then compatibility with self-regarding acts of supererogation becomes an advantage rather than a problem. Rather than viewing MIR2 as too permissive we might now view AIR as too restrictive in what acts can count as supererogatory.²⁰

While this does seem like a genuine advantage for my view I do not want my argument to be dependent on Kawall's. In order to convince those that are not

¹⁹ Note that this problem only arises if we have a teleological conception of value, that is, one that holds that it is states of the world that are the primary bearers of value. If we have a buck passing view of value then we can ignore this problem. A buck passing view is one that holds the following: "Being valuable is not a property that provides us with reasons. Rather, to call something valuable is to say that it has other properties that provide reasons for behaving in certain ways with regard to it," (Scanlon 1998 p.96). If we accept this view of value then we can hold that Act A might be just as morally valuable as Act B (it provides us with just as many and as strong reasons to respond in certain ways) but less praiseworthy (it provides less reasons to praise the agent performing the act). Thanks to X and an anonymous referee for helpful discussions on this point.

²⁰ Thanks to an anonymous referee for pressing me on this point.

persuaded by Kawall's argument I must find some way of showing that those who reject the existence of self-serving acts of supererogation can nevertheless accept the claim that AIR should be replaced by MIR2. My strategy for doing so will be to show that whatever reason we give for wanting to reject Kawall's argument will prevent self-regarding acts from counting as supererogatory regardless of the intention requirement we include in our definition. To be clear, I am not attempting to show that these responses give us good reasons to reject Kawall's argument. All I am trying to show is that the reasons that might be given to reject it will show commitments that prevent self-regarding acts from counting as supererogatory regardless of our intention requirement.

Before looking at these responses though, it is worth pointing out one way of responding to Kawall's argument that cannot be brought in at this stage of the discussion. One reason that might be given to reject Kawall's argument would be because it is incompatible with AIR. Those convinced of the truth of AIR might see this as providing sufficient reason to reject Kawall's view. Note though, that while rejecting Kawall's claim for this reason is incompatible with my view it does not create a new problem for it. Remember that we are considering whether the fact that my view is compatible with the existence of self-regarding supererogatory acts is an independent problem for my view. If the only reason this is problematic is that it is incompatible with AIR then this is not a new problem for my view, it is simply recognition of the fact that my view is different to AIR.

There are two other ways in which someone might try to respond to Kawall's argument. The first way to respond to this argument would be to claim that the acts described in Kawall's example are not morally valuable as only the agent gains the benefits. Those who think that morality is concerned with our interactions with other people will not view the waiter as having performed an act of moral value. I take it that any acceptable definition of supererogation will make moral value a necessary condition. Given this, those who deny the moral value of self-regarding acts will not count such acts as supererogatory regardless of whether or not AIR is part of our definition of supererogation.

Another response to make would be to argue that acts of supererogation necessarily involve sacrifice.²¹ In self-regarding acts though, it could be argued that the agents are not sacrificing anything, in fact they are benefitting from these acts.²² It is conceivable that someone might want to reject Kawall's claim for this reason. If we think this though, then we ought to make sacrifice a necessary condition in our definition of supererogation. This new condition would then prevent self-regarding acts from counting as supererogatory. Rejecting Kawall's claim for this reason would pose no problem for the claim I am defending.

²¹ As Kawall points out (2003 p.488), this view of supererogation seems to be assumed by both Jacobs (1985) and McGoldrick (1984).

²² Of course, we might think that there is a sense in which the waiter is making a sacrifice. This might suggest that this is not a good way to respond to Kawall's argument.

There are then two ways of responding to the criticism that MIR2 would make self-regarding acts of supererogation possible. First, we can embrace this consequence of this requirement. Note that if we are persuaded by Kawall's argument then the fact that my view would allow self-regarding acts to count as supererogatory is an advantage for the view rather than a disadvantage. Alternatively, we can say that whatever reasons are given for rejecting Kawall's argument will be ones that ought to rule out such acts from counting as supererogatory regardless of the intention requirement.

In this section I have looked at two problems that might be raised against replacing AIR with MIR2. The first problem is that MIR2 may prevent acts that don't increase the overall value but do represent some sacrifice from the agent from counting as supererogatory. I responded to this argument by pointing out that an act can be morally meritorious even if it does not increase overall value. The second worry was that this requirement might allow self-regarding acts to count as supererogatory and gave two ways to respond to this problem, one of which shows that it is not a problem for MIR2 and the other shows that it is an advantage for MIR2. I take these to be adequate responses to these problems.

8. Conclusion

In conclusion, this paper has investigated the claim that altruistic intentions are a necessary requirement for acts of supererogation. I have argued that the requirement has some appeal, as it is able to make sense of the meritorious nature of acts of supererogation. Nevertheless, I have argued this requirement should be abandoned as it prevents cases like Mellema's patriotic prisoner and Simon Wiesenthal from being classified as supererogatory. Completely abandoning any requirement that looks at the intentions of the agent will not do either as this would allow unintentional acts to count as supererogatory. I have argued that we should replace Heyd's Altruistic Intention Requirement with a moral intention requirement. After investigating how best to formulate this requirement, I looked at two problems that might be raised against it and argued that good responses can be given to both of them.²³

Bibliography

Anscombe, G.E. (2000) *Intention* (Cambridge MA: Harvard University Press).

Archer, A. (Unpublished Manuscript) "Are Acts of Supererogation Always Praiseworthy?"

Arpaly, Norm (2003) *Unprincipled Virtue: An Inquiry Into Moral Agency* (Oxford: Oxford University Press).

Bennett, J. (1974) "The Conscience of Huckleberry Finn", In *Philosophy* 49 (188): 123-134.

²³ Thanks to the audience at The 2012 Edinburgh University Ethics Fest for helpful comments on an early draft of this paper. Special thanks to Elinor Mason, Mike Ridge and three anonymous referees for this journal for their detailed and insightful comments.

Ferry, Michael (Forthcoming) "Does Morality Demand Our Very Best? Moral Prescriptions and the Line of Duty" In *Philosophical Studies*.

Graham, P. (2010). "In Defense of Objectivism about Moral Obligation" In *Ethics* 121: 88- 115.

Grice, P. (1991) *Studies in the Way of Words* (Cambridge MA: Harvard University Press).

Hale, S. (1991) "Against Supererogation." In *American Philosophical Quarterly* 28 (4): 273-285

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

Horgan, T. and Timmons, M. (2010). "Untying a Knot from the Inside Out: Reflections on the 'Paradox' of Supererogation", In *Social Philosophy and Policy*, 27: 29–63.

Jacobs, R. A. (1985) "Obligation, Supererogation and Self-Sacrifice," In *Philosophy* 62: 96-101

Kant, I. (1993) *Grounding for the Metaphysics of Morals* Translated by James W. Ellington (Indianapolis: Hackett Publishing).

Kawall, J. (2003). "Self-Regarding Supererogatory Acts", In *Journal of Social Philosophy* 34 (3): 487-498

McGoldrick, P.M. (1984) "Saints and Heroes: A Plea for the Supererogatory," In *Philosophy* 59: 523-528.

McNamara, P. (2011) 'Supererogation, Inside and Out: Toward an Adequate Scheme for Common-Sense Morality.' In *Oxford Studies in Normative Ethics Volume 1* (Oxford: Oxford University Press).

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

Mill, J. S. (2001) *Utilitarianism* (Indianapolis: Hackett Publishing)

Montague, P. (1989). "Acts, Agents and Supererogation." In *American Philosophical Quarterly* 26(2): 100-111.

NYC.gov (2007) "Mayor Bloomberg Hails Heroic Acts By New Yorkers During Weekly Radio Address,"
<http://www.nyc.gov/portal/site/nycgov/menuitem.c0935b9a57bb4ef3daf2f1c701c789a0/index.jsp?pageID=mayor_press_release&catID=1194&doc_name=html%3A%2F%2Fwww.nyc.gov%2Fhtml%2Fom%2Fhtml%2F2007a%2Fpr004-07.html&cc=unused1978&rc=1194&ndi=1> (Accessed 10/12/12).

Peterfreund, S. (1978) "On the relationship between supererogation and basic duty." In *The Personalist* Vol. 59 No. 1 pp.53-57

Ross, W.D. (2003) *The Right and the Good* (Oxford: Oxford University Press).

Ridge, M. (2002) "Mill's Intentions and Motives," In *Utilitas* 14 (1): 54-70.

Scanlon, T. M. (1998) *What We Owe To Each Other* (Cambridge MA: Harvard University Press)

Smith, H. M. (2010). "Subjective Rightness", In *Social Philosophy and Policy* 27 (2): 64-110.

Strawson, P.F. (1968) "Freedom and Resentment", In *Studies in the Philosophy of Thought and Action* (London: Oxford University Press).'

Twain, M. (1966) *The Adventures of Huckleberry Finn* (London: Penguin)

Urmson, J. O. (1958). "Saints and Heroes" Reprinted in Joel Feinberg (ed.) (1969) *Moral Concepts* (Oxford: Oxford University Press)

Weinberg, J. (2011) "Is Government Supererogation Possible?" In *Pacific Philosophical Quarterly* 92 (2): 263-281.

Wiesenthal, Simon (1989) *Justice Not Vengeance* (London: Mandarin)

Williams, B. (1981) *Moral Luck* (Cambridge: Cambridge University Press)

The Wave "Pheffer Honours Heroics of Wesley Autrey"

<http://www.rockawave.com/news/2007-02-09/Community/026.html>