Acting for the Right Reasons

Abstract. This paper examines the thought that our right actions have moral worth only if we perform them for the right reasons. I argue against the traditional Kantian view that morally worthy actions must be performed because they are right, and argue that Kantians ought instead to accept the view that morally worthy actions are those performed for the reasons *why* they are right. In other words, morally worthy actions are those for which the reasons why they were performed (the reasons *motivating* them) and the reasons why they ought to have been performed (the reasons *justifying* them) coincide. I call this the *Coincident Reasons Thesis*, and argue that it provides plausible necessary and sufficient conditions for morally worthy action, defending the claim against proposed counterexamples.

1   The Motive of Duty

Kant writes in the "Preface" to the *Groundwork* that "what is to be morally good … must … be done *for the sake of the law*."[1] He infamously claims that when people "without any other motive of vanity or self-interest … find an inner satisfaction in spreading joy around them," their action, "however amiable it may be, has nevertheless no true moral worth."[2] Only when a good action is performed "without any inclination, simply from duty" does it "first ha[ve] its genuine moral worth."[3] This thesis, which we might call the Motive of Duty Thesis, is one of the less popular elements of Kant's ethics. Kant*ians* have largely responded to it by attempting to make the pill easier to swallow. They have suggested, for example, that there may be overdetermination of our actions by different motivations or incentives—that we may act on the motive of duty *while* also feeling sympathy for others.[4]

I hope to argue that a Kantian need not—indeed, should not—accept the Motive of Duty Thesis. I will put forward a more appealing version of the more general thought that right actions are morally worthy only if they are performed for the right reasons—

one which Kantians can accept without giving up on the more attractive elements of Kantian ethics.

This more general thought concerns *motivating* reasons—reasons for which someone acts—as opposed to *justifying* (or normative) reasons—reasons that determine how someone ought to act. Morally worthy actions (the thought is) aren't just right actions—they are actions for which the agent who performs them merits praise. When we do the right thing because it happens to suit us, or happens to be in our interest, our action has no moral worth. We are not deserving of moral praise if we save a drowning child merely for the sake of claiming the anticipated reward. This is intuitive. Morally worthy actions must be performed for the right (motivating) reasons. I'll call this general thought the Right Reasons Thesis. Which motives *can* endow actions with moral worth? The Motive of Duty Thesis provides one answer to this question: a morally worthy action is one performed "out of respect for the moral law", or, more simply, *because it is right*.[5]

I will argue that the Motive of Duty Thesis excludes some apparently admirable actions from having moral worth. As other critics have noted, it also seems to misidentify what's admirable about the actions it does pick out as morally worthy. The passages from the *Groundwork* with which I began help emphasize the unpalatability of the Motive of Duty Thesis. The Kantian 'truly moral man' seems at best guilty of a kind of moral fetishism (to borrow a phrase from Michael Smith)[6], if not plainly cold. A morally attractive person, objectors maintain, will help others not 'because the moral law demands it,' but because they are in need of help. This is the line of objection Philippa Foot favors. "It will surely be allowed," she writes,

that quite apart from thoughts of duty a man may care about the suffering of others, having a sense of identification with them, and wanting to help them if he can. Of course he must want not the reputation of charity, nor even a gratifying role helping others, but, quite simply, their good. If this is what he does care about, then he will be attached to the end proper to the virtue of charity and a comparison with someone acting from an ulterior motive … is out of place.[7]

Foot's words suggest a version of the Right Reasons Thesis that is not equivalent to the Motive of Duty Thesis: according to this version, morally worthy actions are those not performed for an "ulterior motive." Ulterior motives are, presumably, those generated by facts that are not morally relevant features of the situation in which we act. This version of the thesis simply suggests that when our actions have moral worth, our motivating reasons for acting will be given by features of our situation that are morally relevant. Morally relevant features are those facts about a situation that justify a conclusion about what should be done—that provide *justifying* reasons for action.

When I am faced with a practical decision—for example, when I must decide whether to jump into the water to save a drowning boy—there are many features of the situation that may be morally relevant. The endangered well-being of the boy is relevant, as is the risk posed to my own well-being. When I am motivated by concern for either of these, and not in excess of their moral relevance, then I cannot be accused of acting for an "ulterior purpose." When, however, I am motivated to save the boy solely by a desire to claim the anticipated reward—a feature of the situation that has little or no moral relevance—I am acting for an ulterior purpose, and my action has no moral worth. My motivating reason for acting was not also a significant justifying reason: it was not the prospect of reward that made saving the boy the right thing to do.

Foot seems to think a Kantian – or more generally, anyone interested in defending the view that moral imperatives are categorical (in other words, that anyone has sufficient reason to comply with them) – is committed to something like the Motive of Duty Thesis. Her reasons for thinking this are somewhat obscure, and I'll omit investigation of them now in the interests of saving time. (I'd be happy to come back to the problem in the discussion.)[8] In any case, I'll argue in a moment that the thesis sits *un*comfortably with some central elements of Kantian ethics. As someone sympathetic to Kant's approach, and to the doctrine of the categorical imperative, I ought not accept the Motive of Duty Thesis. I can instead accept a much more plausible version of the Right Reasons Thesis—one the discussion of the example I've just given suggests. According to this version, which I will call the **Coincident Reasons Thesis**, *my actions have moral worth if and only if my motivating reasons for acting coincide with the reasons justifying the action*—that is, if and only if I perform the actions I ought to perform, for the reasons why they ought to be performed.[9] My motivating reason for performing some action in this case will not be the duty-based reason "that the moral law requires it" but the reasons *for which* the moral law requires it.

The Motive of Duty Thesis gained what attraction it held from the plausibility of the thought that morally worthy actions don't just *happen* to conform to the moral law— as a matter of mere accident. There must be some stronger, more reliable connection between the rightness of such actions and their performance. It may have seemed a natural step from this observation to the conclusion that the rightness of such actions itself must be the motive for their performance. Kant himself seems to make this assumption in the "Preface" to the *Groundwork*. He writes:

> in the case of what is to be morally good it is not enough that it *conform* with the moral law but it must also be done *for the sake of the law*; without this that conformity is only very contingent and precarious….[10]

But we have seen that the virtuous agent's actions track the requirements of morality even if he does not act for the reason "that the moral law requires it," but acts instead for the reasons that *make* an act morally required.

3 The Coincident Reasons Thesis

In the passages from the *Groundwork* that I quoted earlier, Kant seems to endorse the Motive of Duty Thesis.[11] But the formula of humanity from Section II of the *Groundwork*—Kant's most appealing formulation of the categorical moral law, and an element of Kant's ethics I am much more interested in preserving—suggests a version of the Right Reasons Thesis that is closer to the Coincident Reasons Thesis. Kant's formula states: "*so act that you use humanity, whether in your own person or in the person of any other, always at the same time as an end, never merely as means.*"[12] It seems natural to read the formula as supplying a test for the wrongness of actions. Though Kant may have intended it as such a test, this would, as Derek Parfit has pointed out, have been an oversight on his part. Not all violations of the formula of humanity constitute wrong actions. Sometimes when we treat people merely as a means, our *actions* are not wrong, although our *attitude* is.[13] We have seen already that if I save a drowning boy as a means to claiming the reward, my action is right, although it has no moral worth. I regard the boy as a mere means, and so fail to comply with the formula of humanity. Read as a test of the moral worth or disvalue of actions, the formula of humanity looks more like the Coincident Reasons Thesis than the Motive of Duty Thesis. It states that our actions have moral worth only if our treatment of others is governed by our recognition of their status

as ends in themselves—as beings with unconditional value. My act of saving the drowning boy is right because he has this value: this is the reason *justifying* my act. When I am motivated by my recognitions of this value—when, that is, I regard him as an end, and not a mere means—then the reasons motivating my act coincide with the reasons justifying it.

There is a deeper reason why Kantians should accept the Coincident Reasons Thesis, as Philip Stratton-Lake has pointed out. Stratton-Lake defends a thesis that is similar to the Coincident Reasons Thesis, this time phrased not as a condition for the moral worth of actions, but rather for the moral worth of agents: according to what he calls the "Symmetry Thesis,"

> The reason why a good-willed person does an action, and the reason why
> the action is right, are the same.[14]

Stratton-Lake claims that some thesis like this must be accepted by Kantians.[15] For to reject it, Stratton-Lake maintains, is to abandon "the central [Kantian] view that there is an essential and direct connection between morality and rationality." He continues:

> For if we abandon the symmetry thesis, there need be no connection at all
> between what it is in virtue of which an action is morally good, and the
> normative reasons why it should be done.[16]

Why should this worry the Kantian? I'll set out the argument as it applies to the Coincident Reasons Thesis, although an analogous argument, rephrased to focus on agents, not actions, could be used to support Stratton-Lake's symmetry thesis. Kantians will be likely to accept the following three premises:

(P1)   Rationality just is responsiveness to subjective[17] reasons: an action is rational
        if and only if it is one the agent has conclusive subjective reason to perform and
        the agent performs it in response to those reasons.

(P2)   Strong moral rationalism: we always have conclusive subjective reason to do as we morally ought. An action ought to be performed by an agent if and only if it is one the agent has conclusive subjective reason to perform (and those reasons are the reasons why it ought to be performed).

(P3)   A good will is= a rational will; an action is morally worthy if and only if it is rational.

It follows from (P1) and (P2) that an action is rational if and only if it ought to be performed by an agent, and the agent performs it for the reasons in virtue of which it ought to be performed. It follows from (P3) that the same holds for a morally worthy action. This gives us:

(C)   An action is morally worthy if and only if it ought to be performed by an agent and the agent performs it for the reasons why it ought to be performed.

(C) is equivalent to the Coincident Reasons Thesis.

As Stratton-Lake notes, it is possible to accept the symmetry thesis (or my Coincident Reasons Thesis) without rejecting the Motive of Duty Thesis. Korsgaard, for one, accepts both. She writes that "the reason why a good-willed person does an action, and the reason why the action is right, are the same," and adds that "[t]he good-willed person does the right thing because it is right."[18] But accepting both the Coincident Reasons Thesis and the motive of duty thesis has the strange entailment that the fact that an action ought to be performed is itself a normative, or justifying, reason why it ought to be performed. And this, as Stratton-Lake points out, is implausible.[19]

One thought at work here is that normative reasons do explanatory work. Justification is a kind of explanation. But, the thought is, facts cannot *explain* themselves. The fact that some action ought to be performed doesn't *explain why* it ought to be performed, so it can't be *a reason why* it ought to be performed. Plausibly, the statement "A ought to ϕ" simply reports the fact that A has (other) overriding reasons

to φ. If we were to take the fact that A ought to φ as an additional reason for A to φ, we would be guilty of double-counting the reasons A has to φ. We don't have reason to save the drowning boy because it is the right thing to do, *and* because he might otherwise have died and his life is of value. It is the right thing to do *because* his life is of value.[20,21]

4   Does the thesis provide *sufficient* conditions for moral worth?

I have argued that it is not necessary that we act on the motive "that it is right" in order for our act to have moral worth. Indeed, as the familiar case of Twain's Huckleberry Finn shows, an act can have moral worth even if it is performed in the belief that it is *wrong*. Huck, finding himself unable to pray on account of a guilty conscience, decides to write a letter to Miss Watson to tell her where to find her runaway slave Jim, Huck's travel companion for some time. After writing the letter, he falls into thought:

> I felt good and all washed clean of sin for the first time I had ever felt so in my life, and I knowed I could pray now. But I didn't do it straight off, but laid the paper down and set there thinking—thinking how good it was all this happened so, and how near I come to being lost and going to hell. And went on thinking … and I see Jim before me all the time: …we a-floating along, talking and singing and laughing. But somehow I couldn't seem to strike no places to harden me against him, but only the other kind. […] and then I happened to look around and see that paper.
>   It was a close place. I took it up, and held it in my hand. I was a-trembling, because I'd got to decide, forever, betwixt two things, and I knowed it. I studied a minute, sort of holding my breath, and then says to myself: "All right, then, I'll go to hell"—and tore it up.[22]

When Huck wrestles with his conscience about whether to turn in or protect Jim, and decides to protect him, despite believing this act to be terribly wrong, he is motivated at least in part by his recognition of Jim's value as a fellow human being—that is, by facts which justify his choice. The Coincident Reasons Thesis rightly lauds Huck's act. Examples like that of Huck make very plausible the claim that the thesis identifies a

sufficient condition for the moral worth of an action, and that a Kantian need not accept the Motive of Duty Thesis.

Stratton-Lake disagrees. He follows Kant in thinking that a motive capable of endowing the action performed from it with moral worth is one which is non-accidentally related to the rightness of the act done from it. He then offers an example intended to show that the motive I have been defending may be only accidentally related to the rightness of the acts done from it. He writes:

> It would be merely accidental if I were such that I was only motivated to do the right thing by the normative reasons why it is right *on the condition that* it was in my interest to do so. As soon as I judge that it is *not* in my interest to do the right thing, the normative reasons which would otherwise motivate me to act, lose their grip on me. …
>   If this sort of case is psychologically coherent, which I think it is, then we have a case in which I may be motivated to do the right thing, by the normative reasons why it is right, but in which the relation between my motives and the rightness of my action is purely accidental.[23]

Stratton-Lake thinks the possibility of this kind of counterexample to the symmetry thesis and my Coincident Reasons Thesis shows that these theses do not on their own provide sufficient conditions for the moral worthiness of agents or actions.[24] A further condition is needed, one he argues is supplied by the addition of the motive of duty as a "secondary motive."[25]

According to Stratton-Lake, then, morally worthy actions are those that are motivated at the primary level by the reasons that make them right, *and* are performed from a secondary motive of duty. What is it to act from duty as a secondary motive? He tells us:

> to act from duty as a secondary motive is to regard myself as having sufficient reason to do some act (at the primary level) solely insofar as I judge that it is morally required.[26]

As it stands, Stratton-Lake's proposed additional condition on the moral worth of actions seems a little strange. For *I* regard myself as having sufficient reason to perform some act whenever I believe it to be morally required—this just follows from my moral rationalism. But nothing seems to be implied by this about my motivations. My belief in moral rationalism is not a motivational state at all—it is simply a belief, and a metaethical belief at that: it's a belief about what moral-ought-statements mean. It seems strange to say that my metaethical beliefs could have any impact on the moral worth of my actions.

I may indeed be such that once I judge that I have sufficient reason to act in a certain way, or that I am morally required to do so, I become motivated to do so as a result. But this will be a further fact about me—one that is not entailed by my moral rationalism alone. Perhaps having duty as a secondary motive amounts to being just this sort of person—someone who finds herself motivated to do whatever she judges she's morally required to do. But if having this sort of a motivational disposition is what it is to act from duty as a secondary motive, I take the case of Huck Finn to show that such a disposition is *not* required for an action, or an agent, to have moral worth.

Let's look again at Stratton-Lake's example, which introduces the gap that the appeal to secondary motives is intended to fill. It is not so easy to see what might be going on in the psychology of the agent he asks us to imagine. How are we to determine, in this man's case, whether prudential considerations are playing a motivating or a "purely regulative" role? What does this difference come to? If prudential concerns are playing a fundamental motivating role, I am inclined to share Stratton-Lake's intuition that the man's actions have no moral worth. But in this case, as Stratton-Lake would allow, the conditions for moral worth established by the symmetry thesis (and the

Coincident Reasons Thesis) would not be met. We can imagine, for example, a self-interested believer in an afterlife in which God rewards good-doers by inviting them into heaven. If our believer is a reliable judge of right and wrong, his actions will reliably coincide with what he ought to do, but this reliability will not give his actions moral worth: his fundamental motivating reason for performing them is not also a justifying reason.

But presumably, we are to imagine a case in which our agent has no such further fundamental primary motive. Perhaps we can imagine self-interest playing instead the role of a *second-order* motive. Perhaps our agent is motivated on self-interested grounds to be motivated by the right-making reasons. So let's now consider a second self-interested believer. Our second believer believes that God will let her into heaven only if she does the right things *for the reasons why they are right*—that is, only if her actions are not only right but *morally worthy*. Her end is still to get into heaven, but she cares (instrumentally) not only about doing the right thing, but also about doing it for the right reasons. Does this sort of case provide a counterexample to the Coincident Reasons Thesis?

I'm still inclined to think that it does not. For the only way our second believer could comply with what she thinks God wants of her is for her to actually start doing the right thing for the reasons that make it right—that is, to start being motivated (fundamentally) not by self-interest but by concern for others. She might have been motivated by self-interest to become someone who is motivated by concern for others, but that just means that self-interest drove her to become a better person. We can think of other similar real-life cases of self-interested self-improvement. We can imagine, for

example, a violent heavy drinker, whose wife threatens to leave him unless he gives up alcohol. It seems plausible to say, of such a person, that he is a better person after he stops drinking than he was before, despite the self-interested motive driving his reformation. One conclusion we should draw from such examples is that second-order motives are discharged as soon as the primary motives they target are formed—they don't linger on to *co*-motivate, so-to-speak, the actions motivated by those primary motives.[27]

We might understand secondary motives not as beliefs agents have about the conditions under which they have sufficient reason to act, or even as dispositions they have to be motivated by certain considerations if they believe them to be of a particular class (e.g., the class of right-making reasons), but rather as dispositions to act on certain classes of reasons, even in the absence of any beliefs about those reasons. Because this understanding doesn't require attributing to the morally worthy agent any true beliefs about what morality requires, it has the advantage that it would make Stratton-Lake's argument less vulnerable to the kind of counterexample posed by cases like Huck Finn's. Understood in this way, secondary motives still play the broadly regulative role that Stratton-Lake, following Barbara Herman and Marcia Baron, identified: they still establish the conditions under which agents act on particular primary motives.

Nomy Arpaly, in her discussion of the conditions under which actions have moral worth, suggests a case in which just this kind of *dispositional* secondary motive seems to be lacking. Arpaly, like Stratton-Lake, thinks a thesis like the Coincident Reasons Thesis provides only a necessary condition for the moral worth of actions. She argues that another factor, which she identifies as an agent's "degree of moral concern," weighs into

the assessment of moral worth. To bring this out, she asks us to compare two characters—the "foul-weather," or "die-hard philanthropist," who "would act benevolently even if severe depression came upon her and made it hard for her to pay attention to others," and her "fair-weather friend, who acts benevolently as long as no serious problems cloud her mind, but whose benevolent deeds would cease, the way some people drop their exercise programs, if there were a serious crisis in her marriage or job."[28]

Arpaly's example looks much like Stratton-Lake's: again we have a case where an agent is motivated to act by the right-making reasons, but only conditionally so. Her motivation could quite easily have been undermined, had her circumstances been different. And Arpaly's case is certainly psychologically coherent—even familiar. But I'm inclined to think that while the relative fragility of the fair-weather philanthropist's action may reflect badly on *her, as a moral agent*, it doesn't reflect badly on her *action*. So while it may undermine the symmetry thesis, which states conditions for the moral worth of agents, it doesn't undermine the Coincident Reasons Thesis. Intuitions that the die-hard philanthropist's actions are particularly morally worthy may be colored somewhat by a complicating factor: for some people might think that the difficulty of an agent's circumstances affects her moral obligations, so that someone whose own life is very difficult is under less of an obligation to help others. If this is so, then their admiration for the die-hard philanthropist may reflect a belief that her actions are supererogatory, rather than required. In that case, the intuition that her actions are worthier than her fair-weather friend's won't tell against the Coincident Reasons Thesis.

In any case, the kind of appeal to counterfactuals on which Arpaly's assessment of this example depends can lead us astray. It is often a mistake to ask, when assessing the moral worth of some action, "would she have still done that if…?". If a fanatical dog-lover performs a dangerous rescue operation to save a group of strangers, at great personal risk, we shouldn't discount the worthiness of his actions because, had his dog required his heroics at the same time, he would have abandoned the strangers. That he would have done so may be a sign of his excessive concern for the dog, rather than of too little concern for the strangers—after all, he was willing to risk his own life to save theirs. And given that the dog was not present to deflect our hero's attention from the reasons he had to perform the rescue, it seems ungenerous to withhold praise for so admirable an act simply because the dog might have been present.

5   Does the thesis provide *necessary* conditions for moral worth?

I have been defending the Coincident Reasons Thesis as a statement of sufficient conditions for the moral worth of actions. But I introduced the thesis as a statement of the *necessary and* sufficient conditions for an action's having moral worth. This is a more contentious claim, which raises complex issues. Is performing an act for the duty-based motivating reason "that the act is right" sufficient for making a right act morally worthy, even if it is not necessary? And is it possible to act for this reason without acting for the right-*making* reasons that justify an action? If the answer to these questions is yes, then the Coincident Reasons Thesis fails to identify *necessary* conditions for the moral worth of actions; it will be possible for an act to have moral worth without its being motivated by any of the reasons that in fact justify the act.

Answering these questions is not so easy. I will only raise a few initial thoughts.
I think that right actions that are motivated (merely) by an agent's true *belief* that they are
right may not have moral worth. Such a belief might be formed through reliance on a
very bad 'authority' on right action—an authority that is usually wrong and will lead our
agent astray more often than not. Imagine that Huck does turn in Jim, because the moral
authorities he accepts as reliable impress upon him the (supposed) rightness of this act. If
he gives alms to the poor because (and *only* because) the *same* authorities tell him it is
right to do so, and he believes them, his act of charity, though clearly right, does not seem
to me to have moral worth. The mechanism that triggers Huck's motivation in this
imagined case seems too distantly and unreliably related to the true justifications for such
charity. Merely having the intention to do *the right thing*, read *de dicto*, is not enough to
make one's right acts morally worthy.

Should we think Huck Finn less morally worthy because he acts as he believes he
ought not? I've argued that the fact that one ought to do something is not itself a
normative reason to do it. But it remains possible that the *belief* that one ought to do
something could provide one with a normative reason to do it.[29] Indeed, I think that our
normative beliefs do sometimes provide us with normative reasons of a particular kind:
like rules-of-thumb, they are reasons to refrain from acting directly on some class of first-
order reasons—in this case, the reasons my beliefs are beliefs *about*—which they are
intended to reflect. However, a belief about what ought to be done can provide an agent
with a reason to do it only if his ought-beliefs generally reliably reflect the underlying
first-order normative reasons that determine what it's best that he do. And this, surely,
will not apply to Huck, whose normative beliefs are badly skewed by the racist opinions

of the society in which he lives.[30] So we shouldn't blame Huck for acting as he believes

he ought not, though he may or may not be blameworthy for having false normative

beliefs.

It is more plausible that an act that is motivated by the *justified* true belief that it is

right has moral worth. We can imagine, for example, an agent who relies on a *good*

authority in believing an act to be right, and is motivated to act by this belief, without

having any knowledge of the underlying first-order reasons on which the authority bases

his judgment. But this kind of case does not provide a counterexample to the Coincident

Reasons Thesis, taken as a necessary condition on moral worth: it is not a case in which

a morally-worth action is not motivated by the reasons that make the act right. For in

defense of the thesis I might now reply that the advice offered by the good authority is

one of the reasons that *justifies* the agent's act—that makes it right. That is, it is one of

the reasons that makes the act right. The moral reasons for us to perform some action are

subjective—we are morally required to do only what we have sufficient epistemic reason

to believe it would be best to do, not what it would (in fact) be best to do. If the advice of

the good authority provides our agent with sufficient evidence for the belief that a

particular act would be best, his acting so may be *made* right by the fact that the good

authority advises him to do it (indeed, this may be true even if the authority advises him

wrongly). His belief that the act is right, which motivates him to act, is based on the

good authority's advice. And his act is right *because* it follows the authority's advice.

So in this case the reasons motivating and justifying our agent's act coincide after all.

6   Degrees of moral worth

I have argued that the Coincident Reasons Thesis provides plausible necessary and sufficient conditions for the moral worth of actions. But I have proceeded mainly by trying to defuse possible counterexamples, so I don't take myself to have established, beyond doubt, that no more damaging counterexample to the thesis can be found. My discussion also oversimplifies the nature of moral worth, as well as the structure of our reasons (both motivating and justifying) at a critical place: I've been talking as if moral worth and the coincidence between motivating and justifying reasons were all-or-nothing phenomena; but both the moral worth of actions, and the overlap between the reasons motivating and justifying them, is often, perhaps usually, a matter of degree. Are actions morally worthy to the degree that the reasons motivating them coincide with the reasons justifying them? Unfortunately, this question has no straightforward answer. If we can identify different *non-instrumental* motives, each of which contributes *independently* to an agent's performance of a right action, then the action's moral worth does seem to me to vary with the degree to which those motivating reasons coincide with reasons justifying the action. But as anyone who has had to answer a persistent four-year-old's string of "why?"s can attest, the reasons for which we act, as well as the normative reasons justifying our actions are often interrelated in complex ways. Both the self-interested reward-seeker and the altruist jump into the water to save the drowning boy. And that jumping in will allow them to rescue him is in both cases a normative reason to do so.

For the purposes of this conference paper (already much too long), I will put forward without defense a working hypothesis about degrees of moral worth: right

actions, I propose, have moral worth to the degree that the non-instrumental motivations for their performance coincide with non-instrumental justifications for their performance. An agent can go wrong in acting rightly by failing to be motivated by some consideration that in fact justifies her action, and she can go wrong by being motivated by considerations that don't in fact justify her action. But there are at least two additional ways in which she can go wrong as well: she can treat as fundamental a reason for acting which is in fact merely instrumental (we do this when, like misers, we treat the fact that some action will help us acquire wealth as a non-instrumental reason to perform it); or, (like the first self interested believer after a heavenly reward, whom I discussed in response to Stratton-Lake[31]) she can treat as a merely instrumental consideration a reason that in fact plays a fundamental justificatory role.

The Coincident Reasons Thesis can, I think, also help explain when *wrong* actions have (at least some) moral worth. We are all familiar with the phrase, "he *meant* well." It picks out, I think, a different kind of overlap between justifying and motivating reasons: an overlap in the *more* fundamental links of the reasons chain, and a divergence at the less fundamental links. This occurs when someone is right about what, fundamentally, makes actions right—he recognizes, let's say, the importance of treating others as ends and not just means—but he makes (unwarranted) errors in, for example, his *instrumental* reasoning; he is wrong about what these justifying reasons require him to do. Such a person might be said to perform the wrong action, but for the right reasons (or rather, *some* of the right reasons). And I'm inclined to say such an action has some moral worth as well.

All this intended only as an initial sketch of what morally worthy actions look like. But even an initial sketch will, I hope, have sufficed to show that the Kantian needn't—in fact, shouldn't accept one of the less attractive views often associated with Kantian ethics: that morally worthy agents must act from the motive of duty. At the very least, the Coincident Reasons Thesis identifies other possible motives on which the morally worthy agent can act.

Works Cited

Arpaly, Nomy. *Unprincipled Virtue: An Inquiry into Moral Agency*. New York: Oxford University Press, 2003.

Foot, Philippa. "Morality as a System of Hypothetical Imperatives," in *Virtues and Vices*. Berkeley, Ca: University of California Press, 1978, pp. 157-173.

Herman, Barbara. "On the Value of Acting from the Motive of Duty," in *The Practice of Moral Judgment*. Cambridge, MA: Harvard University Press, 1993, pp. 1-22.

Kant, Immanuel. *Groundwork of the Metaphysics of Morals* (1785), translated by Mary Gregor. Cambridge: Cambridge University Press, 1997

Parfit, Derek. *Climbing the Mountain* (unpublished manuscript), 2007.

Raz, Joseph. *Practical Reason and Norms*. Oxford: Oxford University Press, 1990.

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

Stratton-Lake, Philip. *Kant, Duty, and Moral Worth.* London: Routledge, 2000.

Twain, Mark. *The Adventures of Huckleberry Finn* (1885). New York: Oxford University Press, 1996.

[1] Kant, *Groundwork of the Metaphysics of Morals* (hereafter, *Groundwork*), p.3 (4:390).
[2] *Ibid.*, p. 11 (4:398).
[3] *Ibid.*, p. 12 (4:398).
[4] Thus Barbara Herman, for example, writes in "On the Value of Acting from the Motive of Duty,"

> an action has moral worth if it is required by duty and has as its primary motive the motive of duty. The motive of duty need not reflect the only interest the agent has in the action (or it's effect); it must, however, be the interest that determines the agent's acting as he did. (p. 16)

She adds, a little later, "[o]verdetermined actions *can* have moral worth so long as the moral motive is the determining ground of the action—the motive on which the agent acts." (p. 21) See also Marcia Baron, *Kantian Ethics Almost without Apology*.

⁵ We might think an action performed "out of respect for the moral law" is one performed because it is *believed* to be right. I set this thought aside here, because it seems to me less promising that the version of the Motive of Duty Thesis I focus on above. I come back to this alternative version, and the reasons why it seems to me to be so unpromising in §5 below.

⁶ For an argument for the claim that action performed from the motive of duty involves a kind of "moral fetishism", see Smith's *The Moral Problem*, pp. 71-76.

⁷ Foot, p. 165.

⁸ Philippa Foot links the acceptance of the Motive of Duty Thesis with the acceptance of the central Kantian doctrine of the categorical imperative. She writes, "the doctrine of the categorical imperative has owed much [to the persuasion of that thesis]." (Foot, p. 164) It is not at all clear to me why Foot thinks the doctrine of the categorical imperative relies on the Motive of Duty thesis, and Foot does little to elucidate her claim. But the Motive of Duty thesis should, I think, seem plausible to a Kantian only if the fact that an action is right is *itself* an overriding reason to perform it. Otherwise, the agents the thesis identifies as morally worthy – those motivated by the rightness of their actions – would be acting irrationally, or at least *a*rationally – not in response to the balance of reasons – and a Kantian, who identifies a good will with a rational will, should not want this. We might think (as Foot seems to) that a commitment to the doctrine of the categorical imperative, according to which rational beings always have overriding reasons to perform right actions, must also identify the rightness of an act as an overriding reason, and this *may* be why Foot links the two theses together. (Foot identifies denial of this proposition as the chief think distinguishing the proponent of the view that morality is a system of *hypothetical* imperatives from the Kantian. See pp. 166-7.) But this would, I think, be a mistake. I argued in my thesis that the reason every rational being has to do as the moral law requires is not simply the *de dicto* reason "that the moral law requires it." Rather, our reasons to be moral are generated by the operation of the standards of procedural rationality on our actual ends and commitments. If this is right, then we will always have conclusive reason to be moral because being moral will simply be a matter of doing whatever it is we (anyway) have conclusive reason to do. This is, I think, Kant's own view, at least in portions of the argument of the *Groundwork*. And there are, of course, other accounts of the categorical nature of moral reasons. We hope to give the moral skeptic, who asks, "what reason have I to do as morality requires?", a more satisfying answer than a brute insistence "that morality requires it." The goal is to persuade him not through the sheer force of repetition, but by reasoned argument.

⁹ In my discussion of other versions of the Right Reasons Thesis, I have taken the thesis to be about *necessary* conditions for the moral worth of actions. Here I expand the thesis to state *necessary and sufficient* conditions for the moral worth of actions. I think it very plausible that the Coincident Reasons Thesis provides at least sufficient conditions for an action's having moral worth, though I will raise and respond to some doubts about this claim. I return to the question of necessary conditions below.

There is, of course, considerable debate about what it takes for agents to be *morally responsible* for their actions, and it is very plausible that an action can have moral worth only if it is one for which the agent is morally responsible. So the Coincident Reasons Thesis provides sufficient conditions for the moral worth of actions

only given that the conditions for moral responsibility have already been met. I don't want to take any stand on what these are. On some views on moral responsibility (such as Susan Wolf's, perhaps), according to which agents can be morally responsible for their actions only if they are appropriately responsive to reasons, the conditions for moral responsibility may automatically be met whenever the conditions established by the Coincident Reasons Thesis are met. Such views are, of course, controversial. We should understand the Coincident Reasons thesis as describing only *free* actions: according to the thesis, a *free* action can have moral worth if and only if the reasons motivating it and the reasons justifying it coincide.

[10] Kant, *Groundwork*, pp. 3-4 (4:390).

[11] See Kant, 4:390 (in the "Preface") and Kant's discussion of acting from duty at 4:397-398 of Section II (all quoted above).

[12] Kant, p. 38 (4:429).

[13] See Parfit, *Climbing the Mountain* (unpublished manuscript), Chapter 8, §25 (p. 173).

[14] Stratton-Lake, *Kant, Duty, and Moral Worth*, p. 16. Stratton-Lake extracts this statement of the thesis from Korsgaard, *Creating the Kingdom of Ends*, p. 60

[15] Stratton-Lake actually thinks the thesis is too strong, as it stands, to provide necessary conditions for the moral worth of actions. At least one of his worries is relevant to the Coincident Reasons Thesis as well, but I'll omit discussion of it in this version of this paper, in an effort to keep the paper shorter. I don't think Stratton-Lake's worry withstands closer scrutiny. (See Stratton-Lake, pp. 16-17.)

[16] Stratton-Lake, p. 60.

[17] We have subjective reason to do what we ought to believe we have objective reason to do—subjective reasons are relative to our current epistemic positions, objective reasons are not. Parfit, for example, calls subjective reasons "apparent reasons."

[18] Korsgaard, p. 60.

[19] Stratton-Lake, p. 20.

[20] At most, the fact that I ought to ɸ provides me with a *derivative* reason to ɸ: a reason that adds no normative weight to me primary reasons for ɸing and does no work in explaining why ɸing is right.

[21] Could the fact that I ought to ɸ be a *second-order, dependent* reason for me to ɸ, of the sort described by Joseph Raz? My *belief* that I ought to ɸ might sometimes provide me with a reason to do so—a reason whose weight is not to be added into the balance of first order reasons, but is rather to *replace* those reasons in our deliberation. I return to this issue in section 5. The fact that I ought to ɸ might also provide me with such a second-order, dependent reason if I'm in circumstances in which I have epistemic access to that fact without having any access to the underlying first-order reasons that explain it. I'm not sure when (or even whether) such circumstances could obtain. (See Raz, *Practical Reason and Norms*, pg. 35.)

[22] Twain, Mark, *The Adventures of Huckleberry Finn*, pp. 271-272. Nomy Arpaly, in her book *Unprincipled Virtue: An Inquiry into Moral Agency*, also appeals to the Huck Finn example in her defense of a similar condition on morally worthy action to the one I propose above, in the form of the Coincident Reasons Thesis. Arpaly thinks her similar condition, which she calls "Praiseworthiness as Responsiveness to Moral Reasons" (p. 72), tells only part of the story about when actions have moral worth. I will discuss her

proposed further condition, and the examples she thinks show it to be necessary, below. See Ch. 3 of *Unprincipled Virtue* for a nuanced discussion of the problem of moral worth.

[23] Stratton-Lake, p. 65.

[24] Although Stratton-Lake's argument mainly concerns the symmetry thesis, and the moral goodness of *agents*, he makes it clear that he also sees the type of counterexample described as having implications for the moral worth of *actions*. On the page from which the example is taken, he presents himself as answering the question "why think that duty must function as a secondary motive if someone, *or some act* is to be morally good?" (p. 65, italics added)

[25] Stratton-Lake takes the phrase from Barbara Herman, *op. cit*.

[26] Stratton-Lake, p. 62.

[27] In other words, motives aren't transitive: If M (say, a self-interested motive) motivates the formation of P (say, a worthy primary motive), and P motivates some action ϕ, it doesn't follow that M motivates ϕ. (Thanks to Stephen Kearns for pointing out this way of putting the point.)

[28] Arpaly, p. 87.

[29] Stratton-Lake denies that "verdictive" normative beliefs can give us reasons in this way (see, e.g., p.12, p.20). I'm not sure from what argument he takes this conclusion to follow. It is not, it seems to me, entailed by the conclusion that the *fact* that I ought to ϕ is not a normative reason to ϕ. Nor does it seem to me to follow from Stratton-Lake's other claims that "a verdictive moral consideration cannot be cited in support of itself" (p. 20) because "no verdict constitutes evidence for itself" (p. 19). I have my doubts about these claims: if my beliefs are generally reliable, the fact that I believe something may well be evidence of its truth. In any case, even if we accept that my belief that I ought to ϕ is never an epistemic reason to believe I ought to ϕ, it may nonetheless be a normative reason for me to ϕ.

[30] Krister Bykvist has pointed out that my defense of the Coincident Reasons Thesis as a sufficient condition for the moral worth of actions has the perhaps surprising implication that someone's actions could have moral worth even if she has no normative beliefs at all. Is this grounds for objecting to the thesis? I can see why we might take it to be worrying. The law, for example, considers an agent legally accountable for her actions only if she is judged to have the capacity to distinguish right from wrong, and we might think that this capacity requires having normative beliefs. But I think we should resist this intuition. If we are inclined to think Huck can be morally worthy despite having consistently *false* normative beliefs (and so poor judgment about what is right and what is wrong), I don't think that the *absence* of normative beliefs could rule out an agent's moral worth.

It's interesting to note how very *un-Kantian* this worry is: the ideal Kantian agent is, after all, usually accused of *over-rationalizing*—of having, to quote Bernard Williams, "one thought too many."

[31] See p. 11 (?) above.