

## **Motivational Judgement Internalism and The Problem of Supererogation**

**Alfred Archer**  
**Department of Philosophy/ Tilburg Centre for Logic, Ethics and**  
**Philosophy of Science**  
**University of Tilburg**  
**PO Box 90153,**  
**5000 LE Tilburg**  
**The Netherlands**  
[\*\*atmarcher@gmail.com\*\*](mailto:atmarcher@gmail.com)

**Abstract:** Motivational judgment internalists hold that there is a necessary connection between moral judgments and motivation. There is, though, an important lack of clarity in the literature about the types of moral evaluation the theory is supposed to cover. It is rarely made clear whether the theory is intended to cover all moral judgements or whether the claim covers only a subset of such judgements. In this paper I will investigate which moral judgements internalists should hold their theory to apply to. I will argue that the possibility of the supererogation amoralist, someone who makes genuine supererogation judgements but remains unmotivated by them, makes it implausible to be an internalist about moral goodness. As a result, internalists should restrict their claim to moral requirement judgements. I will then argue that this creates an explanatory burden for Internalism. In order for their view to be plausible they must explain why some moral judgements and not others are necessarily connected to motivation.

## **I. INTRODUCTION**

Motivational Judgment Internalism (henceforth, ‘Internalism’) is the view that there is a necessary connection between moral judgements and motivation. According to this theory, there is a necessary link between making a moral judgement and being motivated to act in line with this judgement. Motivational Judgement Externalism (henceforth, ‘Externalism’), on the other hand, is the view that there is no necessary connection between the two.

This debate is important for the role it plays in other metaethical debates. When Internalism is combined with a Humean theory of motivation it lends natural support to Non-cognitivism about moral judgements (See Smith 1994 p.12). The Humean Theory of Motivation states that beliefs by themselves are incapable of motivating. If we accept Internalism about moral judgements then we accept that there is a necessary connection between moral judgements. This suggests, for those who accept the Humean view of motivation, that they cannot be purely cognitive states. Of course, things are not this simple and it is possible to combine Internalism with a cognitivist view of moral judgements (eg Smith 1994). An interesting development in recent years has been the attempt to develop hybrid views that seek to combine the strengths of both cognitivism and non-cognitivism.<sup>1</sup> Nevertheless, it is fair to say that anyone attempting to combine Internalism with a Humean theory of motivation and a cognitivist view of moral judgements will have some explaining to do.

My first aim in this paper is to raise an awareness of an important gap in the literature on Internalism concerning the kinds of moral evaluations that the proposed necessary connection might apply to. In §2 I will give a quick overview of the literature on

Internalism and explain the importance of this gap in the literature. In §3 I will investigate the different evaluations that Internalism has been claimed to cover. I will argue that these various evaluations can be grouped into two categories. One way to understand the internalist claim would be that it covers all first-person moral judgments about how to act. Another way in which we might understand the internalist claim would be as a claim regarding some subset of such judgments. In §4 I will raise the problem of the supererogation amoralist for any version of Internalism committed to the truth of Internalism about moral goodness judgements. I will then, in §5, argue that this objection is one that traditional internalist responses are powerless against. In §6 I will argue that the only form of Internalism that survives this problem is one restricted to moral requirement judgements. However, as I will argue in §7, those who defend this form of Internalism owe an explanation as to why it is only a subset of moral judgements that are internally connected to motivation.

## **II. MOTIVATIONAL INTERNALISM: THE STATE OF THE DEBATE**

Many have found Internalism attractive because it is able to explain the strong connection that appears to exist between moral language and motivation.<sup>2</sup> As a number of authors have pointed out, there seems to be something puzzling about someone who claims that an act is obligatory but fails to be motivated to perform it.<sup>3</sup> Suppose Jenny tells us that eating meat is wrong. It would be puzzling if she then told us that she felt no motivation whatsoever to cease being a carnivore. We would be likely to doubt Jenny's sincerity in such a case. Internalists claim that the reason why this seems puzzling and the reason why we are tempted to question Jenny's sincerity is that there is an *a priori* necessary connection between moral judgements and

motivation, so that it is impossible to make a genuine moral judgement without being motivated to act in line with it. Before we evaluate this claim it is worth noting that internalists tend to only claim the existence of a necessary connection between moral judgements and a *pro tanto* motivation to act (i.e. one that is capable of being overridden by motivation deriving from other sources).<sup>4</sup>

In response to arguments like this one, externalists have argued that amoralists, people who make genuine moral judgements but remain completely unmotivated by them, are conceptually possible.<sup>5</sup> This would mean that there is no conceptual connection between moral judgments and motivation.<sup>6</sup>

The most common internalist response to the challenge posed by the possibility of amoralists is to restrict the internalist claim to a subset of moral agents. Different kinds of agent restrictions have been proposed. Some restrict the claim to rational agents, others to normal agents and the claim might also be made about virtuous agents. While these agent restrictions help the internalist respond to the challenge of the amoralist, they do so at a price. The price that must be paid for using this strategy is that the further we restrict the internalist claim the less influence the claim will have for the debate between cognitivists and non-cognitivists. Certainly for some ways of restricting Internalism it seems reasonable to worry that what started as an interesting claim about moral judgements, has become a less interesting (at least for those interested in the nature of moral judgement) claim about certain kinds of moral agent.<sup>7</sup>

There is, though, an important question that Internalists have largely ignored. While there has been a great deal of discussion about the kinds of *agents* Internalism applies to, there has been far less discussion about the kinds of moral *evaluations* this

connection holds for. One recent development in the literature has been the explicit recognition of a point that had previously been left implicit, that Internalism is a theory about first-person judgements (see Manne (2015 p. 261) and Wedgwood (2007 p.25). However, this leaves open the question of which first-person moral judgements Internalism applies to. While different definitions of Internalism focus on different moral evaluations there has been little discussion about which evaluations the theory should cover. Indeed most contributions in the literature seem oblivious to the fact that different evaluations have been given and the importance that this has for the debate.<sup>8</sup> The primary aim of this paper will be to investigate whether Internalism is a plausible view to hold about moral judgements in general or whether it should be understood only to be a claim about moral requirements.

I will limit myself in this discussion to focussing on the connection to motivation that most internalists have focussed on up to now, that is a motivation to perform the action in question. In other words, being motivated to perform the act the agent has made a moral judgement about. Other forms of internalism have been defended.<sup>9</sup> For the purposes of this paper, however, I will focus on the standard approach, which is to link moral judgements to motivation to perform the act being judged.

### **III. A NEGLECTED QUESTION FOR INTERNALISTS**

The question for internalists, then, is which first personal judgments about what to do is their theory supposed to cover? Different definitions of Internalism make the claim about different forms of moral evaluation. There are, at least, five different forms of moral evaluation that the internalist claim has been made about. Some authors have claimed that the connection holds for judgements of moral

obligation.<sup>10</sup> Others have defined Internalism as applying to ‘moral ought’ judgements (Kauppinen, 2008 p. 5). The claim has also been made about judgements of ‘moral rightness’ (Smith, 1994 p.61), ‘moral goodness’ (Dreier, 1990 p.11 and Blackburn, 1984 p.187-8) and about all moral judgements (McNaughton, 1988 p.23 and Dancy, 1993 p.7).

The question we must ask then, is whether Internalism might plausibly be held to cover all moral judgements or whether it is only a subset of moral judgements that are covered by the theory. Before we go about answering this question, it is worth remembering that we have already restricted our discussion to first personal judgements about what to do. As a result, we have already restricted the set of moral judgements being considered. For example, we will not be investigating whether judgements concerning the possession of a virtue or vice are necessarily connected to motivation, as these are not judgements about what to do (though they may lead to such judgements).<sup>11</sup> It is also worth pointing out that among the first personal moral judgements about what to do will be judgements that an act is morally neutral. For example, suppose I am trying to decide whether to drink tea or coffee. I may well decide that it is just morally irrelevant which option I pick. This is, in a sense, a first personal moral judgement about what to do. This, though, is clearly not the kind of judgement that internalists have in mind when they claim that there is a necessary connection between moral judgements and motivation. A more precise version of the question we will be asking then is the following: which first-personal moral judgements that provide some form of recommendation might plausibly be thought to be necessarily connected to motivation?

One answer to this question is that they all are. According to this view, any positive or negative first personal moral judgements about how to act will be accompanied by a corresponding positive or negative motivation.

*Wide Judgement Internalism (Wide Internalism)*: Necessarily, if an agent makes a positive first personal moral judgement concerning an act that she could perform  $\phi$  then she will be *pro tanto* motivated to  $\phi$ . Necessarily, if an agent makes a negative first personal moral judgement concerning an act that she could perform  $\psi$  then she will be *pro tanto* motivated not to  $\psi$ .<sup>12</sup>

According to this view, there is a necessary connection between moral judgements and motivation, as all positive and negative first personal moral judgements about what to do will motivate agents to act accordingly.

We might think that this is how all Internalists intend their view to be understood.

However, some discussions of Internalism seem to suggest that it should be understood in a narrower way. For example, William Frankena describes the debate between internalists and externalists as, “A problem with the analysis of judgments of moral obligation,” (1958 p.40). Similarly, Michael Cholbi describes Internalism in the following way: “If an agent judges that she is morally obligated to  $\phi$ , then, that agent is, to at least some minimal extent, motivated to  $\phi$ ,” (2011 p.28). These authors could be interpreted as making a claim about all moral judgments and focussing on the case of moral obligation for simplicity. Nevertheless, I think an equally acceptable interpretation would be that it is a claim specifically about moral requirement judgements. Similarly authors who define Internalism in terms of moral ought judgements, moral rightness judgements or moral goodness might also be understood to either be making a wide claim about all first personal moral judgements and

focussing on one evaluation for the sake of simplicity or as making a narrow claim specifically about the evaluations they pick out in their definitions. An alternative answer, then, to the question of which moral evaluations Internalism covers is the following:

*Narrow Judgement Internalism (Narrow Internalism):* Necessarily, if an agent makes a positive first personal moral judgement of sort S concerning an act that she could perform  $\phi$  then she will be *pro tanto* motivated to  $\phi$ . Necessarily, if an agent makes a negative first personal moral judgement of sort S concerning an act that she could perform  $\psi$  then she will be *pro tanto* motivated not to  $\psi$ .

According to this view, there is a necessary connection between some subset of positive and negative first personal judgements about what to do. There are, of course, a range of possible views about what the appropriate subset would be. One option would be to focus, as Cholbi and Frankena do, on moral requirements.

According to this view whenever an agent makes a first personal judgement about what she is required to do she will be motivated to act in line with her judgement.

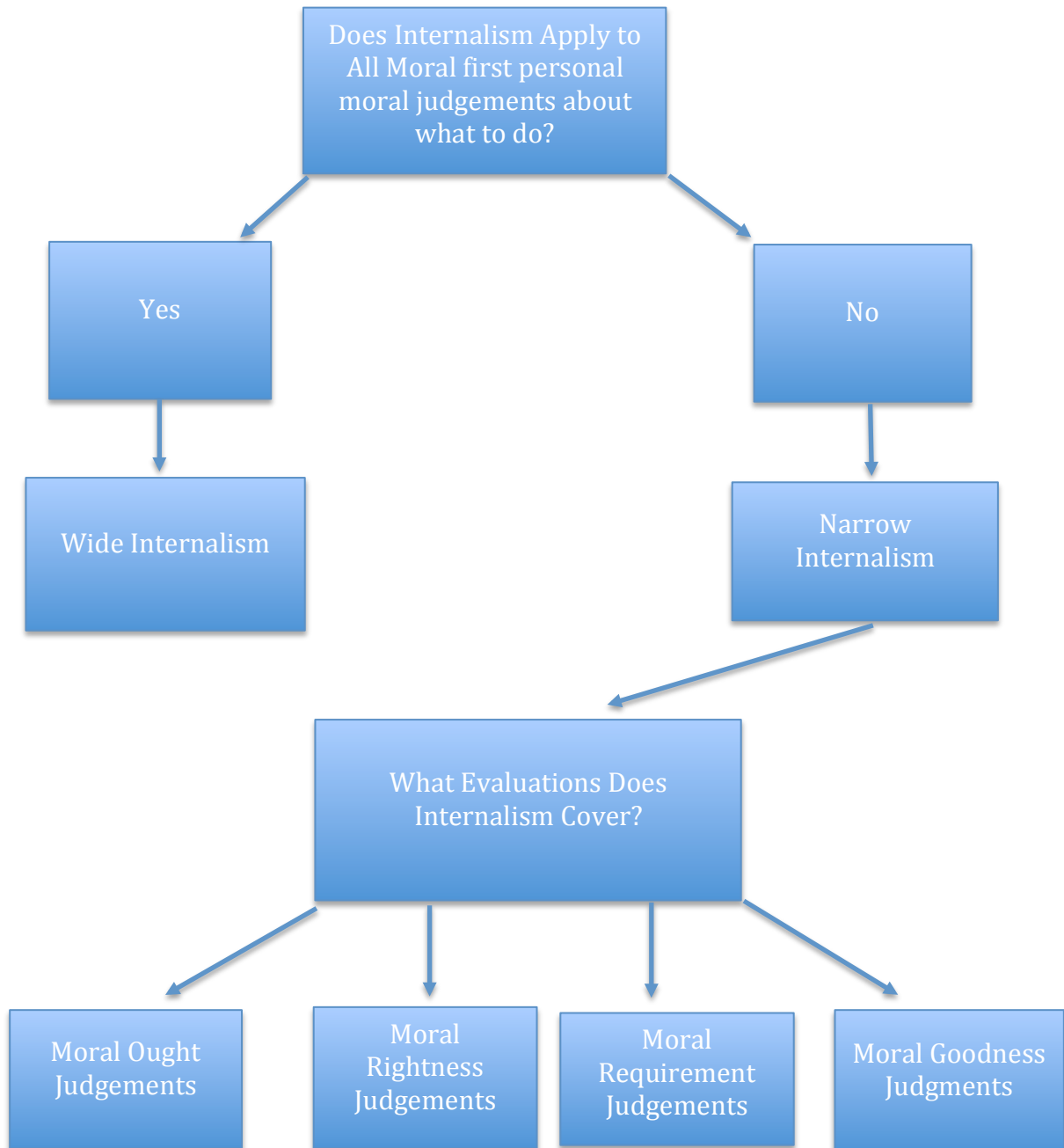
Alternatively, we might restrict the internalist claim to moral ought judgements, moral rightness judgements or moral goodness judgements. We might wonder what the difference between these different judgements is. This is a question that there is some disagreement about in the literature. Some philosophers take 'morally ought to' to be equivalent to 'morally required to' while others have argued that this is not how 'ought' should be understood.<sup>13</sup> Similarly, we might think 'morally right' is equivalent to 'morally required'. Alternatively we might take it to mean 'morally permissible'.<sup>14</sup>

We will return to this issue in §6, for now though it is enough to point out that for at least some philosophers these judgements may not be equivalent.



The following chart displays the options available to internalists:

**Fig. 1 Options for Internalists**



#### **IV. INTERNALISM AND THE SUPEREROGATION AMORALIST**

In the previous section I laid out a range of options for how internalists might answer the question of which moral evaluations their view covers. In this section I will

argue that we have good reason to reject any form of Internalism that is committed to the following view:

*Goodness Internalism:* Necessarily, if an agent judges that it would be morally good for her to  $\phi$  then she will be *pro tanto* motivated to  $\phi$ . Necessarily, if an agent judges that it would be morally bad for her to  $\psi$  then she will be *pro tanto* motivated not to  $\psi$ .

The reason I will argue that any view which endorses Goodness Internalism should be rejected is that any such view is committed to the implausible claim that there is a necessary connection between judgements of supererogation and motivation.

A supererogatory act is one that goes beyond the call of duty.<sup>15</sup> These acts are morally good but morally optional, meaning that they are neither required nor forbidden.<sup>16</sup>

Any internalist view that includes moral goodness judgements then is committed to the view that there is a necessary connection between supererogation judgements and motivation. This is a bad result, as positing a necessary connection to motivation for these judgements seems far less plausible than for other moral judgements.

To see the problem for these forms of Internalism consider the supererogation amoralist, someone who makes genuine supererogation judgments but remains completely unmotivated by them. This is problematic, as it seems possible to judge that an act is supererogatory and remain completely unmotivated to perform it.

Suppose Judy says the following: “it would be beyond the call of duty for me to give away all of my disposable income to charity.”<sup>17</sup> Now imagine we find out that she has no motivation to do so. This does not seem odd in the way that Jenny’s eating meat seemed odd. The reason for this is that we do not expect people to be motivated by a judgement that an act is supererogatory. Suppose we asked Judy why she isn’t acting

in line with her judgement. A perfectly reasonable response for Judy to make is that she judges it to be supererogatory not obligatory.

Let me say a little about why I take this to be a particular problem for Internalism and not simply another battle in what David Enoch has called “The Amoralist Wars,” (2011 p.250). The first reason is that, unlike traditional amoralist objections, being unmotivated by a supererogation judgement is likely to be a recognisable feature of many people’s moral experience. Certainly, when reflecting on my own experience of judging an act to be beyond the call of duty I would go so far as to say that the occasions where these judgements are accompanied by motivation are the exception rather than the norm. As I write this I judge that it would be morally good for me to stop writing and go and help out at a soup kitchen but that I have no duty to do so. I feel no motivation at all to go and help at the soup kitchen and do not find this odd. Of course, we do not need to go as far as asking whether or not someone who is never motivated by supererogation judgements exists. All we need to know is whether it is possible for someone to sometimes be unmotivated by supererogation judgements.<sup>18</sup> Nevertheless, the fact that this seems like a recognisable feature of moral experience means that internalists are going to have a difficult job persuading us that supererogation amoralists are impossible. One difference, then, between this challenge and traditional amoralist challenges is that supererogation amoralism is a condition that many ordinary agents will have experienced in their own lives.

The second reason why this is a particularly important problem for internalists is that the connection between moral language and motivation that is used to provide support for Internalism does not apply for talk of supererogation judgements. When Judy from the previous example utters the sentence, “Donating my disposable income to charity would be beyond the call of duty,” this sentence does not seem to bring with

it the presumption that Judy will be motivated to donate her money. If anything it seems that our expectation would be that she would not donate the money. Note, as we have already seen, the same is not true about obligations. If Judy were to say instead that she viewed donating the money as morally obligatory then it would be somewhat puzzling if she lacked motivation to do so. So, while there may be a connection between moral language and motivation for moral requirements the same does not appear to be the case for supererogation.

Accepting that supererogation amoralists exist gives us good reason to reject any internalist view that holds that there is a necessary connection between moral goodness judgements and motivation. After all, judging an act to be supererogatory involves judging the act to be morally good. If it is possible to make such a judgement without being motivated to perform the act in question then it is possible to judge an act to be morally good without being motivated to perform it. If this is the case then we should reject the claim that there is a necessary connection between moral goodness judgements and motivation. This means that we should reject both Wide Internalism and any narrower form of Internalism that includes goodness judgements within the set of judgements that are necessarily connected to motivation. Before we consider the options that this leaves open for internalists, let us first consider how an internalist might respond to this argument.

## **V. OBJECTIONS AND RESPONSES**

The important response to consider is whether some form of agent-restriction can deal with this problem. To test this possibility we must ask whether for any of the available agent restrictions that have been suggested for Internalism, it is true that

everyone who meets the requirements of the restriction and makes a supererogation judgement will necessarily be motivated to act in line with it. There are three forms of agent-restriction that have been proposed in the literature and I will look at each of these restrictions in turn. I will argue that none of these restrictions can provide a plausible response to the problem of the supererogation amoralist.

*A. Restrict Internalism To Rational Agents*

The first agent-restricted form of Internalism I will consider is one restricted to practically rational agents.<sup>19</sup> For this to be an effective response to the problem we will need to be persuaded that it is necessarily the case that in so far as an agent is acting in a practically rational way she will be motivated to perform acts that she judges to be supererogatory. I take it that many will share the thought that this is an implausible claim to make about supererogation judgements. It seems absurd to think that Judy, who judged she should give all her disposable income to charity but was not motivated to do so, is irrational if she is not motivated to act in line with her judgement. This I take to be the case whether we opt for a procedural or a substantive account of practical reasoning. Proceduralists think that we can only be rationally blameworthy for failing to have a desire we could have rationally reached from our existing set of beliefs and desires. According to substantivists we might be rationally blameworthy for failing to have some desire even if it would not have been possible for us to rationally reach it through our existing beliefs and desires.<sup>20</sup>

I take it to be uncontroversial that proceduralists about practical reason should accept that supererogation amoralists might not be irrational. There seems no reason to think that everyone who judges an act to be supererogatory could rationally reach the motivation to perform the act from his or her existing beliefs and desires. As a result,

internalists are going to have to look to a substantive view of rationality to declare supererogation amoralists as irrational.

The problem with this, though, is that it is plausible to think that many cases where an act of supererogation is available are cases where two or more rationally permissible options are available.<sup>21</sup> In fact this seems like a feature of supererogation that any plausible account must be capable of incorporating. If it is rationally permissible not to perform the supererogatory act then surely it will also be rationally permissible not to be motivated to perform the act. So declaring supererogation amoralists as irrational using a substantive account of irrational is not going to be an option for accounts that allow for rational options in cases where one or more available acts are supererogatory.

This objection connects to the second more general worry that others have voiced about Internalism. As a number of authors have pointed out it seems implausible to think that people will necessarily be motivated to some extent by a judgement that there is some reason to do something when this reason is overridden by other concerns.<sup>22</sup> To see why consider the following case, adapted from an example given by Caj Strandberg:

*Life vs. Coffee:* Clara has a serious illness and knows that she must undergo a medical procedure in order to save her life but that if she gets the treatment she will be unable to drink coffee for one minute, (2013 p.32).

Assuming that Clara likes drinking coffee, it seems that in this case she has some normative reason not to undergo the procedure, albeit one that is vastly outweighed by considerations that count in favour of doing so. According to Internalism about normative judgements, if Clara's judgement is a genuine one then she will necessarily

be motivated to some extent to avoid the procedure. Now of course it might be the case that Clara does have this motivation but there does not seem to be any reason to think that she *necessarily* will. There certainly seems nothing irrational about someone who made this judgement of the balance of reasons and had no motivation whatsoever to avoid the procedure.

To connect this point to our discussion, when an agent is faced with a choice between two morally good and morally optional alternatives there does not appear to be anything irrational about being completely unmotivated to perform one of the options. To see why consider the following case:

*Children vs. Donkeys:* John has five pounds he has set aside to donate to a charity. He judges that he has some moral reason to donate it to a charity that cares for donkeys but much stronger reason to donate it to a charity that feeds starving children. He also judges that either act would be morally good but not morally required.

In this case John judges he has some moral reason to perform one supererogatory act but much greater reason to perform another. Again there does not seem to be any reason to think that unless he is necessarily motivated to some extent to give the money to the donkey charity then he cannot be classed as a rational agent. After all, he is choosing to perform the act that he judges he has most reason to do. Why would it then be irrational if he were not motivated to perform an act that he judges he has less reason to do?

Another problem with attempting to declare supererogation amoralists as substantively irrational is that it seems to rule out as incoherent philosophical positions that appear to be intelligible.<sup>23</sup> Take, for instance, the view that Susan Wolf defends in her influential article 'Moral Saints' (1982). One of the claims that Wolf defends here

is that the life of a moral saint, someone who always acts in the morally best possible way, is an undesirable one. Presumably, those who accept this line of thought would also accept that it is not irrational to be unmotivated by a supererogation judgement. After all, those who are always motivated to perform acts they judge to be supererogatory will, according to Wolf, lead a worse life than those who are not.<sup>24</sup> Of course, not everyone endorses the view Wolf defends here and we might think that the force of this point is contingent upon its acceptance.<sup>25</sup> However, even if we are not persuaded by Wolf's claims, if we find her view intelligible then it is hard to see how it could be a *conceptual* truth that rational people will be motivated by their supererogation judgements. If this were the case then it would surely appear incoherent.

I have argued that restricting Internalism to rational agents will not solve the problem of the supererogation amoralist for three reasons. First, it seems to clash with the widespread view that cases of supererogation are cases where two or more rational options are available. Secondly, there does not appear to be anything irrational about someone who fails to be motivated to perform one act she judges to be supererogatory when another act that she judges to be supererogatory is judged by her to be morally better. Finally, responding to the problem in this way appears to render perfectly intelligible views as incoherent. For these reasons it seems reasonable to conclude that restricting the internalist claim to rational agents will not provide supporters of Goodness Internalism with an acceptable response to the possibility of supererogation amoralists.

*B. Restrict Internalism To Virtuous Agents*



The next agent-restricted form of Goodness Internalism I will consider is one restricted to virtuous agents.<sup>26</sup> In order for this to be an effective response to the problem it must be the case that virtuous people will always be motivated by their supererogation judgements. This claim strikes me as implausible. It seems perfectly possible for a virtuous agent to be unmotivated by a judgement that an act is supererogatory.

For those that don't immediately share this thought it is important to bear in mind that we are concerned with *judgements* of supererogation rather than *acts*. I am not claiming that it is possible for a virtuous agent never to be motivated to perform a supererogatory act. All I am saying is that a virtuous agent need not be motivated to perform any acts that she *judges* to be supererogatory. This is a much more plausible claim, as there may be some people we consider virtuous because they have a very demanding view of what is morally required. In fact, many people who do perform acts of supererogation claim to have only performed their duty.<sup>27</sup> Take the example of John Weidner, who during the Second World War put himself at great risk while working for the Resistance helping Jews escape the Gestapo. When asked whether he had performed a particularly worthy act his response was, "Absolutely not. I did my duty. That is all," (In Monroe, 2004 p.117). Many might think that Weidner is wrong about this and that his acts should be seen as supererogatory. Now let's imagine that Weidner does judge some acts to be supererogatory; perhaps he spent six days a week working for the resistance and judged that to spend a seventh would be supererogatory. If Weidner was occasionally left unmotivated by a supererogation judgement would that mean that he could not be classed as a virtuous person? Clearly not, after all he is motivated to perform acts that many people would consider to be

supererogatory; he just happens to think that they are obligatory. It is possible, then, for a virtuous agent to be unmotivated by a supererogation judgement.

Those sympathetic to this form of agent-restricted internalism might respond by saying that in this context we must understand ‘virtuous’ to mean ‘fully virtuous’.<sup>28</sup> This form of Internalism avoids the problem considered above, as it is compatible with those typically regarded as saintly failing to be motivated by their supererogation judgements. However, it does so at a price. The price is that this form of Internalism, insofar as it is true, no longer seems to be telling us anything interesting about our moral judgements in general. After all, it is possible that this form of Internalism may not apply even to exemplars such as Weidner. Given this, such a view would not seem to be telling us anything about moral judgements made by human agents. Rather, it is making a trivial claim about what we mean by ‘fully virtuous agent’. For these reasons, this form of agent restriction does not offer an adequate solution to the problem.

### *C. Restrict Internalism To Normal Agents*

The final form of restriction I will look at is restricting the claim to normal agents. According to this version of Internalism, moral judgements are necessarily connected to motivation for normal agents. If this form of restricted Internalism is plausible then we will be able to explain why it is possible for supererogation amoralists to exist. The existence of such people is compatible with this form of Internalism as they are not normal agents. On this account all normal agents will be necessarily motivated by their supererogation judgements.

This position is called ‘Communal Internalism’ by Jon Tresan (2009), as it is supported by focusing not on the individual making the judgment but on the

community that she is a part of. According to Communal Internalism, there must be some internal connection between moral judgments and motivation even if it is just one that is generally present in the community rather than in every individual member of that community.<sup>29</sup> Even if individual amoralists are conceivable, it is argued, a community of amoralists is not. Perhaps then, internalists can make a similar response to the conceptual possibility of supererogation amoralists. In response to such a problem it might be argued that, while it may be possible for an individual to remain unmotivated by a judgement of supererogation, it would not be possible for a community of such people to do so. The important issue to consider about such a world is whether or not we could say that the inhabitants are making genuine supererogation judgments given that people tend not to be motivated to act in line with them.

The problem for this attempt to accommodate the existence of supererogation amoralists is that it doesn't seem at all clear that normal people are motivated by their supererogation judgements. Of course, it isn't clear how we should understand 'normal' in this context.<sup>30</sup> However, I cannot see any way of understanding normal that is related to its ordinary use for which it would be plausible to think that normal people are generally motivated by supererogation judgements. A major part of what makes it so admirable to perform supererogatory acts is that most people would not have acted in the same way if they had been in that situation.<sup>31</sup> Of course, this is compatible with it nevertheless being the case that normal people are *pro tanto* motivated by their supererogation judgements. However, it is worth noting that claiming that normal people are *pro tanto* motivated to perform acts that they judge to be supererogatory is not enough to save this response to the supererogation amoralist. The internalist must endorse the stronger claim that it is *necessarily* the case that

normal people are motivated by their supererogation judgements. In other words, they must show that it is inconceivable that normal people would be unmotivated by supererogation judgements. This surely gives us good reason to reject this response.

Perhaps, though, the real challenge of Communal Internalism should not be put in terms of normality. After all we might be quite happy to describe judgements as moral judgements so long as *some* people are motivated by these judgements. James Lenman explains the intuitive appeal of this view in the following:

Story time. Externalism is true. There are moral facts to which one may, in principle, knowingly be indifferent. And there is a village called "*Amorality*" -- *no*, let's take "global" seriously and say a *planet* called "*Amorality*". Here they are really *very good indeed* at ascertaining these facts. *Much* better than we are. They're subject to a certain curiosity about what these are but otherwise they *don't care* about them. Once they have successfully detected the facts in question they make a nice scientific record of them but do not allow their discoveries about what these facts are to impinge in any way on their habits of desire and action. In the universities of Amorality there are departments, called "*morality departments*" where people called "*moral scientists*" are engaged full time to detect and record the moral facts. But this is an activity rather like the more refined branches of astrophysics or pure mathematics here on earth. It's *really jolly interesting* to know what these moral facts are but it's of no practical significance whatever, (1999 pp.445-446).

Lenman describes this as "a preposterous story," (1999 p.446). The reason why this is so ridiculous is that it doesn't look as if the inhabitants of this world are engaging in any kind of moral practice. As the practice that the inhabitants of this world engage in is in no sense connected to guiding action it simply cannot be regarded as a moral practice. This, according to Lenman, shows that even if we accept that individual amoralists are possible we should still reject Externalism. In order for any judgements to be considered moral they must be connected to motivation, even if only at the level of the community.

To relate this objection to our discussion of the supererogation amoralist, we might wonder if such a person is only possible against a background of at least *some* people

being motivated by supererogation judgements. However, unlike with moral judgements in general, there seems to be no reason to think that a judgment of supererogation can only count as genuine against a background of people being motivated to act in line with these judgements. We can see why if we imagine a world, Planet Duty, where no one is ever motivated by supererogation judgements. On this planet, the only moral judgements that motivate people are requiring ones. While people accept that there are acts they could perform that would be morally good but not morally required, this judgement never motivates them to perform such an act and indeed, has never motivated anyone to act in such a way. People can be heard saying things like, “Although it would be good to dedicate my life to helping the worst off, I am not required to do so.” If someone were to perform such an act they would certainly applaud it but they would never perform such an act themselves. It is hard to see how someone might argue that such a world is inconceivable. There does not appear to be anything about this world that would make us think that the inhabitants are not making genuine supererogation judgements. After all, it could well be the case that inhabitants may judge supererogatory acts to be good while judging that they involve too great a sacrifice to be morally required.<sup>32</sup> As a result, they remain entirely unmotivated to perform them. Unlike the planet *Amorality*, this planet does not seem at all preposterous. What this shows then, is that supererogation amoralists seem possible even if we accept Communal Internalism.

To sum up this section, I have argued that the possibility of the supererogation amoralist creates an important problem for the supporter of Goodness Internalism. I have considered three ways in which internalists typically respond to amoralist challenges and found that they do not succeed in providing an adequate response to this objection.

## **VI. ASSESSING THE INTERNALIST OPTIONS**

I have argued that we should reject any form of Internalism committed to the existence of a necessary connection between moral goodness judgements and motivation. If we accept this argument then we remove a number of options from the list of possible answers to the question of what moral judgements Internalism might apply to. Most obviously we should reject Narrow Internalism about moral goodness judgements. In addition, we should also reject Wide Internalism, as this is also committed to Goodness Internalism. On the other hand, Internalism about moral requirement judgements is able to avoid the problem of the supererogation amoralist. If we understand Internalism to apply only to requiring judgements then the existence of people who are unmotivated by supererogation judgements is unproblematic. After all, judging that an act is supererogatory involves judging that it is *not* required.

The case is less clear for the other options for Internalism. What about Internalism about moral ought judgements and moral rightness judgements? As we have already seen, some philosophers take ‘morally ought to’ to be equivalent to ‘morally required to’ while others have argued that this is not how ‘ought’ should be understood.<sup>33</sup>

Similarly, we might think ‘morally right’ is equivalent to ‘morally required’.

Alternatively we might take it to mean ‘morally permissible’.<sup>34</sup> For my purposes it does not matter which of these interpretations of ‘ought’ and ‘rightness’ we accept. It is enough for my purposes to point out that if ‘morally ought’ and ‘morally right’ are understood as being equivalent to ‘morally required’ then a form of Internalism that applies to these judgements will avoid the problem of the supererogation amoralist. However, if they are understood to be equivalent to ‘morally permissible’ then they will be vulnerable to this objection.

In effect, then, we are left with one form of Internalism that can avoid the problem of the supererogation amoralist. Only Narrow Internalism about moral requirements is uncommitted to the existence of a necessary connection between supererogation judgements and motivation. We can define this view in the following way:

*Requirement Internalism:* Necessarily, if an agent judges that she is morally required to perform  $\phi$  then she will be *pro tanto* motivated to  $\phi$ . Necessarily, if an agent judges that she is morally required not to perform  $\phi$  then she will be *pro tanto* motivated not to  $\phi$ .

Narrow Internalism about other forms of moral judgement may also avoid this problem but only if these judgements are understood as being equivalent to Requirement Internalism.

## **VII. REQUIREMENT INTERNALISM AND MORAL RATIONALISM**

In the previous section I argued that only Requirement Internalism is able to avoid the problem of the supererogation amoralist. However, supporters of this view also face a problem. Why is it that there is a necessary connection between requiring moral judgements and motivation but not for other moral judgements? Given that supporters of this view wish to restrict the internalist claim to a narrow subset of our moral judgements, they need to provide some form of explanation as to why this is.

Of course, one explanation that could be given for restricting the claim in this way is that it allows internalists to avoid the problem of the supererogation amoralist. This is an unacceptable explanation. The problem of the supererogation amoralist may make us question whether we should be internalists at all. If we have to restrict Internalism to only a narrow subset of moral judgements then perhaps we might be tempted to

simply abandon it altogether. What we are looking for, then, is a reason to think that some moral judgements but not others are necessarily connected to motivation. We are looking for the internalist to tell us what it is about moral requirement judgements that makes it plausible to think that they are necessarily connected to motivation when other moral judgements are not. Simply stipulating that this is the case in order to avoid the problem of the supererogation amoralist is not an acceptable response to this problem.

The good news for internalists is that there is an explanation that internalists can appeal to here. Both Michael Smith and Mark Van Roojen argue that a form of Internalism follows from the following view:

*Moral Rationalism:* If an act,  $\phi$ , is morally required then, from an all things considered normative perspective,  $\phi$ -ing is in line with the balance of reasons. If an act,  $\psi$ , is morally wrong then, from an all things considered normative perspective,  $\psi$ -ing is against the balance of reasons (See Smith, 1994 pp. 61-62 and Van Roojen, 2010).<sup>35</sup>

This view is one that many philosophers have found attractive, as it appears to follow from our practice of blaming those who fail to act in line with their moral requirements. If we accept that blame is inappropriate for agents who have acted in line with what they had most reason to do and we accept that blame is always appropriate for moral wrongdoers who lack an excuse then it follows that moral wrongdoing always involves acting against the balance of reasons.<sup>36</sup>

Both Smith and Van Roojen argue for Requirement Internalism by appealing to Moral Rationalism. The essence of the argument is that if moral requirements are requirements of practical reason then this provides support for a form of Internalism



restricted to rational agents. The reason for this is that if we accept Moral Rationalism then we should accept that judging an act to be morally required involves judging that one has most reason all things considered to perform that act.<sup>37</sup> This means that an agent's failure to be motivated by a moral requirement judgment is a failure to be motivated to act in the way that she has judged that she has most reason to act. If she fails to be motivated by this judgement then she is irrational, as she has acted against her judgement of what she has most reason to do.

As Strandberg notes, in order to move from Moral Rationalism to a form of Internalism about moral judgements we need to accept the following form of Internalism about normative judgements in general:

*Normative Judgement Internalism:* If a rational agent judges that  $\phi$ -ing is what there is most reason to do then she will necessarily be motivated to  $\phi$ . If a rational agent judges that  $\phi$ -ing is against the balance of reasons then she will necessarily be motivated not to  $\phi$ .<sup>38</sup>

Unless we accept the truth of this form of Internalism then the argument from Moral Rationalism to Requirement Internalism will be unsuccessful. In the absence of this form of Internalism there would be no reason to think that judging an act to be what there is most normative reason to do would necessarily motivate rational agents.

There is then a response internalists can give to the question of why some moral judgements but not others are necessarily connected to motivation. The response an internalist can offer is that there is a special connection between moral requirements and what there is most reason to do, all things considered. This connection is not one that applies to other moral judgements. This, then, is why it is only moral requiring judgements that are necessarily connected to motivation.

For our purposes, it is important to note two implications about the form of Internalism that this argument vindicates. First, if we accept this argument then Internalism is not a special feature of moral judgements. Rather it is a feature shared by all conclusive normative judgements. Given that the argument from Moral Rationalism to an agent-restricted Requirement Internalism depends upon the truth of Normative Internalism, the internalist must accept that it is all conclusive normative judgements that are necessarily connected to motivation. This is unlikely to worry many internalists and indeed is a point that many internalists already accept.<sup>39</sup>

The more interesting point to note is that appealing to this explanation leaves the plausibility of Internalism dependent on the plausibility of Moral Rationalism. Up to now, those who have posited a connection between the two views have claimed that if we accept Moral Rationalism then we have reason to accept Internalism as well. This claim does not tell us whether accepting Internalism gives us any reason to accept Rationalism, and indeed those who argue from Moral Rationalism to Internalism do not think that the argument works the other way.<sup>40</sup> We might think then that while rationalists should be internalists there is no need for internalists to be moral rationalists. However, the problem of how to explain why it is only requiring moral judgements that are necessarily connected to motivation changes the relationship between Internalism and Moral Rationalism. By appealing to Moral Rationalism in order to respond to this objection the internalist makes the plausibility of her view dependent on the plausibility of Moral Rationalism. This is important, as Moral Rationalism is a controversial view that is far from universally accepted.<sup>41</sup> This response, then, leaves internalists vulnerable to any objections that may be raised against Moral Rationalism.

Of course, internalists do not need to respond to the problem in this way. I have not shown that this is the only way for internalists to explain why there is a necessary connection between requiring moral judgements that does not exist for other moral judgements. Perhaps some other explanations can be given. What we can say, though, is that the upshot of this discussion is that the internalist is left with a choice to make. Either she accepts that the plausibility of her view is dependent on the truth of Moral Rationalism or she provides an alternative response to the question of why it is that only requiring moral judgements are necessarily connected to motivation.

Prior to this section, I argued that the problem of the supererogation amoralist gives us reason to reject all forms of Internalism other than Requirement Internalism. In this section I have looked at how internalists can respond to the question of why it is only requiring moral judgements that are necessarily connected to motivation. One explanation that might be offered is that moral requirements are necessarily rational requirements. However, if internalists respond to this problem in this way then they leave the plausibility of their view dependent on the plausibility of Moral Rationalism.

### **Conclusion**

In this paper I have looked at what a supporter of Motivational Judgement Internalism can say about judgements of supererogation. I have argued that the problem of the supererogation amoralist gives us good reason to think that Internalism cannot plausibly be held to cover moral goodness judgements. As a result, we should restrict the internalist claim to requiring moral judgements. However, some explanation needs to be given for this restriction in order for this form of Internalism to be plausible. I pointed out that Moral Rationalism provides us one way of explaining this restriction. However, by appealing to this explanation internalists leave

the plausibility of their view dependent on the plausibility of Moral Rationalism. Internalists, then, must either embrace Moral Rationalism and defend the truth of this view or provide an alternative explanation for as to why requiring moral judgements are necessarily connected to motivation while other moral judgements are not.<sup>42</sup>

## **Bibliography**

Archer, Alfred. 2014. "Moral Rationalism Without Overridingness." *Ratio* 27: 100-114.

Archer, Alfred and Ridge, Michael. 2015. "The Heroism Paradox: Another Paradox of Supererogation." *Philosophical Studies* 172: 1575-1592.

Badhwar, Neera Kapur. 1993. "Altruism Versus Self-Interest: Sometimes A False Dichotomy." *Social Philosophy and Policy* 10: 90-117.

Björnsson, Gunnar. & Olinder, Ragnar Francén. 2013. "Internalists Beware - we Might all be Amoralists!" *Australasian Journal of Philosophy* 91: 1 - 14.

Blackburn, Simon. 1984. *Spreading the Word* Oxford: Clarendon Press.

Blackburn, Simon. 1998. *Ruling Passions* Oxford: Clarendon Press.

Brink, David. 1989. *Moral Realism and the Foundations of Ethics* Cambridge: Cambridge University Press.

Carbonell, Vanessa. 2009. "What Moral Saints Look Like." *Canadian Journal of Philosophy* 39: 371-398.

Carbonell, Vanessa. 2012. "The Ratcheting-up Effect." *Pacific Philosophical Quarterly* 93: 228-254.

Chisholm, Roderick. 1963. "Supererogation and Offence: A Conceptual Scheme for Ethics." *Ratio* 5: 1-14.

Cholbi, Michael. 2011. "Depression, Listlessness, and Moral Motivation." *Ratio* 24: 28-45.

Chrisman, Matthew. 2012. "'Ought' and Control." *Australasian Journal of Philosophy* 90: 433-451.

Colby, Anne and Damon, William. 1992. *Some Do Care*. New York: Free Press.

Copp, David. 2001. "Realist-Expressivism: A Neglected Option for Moral Realism." *Social Philosophy and Policy* 18: 1-43.

Dancy, Jonathan. 1993. *Moral Reasons* Oxford: Blackwell.

Dancy, Jonathan. 2004. *Ethics Without Principles* Oxford: Oxford University Press.

Darwall, Stephen. 2006. *The Second Person Standpoint* Cambridge MA: Harvard University Press.

Dorsey, Dale 2012. "Weak Anti-Rationalism and the Demands of Morality." *Noûs* 46: 1-23.

Dreier, Jamie. 1990. "Internalism and Speaker Relativism." *Ethics* 101: 6-26.

Enoch, David. 2011. *Taking Morality Seriously* Oxford: Oxford University Press.

Finlay, Stephen. 2010. "What Ought Probably Means, and Why You Can't Detach It." *Synthese* 177: 67-89).

Fletcher, Guy and Ridge, Michael (eds.). 2015. *Having it Both Ways: Hybrid Theories and Modern Metaethics*. Oxford: Oxford University Press.

Frankena, William Klaas. 1958 "Obligation and Motivation in Recent Moral Philosophy." In *Essays in Moral Philosophy*, ed. Abraham Irving Melden. Seattle: University of Washington Press.

Gert, Joshua. 2004. *Brute Rationality: Normativity and Human Action* Cambridge: Cambridge University Press.

Gibbard, Allan. 1990. *Wise Choices, Apt Feelings: A Theory of Normative Judgment*. Cambridge MA: Harvard University Press.

Gibbard, Allan. 2003. *Thinking How To Live* (Cambridge, Mass.: Harvard University Press.

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

Heyd, David. 2011. "Supererogation". In *The Stanford Encyclopedia of Philosophy* (Winter 2011 Edition), ed. Edward N. Zalta, URL = <http://plato.stanford.edu/archives/fall2008/entries/supererogation/>.

Hooker, Brad and Streumer, Bart. 2004. "Procedural and Substantive Practical Rationality." In *The Oxford Handbook of Rationality*, eds. Alfred Mele and Piers Rawling. Oxford: Oxford University Press.

Horgan, Terence and Timmons, Mark. 2010). "Untying a Knot from the Inside Out: Reflections on the 'Paradox' of Supererogation." *Social Philosophy and Policy* 27: 29–63.

Jacobs, Russell A. 1987. "Obligation, Supererogation and Self-sacrifice." *Philosophy* 62: 96-101.

Kauppinen, Antti. 2008. "Moral Internalism and the Brain." *Social Theory and Practice* 34: 1-24.

Kawall, Jason. 2003. "Self-Regarding Supererogatory Acts." *Journal of Social Philosophy* 34: 487-498.

Korsgaard, Christine. M. 1996. *The Sources of Normativity*. Cambridge: Cambridge University Press.

Lenman, James. 1999. "The Externalist and the Amoralist." *Philosophia* 27: 441-457.

Manne, Kate. 2015. "Tempered Internalism and the Participatory Stance"  
In *Motivational Internalism*, eds, Gunnar Björnsson, Caj Strandberg, Ragnar Francén  
Olinder, John Eriksson, and Fredrik Björklund. New York: Oxford University Press,  
260-281.

Mason, Elinor. 2008. "An Argument Against Motivational Internalism." *Proceedings of the Aristotelian Society* 108: 135-154.

McGoldrick, Patricia. 1984. "Saints and Heroes: A Plea for the Supererogatory." *Philosophy* 59: 523- 528.

McNamara, Paul. 1996. "Must I Do What I Ought? (Or Will the Least I Can Do Do?)" In *Deontic Logic, Agency and Normative Systems*, eds, Mark A. Brown and Jose Carmo Berlin: Springer-Verlag pp.154- 173).

McNaughton, David. 1988. *Moral Vision* Oxford: Blackwell.

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

Miller, Christian B. 2008. "Motivational Internalism." *Philosophical Studies* 139: 233-255).

Monroe, Karen. 2004. *The Hand of Compassion: Portraits of Moral Choice During the Holocaust*. Princeton NJ: Princeton University Press.

Nichols, Shaun. 2002. "How Psychopaths Threaten Moral Rationalism." *The Monist* 85: 285-303.

Portmore, Douglas. 2011. *Commonsense Consequentialism*. Oxford: Oxford University Press.

Raz, Joseph. 1975. "Permissions and Supererogation." *American Philosophical Quarterly* 12: 161-168.

Ridge, Michael. 2014. *Impassioned Belief* Oxford: Oxford University Press.

Ridge, Michael. (2015) "Internalism: Cui Bono?" In *Motivational Internalism*, eds, Gunnar Björnsson, Caj Strandberg, Ragnar Francén Olinder, John Eriksson, and Fredrik Björklund. New York: Oxford University Press, 135-149.

Ross, William David. 1930. *The Right and The Good* Oxford: Clarendon Press.

Scanlon, Thomas Michael. 1998. *What We Owe To Each Other* Cambridge MA: Harvard University Press.

Shafer-Landau, Russ. 2003. *Moral Realism: A Defence* Oxford: Oxford University Press.

Smith, Michael. 1994. *The Moral Problem* Oxford: Blackwell.

Stevenson, Charles. 1944. *Ethics and Language* New Haven: Yale University Press.

Stranberg, Caj. 2012. "A Dual Aspect Account of Moral Language." *Philosophy and Phenomenological Research* 84: 87-122.

Strandberg, Caj. 2013. "An Internalist Dilemma - and an Externalist Solution." *Journal of Moral Philosophy* 10: 25-51.



Svavarsdóttir, Sigrun. 1999. "Moral Cognitivism and Motivation." *Philosophical Review* 108: 161-219.

Timmons, Mark. 1999. *Morality Without Foundations* Oxford: Oxford University Press.

Tresan, Jon. 2009. "The Challenge of Communal Internalism." *Journal of Value Inquiry* 43: 179-199

Van Roojen, Mark. 2010. "Moral Rationalism and Rationalist Amoralism." *Ethics* 120: 495-525.

von Fintel, Kai and Iatridou, Sabine. 2008. "How to Say *Ought* In Foreign: The Composition of Weak Necessity Modals." In *Time and Modality*, eds, Jacqueline Guéron and Jacqueline Lecarme, New York: Springer.

Wallace, R. Jay. 2006. "Moral Motivation." In *Contemporary Debates in Moral Theory*, ed, Jamie Dreier, Oxford: Blackwell, 182-195.

Wedgewood, Ralph. 2007. *The Nature of Normativity* Oxford: Oxford University Press.

Wolf, Susan. 1982. "Moral Saints." *Journal of Philosophy* 79:419-439.

Zangwill, Nick. 2008. "The Indifference Argument." *Philosophical Studies* 138: 91 – 124.

Zimmerman, Michael. J. 1996. *The Concept of Moral Obligation* Cambridge: Cambridge University Press.

---

<sup>1</sup> See, for example, Copp (2001), Ridge (2014), Strandberg (2012) and the papers collected in Fletcher and Ridge (2015).

<sup>2</sup> This way of characterizing the appeal of Internalism comes from Strandberg (2012 p.89).

<sup>3</sup> See Stevenson (1944 pp.16-17), Dreier (1990 pp.13-14), Dancy (1993 p.4), Smith (1994 p.60), Blackburn (1998 pp.48, 52-53) and Lenman (1999 pp. 443-446).

<sup>4</sup> Brink first made the distinction between Strong Internalism, which holds that moral judgments provide sufficient motivation to act and Weak Internalism, which holds that moral judgments necessarily provide some motivation to act (1989 p.41).

<sup>5</sup> See Brink (1989); Svavarsdóttir (1999); Shafer Landau (2003); Zangwill (2008).

<sup>6</sup> It is worth noting, though, that such arguments leave room for an *a posteriori* necessary connection between moral judgements and motivation. See Bjornsson and Olinder (2013)

<sup>7</sup> Miller makes this point convincingly about versions of Internalism that are restricted to virtuous agents. This, Miller points out, “Might be the case simply because of what it is to be a ‘virtuous agent’ in the first place,” (2008 p.252). Similar points about weak versions of Internalism are made by Enoch (2011 p.251), Mason (2008) and Svavarsdóttir (1999 p.183).

<sup>8</sup> Notable exceptions include Miller (2008 p.235) Ridge (2015) and Tresan (2009) who acknowledge this gap in the literature.

<sup>9</sup> See, for example, Allan Gibbard’s (1990) view that moral judgements are connected to various kinds of emotions.

<sup>10</sup> See Cholbi (2011 p.28), Frankena (1958) and Korsgaard (1996 p. 81);

<sup>11</sup> Though some virtue and vice judgements may be connected to first personal judgements about what to do. For example, the judgement that it would be courageous to save someone from a burning building. Though I will not be considering these judgements in this paper, I take it that what I have to say about moral goodness judgements would apply to these judgements as well. Thanks to an anonymous referee for pressing me on this point.

<sup>12</sup> The inclusion of the ‘she could perform’ clause is intended to rule out cases where an agent makes a judgement about an act that she is unable to perform. Thanks to an anonymous referee for pushing me on this point.

<sup>13</sup> Zimmerman uses the term ‘ought’ to refer to moral obligation (1996 p.1). Chrisman (2012), on the other hand, points to some problems with understanding the moral ‘ought’ in terms of obligation. More general worries about the identification of ‘ought’ with ‘is required to’ are raised by McNamara (1996), von Fintel and Iatridou (2008), Finlay (2010 p.76) and Ridge (2014).

<sup>14</sup> This claim is made by Ross (1930 p.6). We might equally think, as Scanlon does (1998), that ‘morally right’ means ‘morally permissible’.

<sup>15</sup> For a discussion of the concept see Heyd (1982).

<sup>16</sup> At least not as it is usually understood. Some have sought to defend a ‘Qualified’ account of supererogation in which supererogatory acts are required in a weak sense. See for example, Raz (1975 p.165). However, as Zimmerman argues, given that we think of supererogatory acts as being *beyond* duty it seems that the usual understanding of the term is that they are not required (1996 p.235). Thanks to Richard Rowland for pushing me to consider this point.

<sup>17</sup> It could be objected that Judy has not said that she judges this act to be *supererogatory*. Rather, she has judged that it is *beyond the call of duty*. However, supererogation is generally accepted to be a technical term that is equivalent to the ordinary language phrase ‘beyond the call of duty’. See, for example, Heyd (2011) and Mellema (1991 p.16). Thanks to an anonymous referee for pushing me to clarify this point.

<sup>18</sup> Zangwill says that we should not describe people who are indifferent on a particular occasion as amorlists, as amorlists are people who are never motivated by moral judgements (2008 p.101). To be clear, when I use the term supererogation amoralist I mean someone who on that particular occasion is left unmotivated by a

supererogation judgement. Those who think that this is a misuse of the term can replace ‘supererogation amoralist’ with ‘someone indifferent to their supererogation judgement’.

<sup>19</sup> This form of restriction is advocated by Smith (1994 p.61), Wallace (2006) and Van Roojen (2010).

<sup>20</sup> For a discussion of the literature on Proceduralism and Substantivism about practical reason see Hooker and Streumer (2004). Thanks to Elinor Mason for helpful discussion here.

<sup>21</sup> This claim is made by Gert (2004 p.106), Horgan and Timmons (2010) and Portmore (2011 p.153).

<sup>22</sup> Those who make this point include Dancy (2004 p.17), Wallace (2006 pp. 187-188), Strandberg (2013 p.32).

<sup>23</sup> Thanks to Mike Ridge for suggesting this point.

<sup>24</sup> Of course we might think that the moral saints Wolf describes would judge their acts to be obligatory and so would have no need to be motivated by supererogation judgements. This may well be the case for many of the saints Wolf has in mind. However, I take it that Wolf’s view would also apply to a saint who is always motivated to perform the morally best acts she can, even when she judges them to be supererogatory. Thanks to an anonymous referee for pressing me on this point.

<sup>25</sup> One way to put pressure on this view is to reject this account of ‘moral sainthood’ as Carbonell (2009) does.

<sup>26</sup> This form of agent restriction is considered, though not defended, by Miller (2008 p.250-252). Though I know of no one who defends this form of Internalism it seems worth considering here, as it might be thought to be particularly relevant to the case of supererogation.

<sup>27</sup> This phenomenon is identified by Badhwar (1993), Carbonell (2012) and Colby and Damon (1992). I discuss a puzzle that arises from this phenomenon in Archer and Ridge (2015).

<sup>28</sup> Thanks to Mike Ridge for pressing me on this point.

<sup>29</sup> See Dreier (1990 p.14); Blackburn (1998 pp. 61-68); Lenman (1999) and Gibbard (2003).

<sup>30</sup> We might understand it in a purely statistical way or in a more normative way. Of course, if understood in a statistical way then this seems to provide no support for the internalist claim, as Internalism is supposed to be an *a priori* thesis. This point is made by Miller (2008 p.245).

<sup>31</sup> To be clear, I do not mean that most people never perform supererogatory acts, only that they do not perform all of the supererogatory acts available.

<sup>32</sup> Sacrifice is often taken to be a necessary condition for supererogation. See Dancy (1993 p. 139), Jacobs (1987 p.101), McGoldrick (1984 p.525) and Raz (1975 p.167)). Although, this claim is not universally accepted (See for example Kawall (2003) and Horgan and Timmons (2011)) it at least seems possible to imagine a world where all acts of supererogation do in fact involve agential sacrifice.

<sup>33</sup> See note 13.

<sup>34</sup> See note 14.

<sup>35</sup> This view is sometimes confused with the related claim that morality is overriding. For an explanation of the differences between the two see Archer (2014).

<sup>36</sup> This is a simplified version of the argument given by Portmore (2011 pp.43-44) and Darwall (2006 pp.95-99).

<sup>37</sup> This is roughly the argument Smith gives (1994 pp. 61-62). However, as Van Roojen notes this is too simplistic, even if we take Moral Rationalism to be true, it is possible for rational agents to be unaware of its truth. Such an agent cannot be said to be irrational if she fails to be motivated by her obligation judgement. Nevertheless, Van Roojen's argument argues that it is in the normal cases, those where the judgement that an action is right motivates the agent, that fix the meaning of the term 'right' and that this allows us to conclude that rational agents acting normally will be motivated by a judgement that an act is morally required (2010 pp.518-521).

<sup>38</sup> This is similar to the form of Normative Internalism considered by Strandberg (2013 p.32). A version of this form of Internalism is also endorsed by Wedgwood (2007 p.25).

<sup>39</sup> See, for example, Van Roojen (2010 p.501), Ridge (2015 p.141).

<sup>40</sup> See Smith (1994 p. 62) and Van Roojen (2010 p.501), although it is worth noting that Van Roojen does think that Internalism 'provides evidence for' Moral Rationalism.

<sup>41</sup> For arguments against Moral Rationalism see Dorsey (2012) and Nichols (2002).

<sup>42</sup> Thanks to audiences at The University of Edinburgh's Philosophy Work in Progress Seminar and the 2012 British Postgraduate Philosophy Association Annual Conference. Special thanks to Guy Fletcher, Elinor Mason, Rae Langton, Mike Ridge, Richard Rowland and two anonymous referees for detailed comments on early drafts of this paper.