MIT Open Access Articles

Acting for the Right Reasons

Acting for the Right Reasons[*]

Julia Markovits
Massachusetts Institute of Technology

Introduction

This essay examines the thought that our right actions have moral worth only if we
perform them for the right reasons. On the face of it, views about the conditions of moral worth
seem independent of what first-order moral views we hold. That is, we can debate what *else*
must be true of right actions for them to count as morally worthy without first settling the
question of what it takes for them to be right. My initial aim will be to identify the conditions
under which right actions have moral worth, and I believe the intuitive appeal of my account of
moral worth and the force of most of the arguments I marshal in its support are independent of
our adopting any particular first-order ethical standpoint. Nonetheless, the view of moral worth I
defend turns out to have implausible implications when held in conjunction with any of a class of
first-order ethical views that includes utilitarianism. Because utilitarians would, I think, be hard-
pressed to come up with an account of moral worth as independently plausible as the one I
defend, my argument for this account turns out to provide an objection to utilitarianism.
Thinking about moral worth may tell us something about which actions are right after all.

In *section 1*, I introduce and begin to argue against the traditional Kantian account of moral worth, according to which morally worthy actions must be performed from the "motive of duty," or because they are right. I suggest an alternative formulation of the thought that morally worthy actions must be performed for the right reasons, according to which morally worthy actions are those performed for the reasons *why* they are right. I argue that this alternative account should in fact be accepted by Kantians because it is entailed by some central tenets of Kantian ethics. In *section 2*, I argue that my account provides plausible *sufficient* conditions for an action's having moral worth; that it can explain the moral worth of some actions whose worth the motive of duty thesis excludes; and that it provides a good account of the idea that in the case of morally worthy actions, it is no accident that the agent acts rightly. In *section 3*, I argue that my account also provides plausible *necessary* conditions for the moral worth of actions, defending that claim against proposed counterexamples. In *section 4*, I argue that the plausibility of my account of moral worth, which is largely independent of any particular ethical standpoint, gives us some reason to doubt a class of ethical theories that includes utilitarianism. *Section 5*, the final section before the *conclusion*, considers some issues concerning partially worthy actions, including the case of wrong actions that seem nonetheless partially worthy.

1   The Motive of Duty, the Coincident Reasons Thesis, and Kantian ethics

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

---

[1] Immanuel Kant, *Groundwork of the Metaphysics of Morals*, trans. Mary Gregor (Cambridge: Cambridge University Press, 1997), 3 (4:390), emphasis in original. Parenthetical citations from Kant's work refer to the Royal Prussian Academy of Sciences edition of Kant's Gesammelte

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.

I will argue that the Motive of Duty Thesis should be rejected. I will defend a different way of spelling out the more general thought that right actions are morally worthy only if they are performed for the right reasons – one which should appeal to Kantians and non-Kantians alike.

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*. But not *all* praiseworthy actions have moral worth. We praise many actions for valuable or admirable qualities they have that are not moral – skillful actions, for example, are also praiseworthy. Morally worthy actions are ones that reflect well on the moral character of the person who performs them. This is not to say that only virtuous people can perform worthy actions – it is possible to act, in this sense, out of character.

Schriften (Berlin: George Reimer, later Walter de Gruyler, 1900–), with volume and page numbers separated by a colon.

[2] Ibid., 11 (4:398).

[3] Ibid., 12 (4:398).

But morally worthy actions are the building blocks of virtue – a pattern of performing them makes up the life of a good person.[4]

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. 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 hope to show that the Motive of Duty Thesis runs against the grain of some central and attractive elements of the Kantian approach to ethics and wrongly 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

---

[4] It is important to distinguish, in this context, between actions we have instrumental reasons to praise, and actions that merit praise in their own right. Kant himself makes clear that he believes that an action performed from beneficent motives, such as sympathetic concern for others or the inclination towards honor to "deserves praise and encouragement" because it "fortunately lights on what is in fact in the common interest and in conformity with duty." Such an action is to be praised if praising it makes people more likely to perform similar actions because it is *good* that people perform them. But it is not to be "esteemed" as morally worthy because it may not be good *in* a person that she performs it. According to Kant, it may be merely accidental – "fortunate" – that a person acting on such motives does the right thing; the motive need not reflect a *good will*. Whether Kant is right to characterize sympathetic actions in this way will be discussed further below (see sec. 3). Thanks to an anonymous referee for the *Philosophical Review* for pressing me to be more precise on this point.

I will also return, in discussing the problem my account of moral worth raises for utilitarianism, to the question of whether a distinction between actions we have instrumental reasons to praise and actions that merit praise in their own right must be acknowledged (see sec. 4, n. 65).

[5] We might think an action performed "out of respect for the moral law" is one performed whenever it is *believed* to be right, regardless of whether it actually *is* 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 in sec. 3.

Duty Thesis. The Kantian "truly moral man" seems guilty of a kind of moral fetishism (to

borrow a phrase from Michael Smith),[6] or at best, of having "one thought too many" (to borrow

one from Bernard Williams),[7] 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.[8]

The Motive of Duty Thesis seems designed to exclude from moral worth, quite correctly,

those right actions that were performed from what might be called ulterior motives. But, as I

will argue, the thesis excludes more than it needs to, and another way of spelling out the Right

Reasons thesis does more justice to our intuitions. Ulterior motives are, presumably, those

generated by facts that are not morally relevant features of the situation in which we act. It's

---

[6] For an argument for the claim that action performed from the motive of duty involves a kind of "moral fetishism," see Michael Smith, *The Moral Problem* (Oxford: Blackwell1994), 71-76.

[7] See Bernard Williams, "Persons, Character, and Morality," in *Moral Luck* (Cambridge: Cambridge University Press, 1981), 1-19, at 18.

[8] Christine Korsgaard argues that this worry about Kant's account of moral worth stems from a misreading of it. She maintains that the appropriate bearers of moral value are *actions*, which, she says, include both the act performed and the end to which it was performed. She takes this account of actions to make the Motive of Duty thesis less unpalatable. She writes,

> The idea that acting from duty is something cold, impersonal, or even egoistic is based on the thought that the agent's *purpose* or *aim* is 'in order to do my duty' *rather than* 'in order to help my friend' or 'in order to save my country,' or whatever it might be. But that is just wrong. Sacrificing your life in order to save your country might be your duty in a certain case, but the duty will be to do that act *for that purpose*, and the whole action, both act and purpose, will be chosen as one's duty. (Korsgaard, "Acting for a Reason," in *The Constitution of Agency* (Oxford: Oxford University Press, 2008), 207-229, at 218-219.)

This response seems to me unsatisfactory. As I hope will become clear, even in Korsgaard's version, the thesis retains the tension with other aspects of a Kantian approach to ethics that I'll describe in a moment, and remains vulnerable to other objections and counterexamples – particularly those provided by apparently morally worthy agents who act in ignorance of the moral status of their actions – that I will discuss in sec. 2. But more relevantly here, it does not, in any case, entirely avoid the kind of "one thought too many" worry that the Motive of Duty Thesis inspired. That's because it's not clear how the motive of duty could, on Korsgaard's view, relate to the action of, say, doing something in order to help one's friend, except by providing a more fundamental motive: I did something to help my friend in order to do my duty. The duty, and not the friend, remains the primary target of one's attention. (I'll have more to say about such chains of motivating reasons later on.)

very plausible 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 morally justify a conclusion about what should be done – that provide morally *justifying* reasons for action.

When I am faced with a practical decision – for example, when I must decide whether to rush into a burning house to save a trapped child – there are many features of the situation that may be morally relevant. The endangered well-being of the child 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 child 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 morally justifying reason: it was not the prospect of reward that made saving the child the right thing to do.

According to what I will call the **Coincident Reasons Thesis**, *my action is morally worthy if and only if my motivating reasons for acting coincide with the reasons morally justifying the action* – that is, if and only if I perform the action I morally ought to perform, for the (normative) reasons why it morally ought to be performed.[9] My motivating reason for

---

[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. 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 if meeting the conditions for moral worth established by the thesis also entails meeting the conditions for moral responsibility. It is my view that only agents who can be held morally responsible for their actions have moral reasons – that is,

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, since a ground that is not
> moral will indeed now and then produce actions in conformity with the law, but it
> will also often produce actions contrary to law.[10]

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

---

moral reasons apply to agents only if they've met the conditions for moral responsibility. Since the Coincident Reasons Thesis attributes worth to an agent's actions only when the agent is motivated to act by the moral reasons that apply to her, conditions for moral responsibility will be have been met whenever the conditions for moral worth specified by the thesis have been met.

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

In the passages from the *Groundwork* that I quoted earlier, Kant seems to endorse the Motive of Duty Thesis.[11] But the account of moral worth given by the Coincident Reasons thesis sits much more comfortably with some central tenets of the Kantian approach to ethics. To reject it is to abandon what Philip Stratton-Lake has called "the central [Kantian] view that there is an essential and direct connection between morality and rationality."[12] It is a familiar Kantian thought that moral goodness and practical rationality go hand in hand. Being good, on the Kantian view, is a matter of being practically rational. And being practically rational is a matter of responding in one's actions to the reasons one has for acting. So a Kantian should be sympathetic to the thought, expressed by the Coincident Reasons Thesis, that morally worthy actions are those an agent performs in response to the reasons she has to perform them. It is entailed by Kantian *moral rationalism* – the view that we always have conclusive reason to do as morality requires – in conjunction with the Kantian view that a *good, or morally worthy, will is a rational will* – a will that's responsive to reasons.[13]

---

[11] See Kant, 4:390 (in the "Preface") and Kant's discussion of acting from duty at 4:397-398 of sec. 2 (all quoted above).

[12] Philip Stratton-Lake, *Kant, Duty, and Moral Worth* (London: Routledge, 2000), 60. Stratton-Lake considers a thesis that is similar to the Coincident Reasons Thesis, 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." Ibid., 16. He extracts this statement of the thesis from Christine Korsgaard, *Creating the Kingdom of Ends* (Cambridge: Cambridge University Press, 1997), 60. Stratton-Lake goes on to reject this thesis as "too strong" as it stands and in any case accepts a version of the thesis only as a necessary, and not as a sufficient, condition for moral worth, for reasons I will come to later on (in sec. 3). But he maintains that some thesis like it must be accepted by Kantians.

[13] Although all morally worthy (good-willed) actions will on this view be rational – that is, actions the agent has conclusive reasons to perform, and which the agent performs for those reasons – not all rational actions will count as morally worthy. Some rational actions – actions that lack moral significance (my choice of the most appropriate spoon with which to eat my breakfast cereal, for example) are simply morally neutral. The Coincident Reasons Thesis speaks of "*morally* justifying reasons" and "actions I *morally* ought to perform" precisely to exclude morally insignificant actions from consideration – they fall outside of its scope.

It is possible to accept the Coincident Reasons Thesis without rejecting the Motive of Duty Thesis.[14] 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 is implausible.

One thought at work here is that normative reasons do explanatory work. Justification is a kind of explanation. But 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.[15] 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 trapped child 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.

2   Does the Thesis Provide *Sufficient* Conditions for Moral Worth?

I have argued that anyone who is sympathetic to the Kantian tying together of moral goodness and responsiveness to reasons ought to accept the Coincident Reasons Thesis, and reject the Motive of Duty Thesis. In the next two sections, I will argue that the Coincident Reasons Thesis provides an account of the moral worth of actions that should appeal to Kantians

---

[14] As Stratton-Lake also notes, in relation to his symmetry thesis (see Stratton-Lake, 20, and my n. 12, above). As he points out, Korsgaard seems to accept both. My argument here mirrors his.

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

and non-Kantians alike, because it tracks and explains our best intuitions about which actions are

morally worthy.

I've suggested 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 Mark Twain's Huckleberry Finn

shows, an act can have moral worth even if it is performed in the belief that it is *wrong*.[16] When

Huck wrestles with his conscience about whether to turn in or protect the runaway slave Jim, his

travel companion of some time, and decides to protect him, despite believing this act to be

terribly wrong (he thinks it amounts to stealing from Miss Watson, Jim's 'owner'), he is

motivated at least in part by his recognition of Jim's value as a fellow human being – that is, by

facts which morally 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 we ought not accept the Motive of Duty

Thesis.[17]

---

[16] Mark Twain, *The Adventures of Huckleberry Finn* (New York: Oxford University Press, 1996), 271-272. Jonathan Bennett's discussion of the case of Huck Finn in "The Conscience of Huckleberry Finn," *Philosophy* 187 (1974): 123-134, first focused philosophers' attention on the example. Nomy Arpaly, in her book *Unprincipled Virtue: An Inquiry into Moral Agency* (New York: Oxford University Press, 2003), also appeals to the Huck Finn case, which she discusses at length, in her defense of a similar condition on morally worthy action to the one I propose here, in the form of the Coincident Reasons Thesis. Arpaly thinks her similar condition, which she calls "Praiseworthiness as Responsiveness to Moral Reasons" (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 in a moment. See chap. 3 of *Unprincipled Virtue* for a nuanced discussion of the problem of moral worth.

[17] A defender of the Motive of Duty Thesis might be tempted to respond to Huck's case that he is not *really* doing something he thinks is wrong because he doesn't really understand what "wrong" means. When he describes his action of protecting Jim as "wrong" to himself, he must mean something like "forbidden by society," and we all know that that is not what "wrong" means. (Such an interpretation will be particularly tempting to those philosophers, including, for example, Christine Korsgaard, who are skeptical about the very possibility of clear-eyed akrasia.) But I don't find this description of Huck's circumstances plausible. Huck may indeed think that whenever some act is condemned by society it is wrong, and he may even (though the book surely leaves this underdetermined) believe that such acts are wrong *in virtue of* being condemned by society. But thinking "wrong" *means* "condemned by society" is

Some philosophers, notably ones working from within the Kantian tradition, broadly-

speaking, have considered the possibility that worthy actions are, or at least include, those

motivated by right-making reasons.[18] But they have worried that even when agents are

motivated to do the right thing for the reasons *why* it is right, their actions might *still* be only

accidentally right. Thus Philip Stratton-Lake, for example, has suggested that we can imagine an

agent who is only disposed to be motivated by right-making reasons when doing so is in his

interest. Such an agent's actions, he argues, would lack moral worth, despite being motivated by

right-making reasons, because it would be a matter of mere accident that the conditions under

which he was motivated by such reasons obtained.[19]

Nomy Arpaly considers a similar case in which a second-order, *dispositional* motive to

be motivated by right-making reasons regardless of circumstance seems to be lacking. She asks

---

quite another thing. I think Huck means *wrong* by "wrong" – after all, he responds to his judgment that he is acting wrongly with a terrible attack of guilty conscience – a response to wrong-doing that will be familiar to many of us. Of course, nothing really turns on the correct interpretation of Huck's case – what matters is that a case of the kind I am imagining, in which someone acts rightly while *really* believing he's acting wrongly, is psychologically imaginable. Moreover, the defender of the Motive of Duty thesis must establish more than that Huck doesn't believe himself to be acting wrongly in protecting Jim to credit his action with moral worth – he must show that Huck takes himself to be acting *rightly* (whatever term he would use instead) – to show that he acts from the motive of duty. And this is surely incredible.

[18] As I note above, Philip Stratton-Lake has pointed out the appeal of a view he calls the "symmetry thesis," which he attributes to Korsgaard, and which claims that good-willed agents act for right-making reasons (in *Kant, Duty, and Moral Worth*: see n. 12, above). And as I also note above, Nomy Arpaly has defended a thesis very like the Coincident Reasons Thesis, which she calls "Praiseworthiness as Responsiveness to Moral Reasons" (in *Unprincipled Virtue*: see n. 16, above). Arpaly takes her thesis to provide a necessary, but not sufficient, condition for the moral worth of actions, and Stratton-Lake thinks his similar thesis identifies neither necessary nor sufficient conditions for moral worth, for reasons I will discuss.

[19] According to Stratton-Lake, while morally worthy actions are indeed motivated by right-making reasons, their agents also act from a "secondary" or regulating, dispositional motive of duty: they will be disposed to be motivated by right-making reasons (on the primary level) only if they *judge their actions to be right*. See Stratton-Lake, *Kant, Duty, and Moral Worth*, 60-67. This shares the significant disadvantage of the more traditional interpretation of Kant's Motive of Duty Thesis that it fails to recognize Huck Finn's decision to protect Jim as worthy. Arpaly's version of the same basic idea, which I discuss next, avoids this problem.

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." She concludes:

> The first agent is more praiseworthy for her actions that the second agent, because
>
> to act benevolently for moral reasons while one is depressed takes more concern
>
> for those moral reasons than to do so in happy times.[20]

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.

Arpaly takes her example to show that a thesis like the Coincident Reasons Thesis can

provide 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. But I will argue that while the relative fragility of the fair-weather

philanthropist's good motive suggests that she may be less likely to act well *consistently* than her

foul-weather friend, and so may give us some reason to worry about *her moral character, taken*

as a whole, it doesn't make her *individual action* any less worthy. So it doesn't undermine the

Coincident Reasons Thesis.

---

[20] Arpaly, *Unprincipled Virtue*, 87.

The kind of appeal to counterfactuals on which both Arpaly and Stratton-Lake rely 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, should we 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, the dog-lover 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.

To come to grips with what Stratton-Lake and Arpaly find unsatisfying about actions from which such background, regulating dispositions to act on right-making reasons are absent, it is worth thinking more carefully about the idea that in the case of morally worthy actions, it is no accident that the agent acts rightly. It was this idea that prompted Kant's own response to the problem of moral worth. Kant, remember, worried that in the case of any action not performed from the motive of duty,

> conformity [of the action to the moral law] is only very contingent and precarious, since a ground that is not moral will indeed now and then produce actions in conformity with the law, but it will also often produce actions contrary to law.[21]

And Stratton-Lake worries that if duty is not present as a regulating, dispositional motive, "the relation between my motives and the rightness of my action is purely accidental."[22]

---

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

The worry that an action might be right only contingently or accidentally might take two forms. First, and this seems to be the target of Kant's worry, it may be a matter of accident that the motives on which an agent acted caused him to act rightly. This would be the case, for example, if I were motivated by self-interest to, say, save the child trapped in the burning house. The motive of self-interest could cause me to do as morality requires only by accident – the self-interested thing to do *happened* also to be the right thing to do. The rightness of the action, in such a case, tells us nothing about the character of the agent performing it. The nature of an action (be it right, prudent, helpful, and so on) can tell us something about the motive (and by extension, the character of the agent) behind it *only if* an action with that nature is the *non-accidental* result of acting with that motive.[23] Right actions, however, can result from self-interested motives only accidentally.

As I argued earlier, the Coincident Reasons Thesis (like the Motive of Duty Thesis) rightly excludes such acts from having moral worth. Actions motivated by right-making reasons, by contrast, are not merely contingently or accidentally right. If I am motivated by right-making reasons, it is no coincidence that my motive issues in the right action. So Stratton-Lake is, I believe, mistaken in claiming that the relation between an agent's motives and the rightness of his action is "purely accidental" when, as in the case he imagines, he acts for the right-making reasons, but would not have acted for them had his interests been at stake. The type of

---

[22] Stratton-Lake, *Kant, Duty, and Moral Worth*, 65.

[23] I have intentionally phrased this a necessary but not sufficient condition: the nature of an action may not *always* tell us something about the character of its agent, *even when* that agent's motive non-accidentally results in an action of that type. Some motives might non-accidentally result in right actions without thereby endowing those actions with moral worth. This might be the case, for example, if there were a necessarily existing God, who necessarily rewarded all right (but not necessarily worthy) actions in an afterlife, and punished wrong ones. If the existence of such a God was known, then selfish motives might non-accidentally produce right actions, but such actions would still not be worthy, or reflect well on the character of the agent – and the agent would not be performing them for right-making reasons.

accidentality worrying Stratton-Lake, and Arpaly, is different: they worry not that actions motivated by the right-making reasons are only contingently right but rather that a particular *agent* might be only contingently motivated to act on the right-making reasons.

This kind of contingency is impossible to completely eliminate, no matter what our motives. Agents can act from the motive of duty only if their moral reasoning is good (otherwise, they will be motivated only by their *belief* that their act is right, which, as I will argue, is no guarantee at all of the moral worth of their actions). How good our moral reasoning is will depend on many factors that are beyond our control, including the quality of our moral education and, as Huck's case shows, the culture in which we live. Even if we reason well and are motivated to perform only those acts we perceive to be right, we won't always perceive opportunities to do good – we may not always notice, for example, even when we should, that someone is in need of our help. What's more, *very* few people genuinely have duty as a secondary motive, in the sense that there are *no* circumstances under which they could fail to be motivated to perform the acts they believe right, or could fail to be motivated by right-making reasons. We all have our breaking points, whether they're triggered by threats to our own interests, or to the interests of those we love. So a criterion for moral worth according to which our being motivated by the right-making reasons would have to be completely independent of contingent circumstances for our acts to count as morally worthy[24] entails that virtually no acts at all would qualify.

Arpaly sees this, writing, "If we were to believe that only foul-weather, die-hard philanthropists act for moral motives, we would have to believe that only *very* morally virtuous

---

[24] Like that suggested by Stratton-Lake.

people ever act for moral motives and ordinary people never do."[25] Her appeal to a second

criterion for moral worth – the agent's "degree of moral concern" – makes the moral worth of

actions a matter of degree. How morally worthy an act is depends, according to Arpaly, not only

on whether the agent acts for right-making reasons, but also on how easily (or not) she might

have *failed* to act for those reasons. While both the fair-weather and the die-hard philanthropists'

actions have moral worth, the die-hard philanthropist's acts are worthier than her fair-weather

friend's, because she is less easily deterred from doing the right thing by reasons that aren't

morally relevant.

But how are we to determine how *easily* someone is tempted away from doing his duty?

Consider again the fanatical dog-lover of my earlier example. Is he a fair-weather or die-hard

good-doer? When we consider that it would take only a threat to his dog for the dog-lover to

give up the lives of several strangers, he looks quite fair-weather. But, on the other hand, he was

willing to put his own life on the line to save theirs – is he perhaps die-hard after all? Does the

dog-lover care too much about the dog or not enough about the people he is saving, and his own

life?

The worries these questions raise about accounts of moral worth that appeal to

counterfactuals are not merely epistemological. It is significant that the rescuer cannot plausibly

be described as having been motivated to save the endangered strangers by the fact that his dog

was not in jeopardy at the same time. This thought will likely not even have entered the

rescuer's mind. He responds simply to the facts with which he is presented – the threat to the

strangers' lives, and the means of rescuing them that are available to him – although he would

---

[25] Arpaly, *Unprincipled Virtue*, 88-89.

have responded differently had he been presented with other facts as well. The desire to protect his dog at all costs was not a reason motivating his action (if it had been, he could not have been described as acting for the right-making reasons, and so would not have satisfied the conditions for moral worth imposed by the Coincident Reasons Thesis). How relevant to the moral worth of our actions are facts about how we might have been motivated to act under other, counterfactual circumstances?

I'm not convinced that we have a sufficiently developed network of conditional motivational dispositions for our current motivational configuration to be sufficient to determine, always or even often, how we would be motivated in such counterfactual scenarios, and this doubt is grounds for some skepticism about the relevance of such counterfactuals to determining the moral worth of what we actually do. But we can set that worry to one side: even if my psychological profile provides an answer to the question of how I would have been motivated to act in other circumstances, that answer does not help determine the moral worth of my action in my actual circumstances. We do not think a relatively low-cost right action is made *worthier* by the fact that the agent who performs it (for the right-making reasons) would have done so even had the cost to him been higher. So we should not think it *less* worthy because the agent who performs it (still for the same right-making reasons) might *not* have done so had the cost been higher. If saving the group of strangers requires my dog-lover to do no more than toss them a life-saver from the edge of the pier, we would not call his act heroic just because he would have risked his life for them had the emergency required it – his actual act required no particularly heroic motivations. So we should not deem his selfless rescuing of the strangers any less heroic because he might have been motivated less heroically had his circumstances been different.

A similar conclusion should be reached about the relevance of the moral perceptiveness

of the agent (another indicator, according to Arpaly, of what she calls an agent's "degree of

moral concern") to the moral worth of individual right acts, performed for right-making reasons.

When my "die-hard" and my "fair-weather" friends both remember to call me on my birthday,

simply because they know doing so will make me happy, their actions have the same moral

worth, although the former reliably remembers and the latter usually forgets.[26] These intuitions

provide support for what has been called the "battle-citation" model of moral worth:[27] I should

not be condemned for a crime I had no occasion to commit, or honored for a feat of bravery I had

no chance to perform. We are not to be credited with heroic actions or discredited for bad

actions we weren't given the opportunity to carry out.

Comparison to the epistemic case bears out these intuitions. Something like the

Coincident Reasons Thesis plausibly describes the conditions under which our beliefs are

epistemically worthy or justified. Our beliefs have epistemic worth – are epistemically justified

(not just justifiable) – when we believe them for the epistemic reasons why we ought to believe

them – when, that is, we believe them in response to our evidence.   As in the moral case, the fact

that we epistemically ought to believe *p* is not itself an epistemic reason to believe *p*, but rather

reflects the fact that we have other (conclusive) epistemic reasons to believe *p* – that *p* is

---

[26] What if the reason my friends should call me is not that "doing so will make me happy," but rather that remembering to call is the kind of behavior that is constitutive of loyalty, which is valuable in itself? Doesn't it then make a difference whether my friend is fair-weather or die-hard – whether, for example, she remembers my birthday reliably or not? In this case it may well matter, but it matters in a way that is easily accommodated by the Coincident Reasons Thesis. My friend has reason to be loyal, and being loyal is not something she can do intermittently, when the thought happens to strike her. Being loyal is not something you can accomplish with a single act. So where *loyalty is required*, a fair-weather friend does not just fail to do the right thing for the right-making reasons; she doesn't do the right – the *loyal* – thing at all. Thanks to Daniel Markovits for pressing me on this point.

[27] See Richard Henson, "What Kant Might Have Said: Moral Worth and the Over-determination of Dutiful Action," *Philosophical Review* 88 (1979): 39-54, at 48. Henson uses the comparison to battle-citations in a different context.

supported by the evidence. Our beliefs have epistemic worth whenever we believe them for the

epistemic reasons why we ought (in response to the evidence) even if we might have failed to

believe as we ought had other, misleading, factors been present. If my reasoning is correct, my

belief has worth, even if my reasoning might have been faulty in slightly altered circumstances –

had, say, some further, misleading evidence also been present, or had I been distracted by other

worries and therefore been focusing less well.[28] And in the practical, as in the epistemic case, it

is how well an agent *actually* responds to her reasons that determines her action's moral worth.

I've argued that an appeal to how the agent might have acted counterfactually[29] is not

relevant to the determination of the moral worth of actions. But we might have a different worry

about Huck's act, and correspondingly, a different worry about whether the Coincident Reasons

Thesis identifies sufficient conditions for the moral worth of actions. We might think there is

something unworthy about an agent's performing an action despite *believing* it to be wrong (even

if that normative belief is false). Should we think Huck Finn less morally worthy because he acts

as he believes he ought not? 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.[30] Indeed, I think that our normative beliefs may

---

[28] Imagine that Paul, who is taking a difficult math examination, works out the answer to each question perfectly. His achievement would be no less epistemically worthy if it were true of him that, had the lovely Linda been seated at the next desk, he would have been so distracted that he would have gotten every question wrong.

[29] Such as that suggested by Arpaly and Stratton-Lake.

[30] Stratton-Lake denies that "verdictive" normative beliefs can give us reasons in this way (see *Kant, Duty, and Moral Worth*, 12, 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" (20) because "no verdict constitutes evidence for itself" (19). 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.

sometimes provide us with normative reasons of a particular kind: sometimes the costs of re-opening deliberation on a moral question about which we've already formed a judgment give us reasons not to deliberate further, but act on the basis of that judgment. That is, our moral judgments can serve a function similar to that served by rules of thumb. However, like a rule of thumb, 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 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.

Some philosophers, Christine Korsgaard among them, do believe that if we judge we ought to act a certain way, we have some reason to act (or at least, intend to act) in that way, even if our ought-judgments are mistaken and generally unreliable.[31] I'll follow John Broome in calling such a normative requirement, according to which we ought to, or at least have reason to, act (or intend to act) as we believe we ought to act, a narrow-scope enkratic condition.[32] The condition is *en*kratic, because it tells us not to be *a*kratic (like Huck); it is narrow-scope, because only the consequent of the conditional enters into the scope of the normative requirement (the ought-claim): if you believe you ought to ϕ, then you ought, or have reason, to ϕ. According to

---

[31] See, for example, Korsgaard's response in *The Sources of Normativity* (Cambridge: Cambridge University Press, 1996), 257 to G.A. Cohen's example of the Mafioso, who feels bound by his gangster's code of strength and honor to perform merciless acts:

> I want to say of the Mafioso what I said of the Knight …who felt himself to be obligated to fight a duel. There is a sense in which these obligations are real – not just psychologically but normatively. And this is because it is the endorsement, not the explanations and arguments that provide the material for the endorsement, that does normative work.

[32] Broome calls such a condition "enkratic" in, for example, "Does Rationality Consist in Responding Correctly to Reasons?", *Journal of Moral Philosophy* 4 (2007): 349-374, in sec. 4 (hereafter, "Does Rationality Consist?"). He talks about wide- versus narrow-scope versions of the requirement in, for example, "Wide or Narrow Scope?" *Mind* 116 (2007): 359-370.

a narrow-scope enkratic requirement, our ought-judgments, regardless of their merit, do give us normative reasons to do or intend particular things.

I'll also follow Broome in maintaining that if there is an enkratic condition that normatively binds us, it must be wide-scope, not narrow-scope. That is, it is at most true that we ought to be such that our actions (or intentions) conform to our beliefs about how we ought to act. Wide-scope requirements tell us to maintain certain relations between mental states: in this case, not both to judge we ought to do something, and fail to (intend to) do it.[33] They never tell us to do or intend particular things because there are always two ways to satisfy such requirements: we can change our actions, or we can change our beliefs about what we ought to do. When he decides to protect Jim, Huck clearly violates a wide-scope enkratic requirement, and this may explain why we feel somewhat dissatisfied with him. But it doesn't follow that there is anything wrong with his action – which, I have stipulated, he performs for the right (that is, right-making) reasons. That Huck is failing to comply with a normative enkratic requirement does not entail he has reason to do as he judges he ought, and turn Jim in. It entails that there must be *something* wrong, *either* with his action, *or* with his (theoretical) moral reasoning. And in Huck's case, it is surely more plausible to conclude that the problem lies in his moral reasoning – the process by which he forms moral beliefs – which we know to be deeply flawed.

As Broome, Niko Kolodny, and others have argued, a narrow-scope enkratic requirement seems to entail a highly implausible form of bootstrapping. It says that our ought-judgments can make themselves true, or at least that our beliefs about our obligations, no matter how far off the

---

[33] Broome adds the qualifier, "if [we] believe [we] will [do it] if and only if [we] intend to [do it]." See "Does Rationality Consist?", 361.

mark, can create normative reasons for us where there were none before.[34] And if our ought-

judgments can't (generally) give us normative reasons to act as we judge we ought I'm very

skeptical that there is anything worthier about acting as we judge we ought, or less worthy about

failing to do so. Göbbels' persecution of the Jews is no less despicable because he believed he

was acting rightly.[35] And as I will argue in the next section, even right acts cannot be made

worthy by the fact that an agent (merely) believes himself to be acting as he ought. So we

shouldn't think worse of Huck's action because he believes he is wrong to perform it, though

Huck may be to blame for his faulty moral reasoning.[36]

3   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. I've argued that neither counterfactual considerations

---

[34] See, for example, Broome, "Normative Requirements," *Ratio* 12 (1999): 404; Broome, "Wide or Narrow Scope?"; Nico Kolodny, "Why be rational?", *Mind* 114 (2005): 509-63.

[35] If anything, there is something particularly sinister about agents who act terribly wrongly while claiming to have right on their side.

[36] 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 (as Krister Bykvist has pointed out to me; 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 "one thought too many."). Is this grounds for objecting to the thesis? One might worry that, if, as I've argued, acting worthily does not require awareness of the reasons for which one acts as reasons, or even any normative beliefs at all, then many examples of *non-rational animal* action will suddenly qualify as morally worthy.
    A quick response to the worry posed by non-rational animals is that the Coincident Reasons Thesis does not in fact entail that their actions have moral worth: since such animals lack the higher rational capacities required for moral responsibility, no normative reasons apply to them, so their motivating reasons cannot coincide with their normative reasons (see n. 9). This response, however, feels unsatisfactorily glib, since part of what is at issue is what kind of rational feature such animals' actions lack (the absence of which exempt them from moral obligations and their actions from moral worth). While I feel there is much more to be said about the matter, for example, I suspect that the lower animals also don't properly act for *motivating* reasons, in the way in which I understand the term (see my discussion in sec. 3), I set these further issues aside here.

about how an agent might have acted in altered circumstances nor the agent's normative beliefs about his own actions are relevant to their moral evaluation. 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 morally justify an action? Are there other morally worthy motives, which are not provided by right-making reasons? If the answer to either of 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 morally justify the act.

Right actions that are motivated (merely) by an agent's true *belief* that they are right do 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 have moral worth. The mechanism that triggers Huck's motivation in this imagined case is too distantly and unreliably related to the true moral 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.[37]

---

[37] The problem with Huck's reasoning in cases like this is not just that it would lead him astray in other circumstances. As I argued in the previous section, counterfactual considerations like this don't tell us

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 reasons on which the authority bases her 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 does not provide an example of a morally-worth action that 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 morally *justifies* the agent's act – that *makes it 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. This is why, for example, doctors are morally obligated to prescribe the course of treatment their evidence tells them is most likely to cure their patients, not the treatment that (against all evidence) happens to be best; if all the evidence suggests that I need penicillin, and my doctor has no evidence that I'm allergic, she fails to fulfill her obligations if she refuses me penicillin, even if it turns out I am allergic. Cases like these illustrate that our normative reasons depend on our evidence. So: if the advice of the good authority provides our agent with sufficient evidence for the belief that a particular act would be best, our agent's 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. The agent's belief that the act is right, which motivates him to act, is based on the good authority's advice. And his

---

much about the moral worth of actions; the problem with relying on bad authorities is, as the Coincident Reasons Thesis highlights, that actions can't be made *right* by the fact that they're recommended by these authorities: a bad authority's recommendation to do something is no reason to do it.

act is right *because* it follows the authority's advice. So in this case the reasons motivating and

morally justifying our agent's act coincide after all.[38]

The subjective nature of moral reasons shields the Coincident Reasons Thesis against

another kind of worry – a worry again raised by Philip Stratton-Lake. Such a thesis cannot, he

argues, provide necessary conditions for the moral worth of agents, because some morally

worthy agents may blamelessly fail to comply with it:

> There is nothing about the concept of a morally virtuous agent that entails …that
> she will never be ignorant of the relevant facts. Someone might be perfectly
> virtuous, but nonetheless motivated to do what she thinks she should by the
> thought of something other than the reason, or reasons why she should do this act;
> and this might be because she is ignorant of some of the relevant facts.[39]

Stratton-Lake imagines a case in which an agent's (blamelessly) false beliefs about the source of

her friend's pain seem to cause a disconnect between her motivating reasons for helping him (she

---

[38] There is an important difference between relying on someone else's judgment about non-normative
questions and relying on her judgment about normative questions. Cases in which another person's
verdictive moral judgment *that some act I can perform is right* provides me with sufficient evidence that
performing it would be best, and so with conclusive reason to perform it, without giving me any access to
the underlying reasons on which she forms her judgment, will be very rare. In other words, morality
places unusually strict limits on permissible deference to normative authorities, even accurate authorities.
Usually, we must do our moral reasoning ourselves.
   This is an important feature of the practice of moral judgment that I can't explain satisfactorily here.
But *part* of the explanation of the relative rarity of permissible deference to normative authorities,
compared to other kinds of authorities, might lie in this: in order to be justified in allowing an authority's
judgment to guide our actions, we must be in a good position to judge that the authority is a reliable one –
that is, likely to be right about the subject-matter of her expertise – *without* being in a position to form
reliable beliefs about that subject matter ourselves. Doctors are a good example of authorities we can
judge in this way – their track record, for example, allows us to judge their reliability despite not being in
a position to form sound judgments about the medical facts by ourselves. But we are rarely in the
position with respect to purported moral authorities (whose moral "track records," for example, we
generally can't judge unless we *can* form sound moral judgments ourselves).

[39] Stratton-Lake, *Kant, Duty, and Moral Worth*, 16-17.

thinks he has a headache) and the reasons morally justifying her help (according to Stratton-Lake, that he has an earache).[40]

But the appearance of a disconnect between motivating and justifying reasons in cases of blameless ignorance is misleading. Because the reasons relevant to moral-ought claims are subjective – they depend on what an agent ought to believe about her situation – our normative reasons for acting *can't* be given by facts of which we're *blamelessly* ignorant (such as the fact, in Stratton-Lake's example, that the friend has an earache). I can't have an obligation to do something I have no reason to believe should be done. The normative reason to help in this case must therefore be provided by a different fact – one that the agent ought to believe in (perhaps, in this case, that her friend *appears* to be in pain, or *appears* to have a headache, or even that she justifiably *believes* he has a headache). And each of these is a plausible account of the agent's motivating reasons for offering the help, as well. So this, too, seems to be a case in which motivating and morally justifying reasons coincide.[41]

It is worth thinking more carefully about how what we know can affect the reasons for which we act and stating more precisely what I take motivating reasons to be. Since I argue that normative and motivating reasons can overlap (and do overlap in the case of morally worthy

---

[40] See ibid., 17.

[41] Should we instead think that while the *moral worth* of our actions is relative to our evidence, the rightness of our actions is not? Should we think, for example, that a doctor who offers her patient the course of treatment that her evidence (misleadingly) tells her is most likely to effect a cure acts worthily but not rightly (and not in response to genuine normative reasons)? That an agent who fails to offer aspirin to a friend who gives no outward indication of his discomfort is acting wrongly but not blameworthily? I think it would be a mistake to interpret such cases in this way: it seems clear, in the doctor case, that the doctor fulfills her obligations, acts as she is morally required to act, and is responding to her normative reasons. And we should not say, in this headache case, that the agent is morally required to offer her silently suffering friend an aspirin. Rightness, moral requirements, and normative reasons go hand in hand: they all operate at the same level of evaluation. Of course, it would, in these cases, be fortunate if our doctor irresponsibly chose the more effective treatment, which her evidence did not support, and if our second agent offered her friend the aspirin anyway, on a whim, perhaps.

actions) and since I take normative reasons to be facts, I also take motivating reasons to be facts.

Because it must usually be possible for a particular normative reason applying to an agent to also

serve as the agent's motivating reason[42] and because not all of an agent's normative reasons are

determined by her desires, I cannot be using the term 'motivating reason' as it is often used – to

pick out (exclusively) belief-desire pairs. I propose that motivating reasons are the kinds of facts

we are after when we ask about an agent, "what were *her reasons* for acting as she did?" – those

that appear in what have been called "rationalizing explanations." Other kinds of facts can also,

of course, explain our actions, such as biochemical facts about our brains, or facts about how

much sleep we've been getting, or how much coffee we've drunk; and these might also be called

'reasons.'[43] But they could not be described as "*our reasons* for acting." I might snap at you

because I have gotten enough sleep lately, but this can't be *my reason* for snapping at you – it

can't be my motivating reason.[44] My motivating reason will always be some fact on the basis of

which I chose to snap.[45]

---

[42] Why "usually"? I will discuss in the next section some cases in which it may not be possible for an agent to be motivated to perform an action by the normative reasons why it is right. Such cases, if they are possible, entail that it is *sometimes* not possible for an agent to perform the right action worthily. I will argue that such cases, if they exist, will be rare, on any plausible normative theory.

[43] As is often the case with explanations, these kinds of explanations need not crowd out rationalizing explanations – both may be true at the same time.

[44] It's natural to think of rationalizing explanations as a species of causal explanation. I won't consider this proposal in any detail here. But if motivating reasons are causes, then, if the Coincident Reasons Thesis is right, we should expect (1) that cases in which it is particularly difficult to identify the cause (at the level of motivating reasons) of an agent's action will also be cases in which it's more difficult to assess the action's moral worth; and (2) that the irrelevance of counterfactuals about how an agent might have behaved in altered circumstances to determining the moral worth of her actions will be mirrored by the irrelevance of such counterfactuals to determining the motivating reasons which caused her to act.

It seems to me that our intuitions about moral worth do falter where our intuitions about cause falter. And perhaps the most prominent defender of a causalist account of motivating reasons, Alfred Mele, does argue that appeals to counterfactuals cannot reliably reveal an agent's motivating reasons – see Mele, *Motivation and Agency* (Oxford: Oxford University Press, 2003), 51.

[45] To say this is not to say that all motivating reasons are facts we *take* to be normative reasons, and certainly not to say that they are facts we take to be *sufficient* to justify our actions. We might sincerely

If people were always perfectly self-aware and sincere, their account of what prompted them to choose to act as they did would always provide us with their motivating reasons. But people can, of course, be self-deceptive about their motivations, and just as importantly, agents' awareness of their own motivations can be affected by their ignorance about the world in which they act. We aren't always well-placed to recognize the difference between our knowing some fact and our merely believing it. Let's say I come to believe (as it happens, falsely) that my child has swallowed a marble, and so rush her to the emergency room. If I am asked why I did so, I might well reply, in all sincerity, and without self-deception, that I did it because she swallowed a marble. I might believe this was my motivating reason for acting. Had my belief amounted to knowledge, the fact that my child swallowed a marble (that is, the propositional content of my belief) would have constituted my motivating reason. But when, as in my imagined case, our beliefs fall short of knowledge, it may be a fact about how things appear, or sometimes even the fact *that we believe*, rather than the content of our belief, that serves as our motivating reason.

This does not undermine the claim, to which I appealed a moment ago, that motivating reasons are the kinds of facts that typically appear in rationalizing explanations. When we explain our actions, we generally report *either* our beliefs *or* their contents. Usually, which one we appeal to will be partly a matter of the level of credence we have in our beliefs, and of how conscious we are of the perceptual stimuli to which we are responding. Thus I might say that I rushed my child to the emergency room because she swallowed a marble, or because it looked like she swallowed a marble, or because I believed she swallowed a marble. (Similarly, although

---

offer a "rationalizing" explanation even when we believe our reasons for acting did not justify our action. I might offer as my reason for snapping at you the fact that your voice is rather shrill for this time in the morning, even if I know that the tone of your voice is no justification for snapping at you at all. And Huck could offer a sincere rationalizing explanation of his decision to protect Jim.

the woman in Stratton-Lake's example might *think* she offered her friend an aspirin because he had a headache, if we pointed out to her that her friend had an earache, and not a headache, she would correct herself, and offer a revised rationalizing explanation for her action – one that appealed instead to his *appearing* to have a headache, or her *belief* that he had one.)

Now let's look again at the possibility of overlap between normative and motivating reasons in cases of blameless ignorance. If my child swallows a marble, and I *know* it, I may be motivated by the fact that she did so to rush her to the emergency room, and this fact will also morally justify my action – the action is morally worthy. But let's say instead that I believe but *don't know* that my child swallowed a marble. And let's say I am motivated by the fact that I believe this to take her to the emergency room. Does my act have moral worth? According to the Coincident Reasons Thesis, it has moral worth only if the fact that I believe she swallowed a marble is also the reason morally justifying my action. Is it? Well, this will depend on the details of the case. If I'm not generally given to paranoia, then the fact that I believe she swallowed a marble probably is a good (normative) reason to go to the emergency room, and the thesis rightly lauds the act. If I am given to paranoia, then my belief may be a normative reason to take *myself* to see a doctor, not my child. Now there is no overlap between motivating and justifying reasons (the act is not justified), and the thesis rightly declares the act not worthy. (It does not follow, of course, that a parent who does nothing to help his child when he irrationally but firmly believes her health to be threatened, out of some kind of callousness or laziness, acts worthily. Such a parent is certainly not motivated by right-making reasons.)[46]

---

[46] **A** paranoid parent who frequently rushes his perfectly healthy child to the doctor may mean well, but his actions are not fully worthy; in sec. 5, I show how the Coincident Reasons Thesis can explain why some well-meaning but wrong actions can be partially morally worthy.

Blameless ignorance does not prevent agents from acting for right-making reasons, so we needn't appeal to an alternative account of moral worth to explain the worthiness of actions performed out of blameless ignorance. But it remains possible that some actions that are not motivated by right-making reasons are nonetheless morally worthy. The Motive of Duty Thesis appeals to us in part because we tend the think that virtue is a matter of living according to one's conscience. I've argued that this is a mistake because our consciences do not always guide us well: Huck Finn's sheltering of Jim is virtuous despite the fact that he does it with a guilty conscience, and the fact that Göbbels was driven by his conscience to persecute the Jews does not exonerate him, much less endow his acts with moral worth. But what if our moral judgment is sound? If we are reliable judges of right and wrong and perform right acts out of a desire to keep a clear conscience, do our acts have moral worth? And what about those people, like the "sympathetically attuned" soul Kant imagines in the passages I quote at the outset of this paper, who do good out of the satisfaction they get at spreading joy around them? That doing so will make us happy is not the reason we ought to help those in need, but isn't there nonetheless something to be said for the person who does the right thing for the sheer joy of it? Do the conscientious do-gooder and the joyous one show that we need not act for right-making reasons for our actions to have moral worth?

Kant was right to exclude right actions performed just for the satisfaction of doing right from having moral worth, and the Coincident Reasons Thesis helps explain why such actions, and similar actions performed out of a nagging conscience, lack moral worth. Kant's defenders are quick to point out that doing good because it makes you happy is, on his view, neither *wrong* nor *vicious* (indeed, Kant considered such actions deserving of "praise and encouragement"[47]);

---

[47] Kant, 11 (4:398).

that the fact that you are made happy by acting rightly (or feel guilty if you fail to do so) does not preclude your act from having moral worth so long as this is not your motivating reason for acting rightly; and that right actions might be over-determined by an agent's motives.[48] Similarly, there is nothing in the Coincident Reasons Thesis that entails that only the miserable can be virtuous.

But people who act rightly *just because* it makes them feel good, or *just because* failing to do so makes them feel bad, act for reasons that are fundamentally selfish – self-directed – and not (significantly) morally relevant. We should be worried about someone who does the right thing only or primarily because of the pleasure it affords her, rather than because of the reasons why it is the right thing to do. Such a person cares not simply that good be done, but rather that *she personally* be the one to do it. Her preferences would lead her to supplant other agents in a position to do good even when they're better-placed to help than she is, and so could do good more effectively. This is not a morally admirable characteristic, and the Coincident Reasons Thesis explains why.[49]

---

[48] I won't tackle here the difficult problem of how to identify which of various "incentives" an agent has to perform an action serves as her motive, or the equally difficult problem of what it is for an action to be over-determined by an agent's motives. Both topics have received significant attention from Kant scholars writing about the motive of duty. See, for examples, Barbara Herman, "On the Value of Acting from the Motive of Duty," in *The Practice of Moral Judgment* (Cambridge, MA: Harvard University Press, 1993); and Marcia Baron, *Kantian Ethics Almost Without Apology* (Ithaca, NY: Cornell University Press, 1995); as well as Stratton-Lake, *Kant, Duty and Moral Worth*; and Henson, "What Kant Might Have Said.")

[49] Although pleasure is the motive that interests me here, it is not, of course, the only possible selfish motive an agent might have for doing good. There are a number of reasons agents might care that *they do good*, rather than that *good be done*. They might, for example, want the reputation of a do-gooder. Or they might simply be too exclusively focused on their own virtue – and self-defeatingly so. As Hume pointed out, to suppose that a virtuous (that is, virtuous-making) action is one performed out of regard for its own virtue is to reason in a circle. Hume, *A Treatise of Human Nature*, ed. by L.A. Selby-Bigge and P.H. Nidditch (Oxford: Oxford University Press, 1740/1978), 478. And in any case, the fact that performing some act is what a virtuous person would do is, like the fact that an act is right, not a reason to perform it, but rather reflects the fact that there are other, conclusive, reasons to perform it. So the

But importantly, the Coincident Reasons Thesis can help explain not just why such actions lack moral worth but also why we nonetheless tend to think (rightly) that people who have sensitive consciences or who feel joy at doing right are virtuous.

Someone who is made happy by doing good for others usually cares about them noninstrumentally as well – not just as tools useful for promoting her own well-being. Otherwise, it is not clear why doing good for others should make her happy. Nor would someone be likely to feel guilt at failing to help others in need if she did not care noninstrumentally about their well-being, too. People who exhibit sensitive consciences or joy at acting rightly are reliably also motivated directly by the reasons why their actions are right. This is borne out by the fact that most people who are motivated to do right by a desire to "sleep easy" would not feel that they could accomplish their goal equally well by, say, taking a pill that could make them lose their scruples – or even a pill that would make them believe they had acted rightly. It is certainly possible for agents to act worthily despite not being susceptible to a guilty conscience and despite not feeling joy at doing right. Such emotional propensities are not required for moral worth.[50] But the Coincident Reasons Thesis helps explain why we nonetheless take a conscience that is sensitive to right-making reasons or a temperament that gets

---

Coincident Reasons Thesis helps explain why acts motivated (soley) by an agent's desire for her own virtue aren't worthy. But I don't mean to suggest that wanting *not just* that good be done but *also* to play a role in doing it is always a morally unattractive desire. Sometimes we might have good normative reasons to ensure that we be the ones to bring about some good – for example, when fairness demands that we contribute our share to a collective good.

Thanks to Daniel Markovits and to an anonymous reviewer for the *Philosophical Review* for helpful discussion of these issues.

[50] Here again, my account of moral worth departs from that offered by Nomy Arpaly. She asserts, in stark contrast to Kant, that "the cold-hearted philanthropist is less praise-worthy than he would be if it were not for his cold-heartedness." *Unprincipled Virtue*, 89. This is another example of a case in which Arpaly thinks a thesis like the Coincident Reasons Thesis cannot, on its own, account for the moral worth of actions.

satisfaction out of doing right as *evidence* of moral worth: such qualities are evidence of

noninstrumental motivational responsiveness to right-making reasons as well.

There is a related worry about the supposed coldness of morally virtuous agents on

accounts of moral worth like Kant's and like the one I have been defending. The worry shows

up particularly in the case of actions in which the motivating role played by emotions the are not

impartial, such as love, seems at least appropriate – even morally attractive – and in which

partiality on the part of agents seems, if not required, at the very least permissible. Thus Bernard

Williams considers the example of a man who is in a position to rescue one of two people in

peril, and one of them is his wife. Williams worries that even if morality allows, or even

requires, that the man save his wife under these circumstances, the justification for this choice –

the normative moral reasons why it is right – will be, at the most fundamental level, impartial.

But, Williams contends,

> This construction provides the agent with one thought too many: it might have
>
> been hoped by some (for instance, by his wife) that his motivating thought, fully
>
> spelled out, would be the thought that it was his wife, not that it was his wife and
>
> that in situations of this kind it is permissible to save one's wife.

> Perhaps others will have other feelings about this case. But the point is
>
> that somewhere (and if not in this case, where?) one reaches the necessity that
>
> such things as deep attachment to other persons will express themselves in the

world in ways which cannot at the same time embody the impartial view, and that

they also run the risk of offending against it.[51]

If a man chooses to rescue his wife, and does it for the simple reason, fully spelled out,

that she's his wife, or even for the reason, fully spelled out, that *that's Sarah*, does the

Coincident Reasons Thesis improperly imply that his act has no moral worth? Even if moral

reasons are always, fundamentally impartial, must morally worthy action always be impartially

motivated? Need we, if our actions are to be morally worthy, learn to think of anyone we love as

just one person among others – perhaps one person among others whom we happen to love,

where this may be a morally relevant fact about them – whenever we act on their behalf?

I'm not convinced that our ability to form meaningful relationships based on emotions

like love is undermined by our learning to recognize that those we care most about have no

special moral status, in virtue of their relationship to us, which strangers (who are, perhaps, just

as loved by their own family and friends) lack. After all, learning to see ourselves this way is an

essential part of our moral education, and yet even if we learn to recognize this fact, we do not

become alienated from our own interests – we do not cease to be able to think about them as

special *to us*. We have the capacity to view ourselves, as Thomas Nagel puts it, "simultaneously

as 'I' and as *someone*."[52]

But in any case, if we allow, as Williams does, that saving one's wife in such

circumstances can be justified from the impartial point of view, then it may well turn out that

"she's my wife" (or even: "that's Sarah!") *is* a sufficient normative moral reason for this man to

---

[51] Williams, "Persons, Character, and Morality," 18.

[52] Thomas Nagel, *The Possibility of Altruism* (Oxford: Oxford University Press, 1970), 19.

choose to save one of two people, rather than the other. And if this is a sufficient normative moral reason, then if it is also the man's motivating reason, this may be enough for his act to count as morally worthy, whether or not he is motivated by additional facts explaining *why* he is, in these circumstances, permitted to choose to save his wife.

Whether this does turn out is a complicated matter, and one that points to a need to refine the Coincident Reasons Thesis. 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. We generally act for chains of dependent motivating reasons, running from the less to the more fundamental. Both the self-interested reward-seeker and the altruist may rush into the burning house to save the trapped child. The fact that rushing in will allow them to rescue it is, of course, a normative moral reason – indeed, a sufficient one – to do so. So both the reward-seeker's and the altruist's motivating reasons overlap in some sense with the normative reasons morally justifying their act. But the reward-seeker's fundamental motivating reason is *not* also a significant normative moral reason to save the child – her chain of motivating reasons diverges in its more fundamental links from the chain of normative moral reasons, whereas the altruist's does not. And the reward-seeker treats as *merely instrumental* a consideration (that rushing in may save a life) that in fact provides a *noninstrumental* moral reason for acting as she did. *This* seems to be why her action has no moral worth.

But should we conclude that, for our actions to have moral worth, our most fundamental motivating reasons for acting must coincide with the most fundamental moral justification of our action? The example of Williams' loving husband suggests otherwise. And his "one thought too many" worry has important wider implications for the plausibility of the claim that the

Coincident Reasons Thesis identifies necessary conditions for the moral worth of actions. For the conditions established by the thesis would be far too difficult to meet if they required that every agent recognize the kind of fundamental justification for their actions that even moral philosophers, who spend their lives arguing about such matters, struggle to identify. Let's assume, for the sake of argument, that Kant's formula of humanity provides the right fundamental account of what makes actions right or wrong: the most fundamental reason why performing some act is right is that doing so amounts to treating those people affected by our actions as ends in themselves, not mere means, and to respecting the unconditional value they have as rational beings. But surely an agent needn't be motivated by *this* fact in order for her act to have moral worth. Committed utilitarians, for example, are certainly capable of worthy action. We need a more nuanced way of explaining why the loving husband (probably) and the altruist (certainly) and the utilitarian (usually) act worthily, while the reward-seeker does not.

Here a distinction to which Christine Korsgaard has drawn attention is very useful. Korsgaard points out that although moral philosophers often contrast intrinsic with instrumental value, there are in fact two distinctions in goodness: to say that something is *noninstrumentally* valuable, as opposed to *instrumentally* valuable, is not the same as attributing to it *intrinsic*, as opposed to *extrinsic*, value. To describe something as noninstrumentally valuable – valuable as an end – is to say something about the *way* we value it: we value it for its own sake. To describe something as intrinsically valuable is to say something not about the way we value it but about the conditions for its value. More particularly, its value is unconditional – it doesn't get its value from the value of any other thing, but rather, as Korsgaard puts it, "has its goodness in itself." Merely instrumentally valuable things or actions are never intrinsically (unconditionally) valuable. But noninstrumental value need not be unconditional: we value many things as ends,

but most of those ends are not unconditionally valuable. On the Kantian view, only rational

nature is unconditionally (intrinsically) valuable. On the classical utilitarian view, by contrast,

only happiness is intrinsically valuable.[53]

However, utilitarians, ordinary moral thinkers, and Kantians can agree more about what's

noninstrumentally valuable, and thus a source of noninstrumental reasons, than they can about

what's intrinsically (unconditionally) valuable. Happiness provides an important example of

this. Though happiness is not always reason-giving (it's plausible, for example, that the

happiness of the sadistic torturer does not give us a reason help him achieve it) and not

unconditionally reason-giving (if Kant is right, then the value of rational nature is a condition for

the value of happiness), happiness, when it *is* valuable, is plausibly valuable as an end –

noninstrumentally valuable. It is a source of noninstrumental reasons.

Noninstrumental reasons are those reasons to act that are provided by the ends of our

actions that are worth pursuing for their own sake (in the case of normative reasons), or that we

pursue for their own sake (in the case of motivating reasons). Thus, as I'm using the term, we

have noninstrumental reasons to pursue even purely instrumentally valuable actions – they're

provided by the end, valued for its own sake, to which we perform those actions. Our

fundamental reasons for acting, normative and motivating, are always noninstrumental – our

reasons cannot be instrumental all the way down.[54]

---

[53] Korsgaard, "Two Distinctions in Goodness," in *Creating the Kingdom of Ends*, 250.

[54] Noninstrumentally valuable ends can sometimes be sources of noninstrumental (normative and motivating) reasons to perform actions even when those actions won't actually promote the ends in question (for example, when appearances are misleading). For example, the noninstrumental value of relieving someone's pain can give me a noninstrumental normative reason to offer a friend an aspirin even when she merely *appears* to have a headache, or a noninstrumental normative reason to give my child the medications recommended by her (generally reliable) doctor even when the doctor's prescription

It is very plausible that, on the correct normative moral theory, the fact that one of the people drowning is his wife will provide the man in our example with a noninstrumental normative moral reason to save her, rather than a stranger, even if it does not provide the most fundamental justification for doing so. In that case, the Coincident Reasons Thesis would allow that such an act has moral worth. And on *any* plausible moral theory, the fact that rushing into a burning house will allow one save a child from much pain, suffering, and an early death provides a noninstrumental moral reason to rush in. By contrast, as we've seen, our imagined self-interested reward-seeker sees the fact that rushing in will allow her to save a child's life and protect her from pain and suffering as providing a merely instrumental reason to do so, since she values saving the child from suffering and death merely instrumentally. The fact that rushing in is a means to promoting her own well-being is her sole noninstrumental motivating reason. But this fact is not a (significant) noninstrumental, normative moral reason.

We should understand the Coincident Reasons Thesis as pronouncing an action morally worthy whenever the noninstrumental reasons for which it is performed coincide with the noninstrumental reasons that morally justify its performance. So if the Coincident Reasons Thesis identifies the correct criteria for the moral worth of actions, and if Kant's formula of humanity identifies the correct criteria for their rightness, most right actions motivated on utilitarian grounds, or by common-sense morality, will still have moral worth. In those (many) cases in which the utilitarian is right to promote the general happiness, and the ordinary moral

---

won't in fact help. My end of relieving pain, when I value it for its own sake, can also provide me with noninstrumental motivating reasons to do these things. My (normative and motivating) reasons in each case will be the fact that my friend *appeared* to be in pain and the fact that the medication *appeared* to be most likely to effect a cure.

thinker is right to protect those he loves, they are noninstrumentally motivated by noninstrumentally morally justifying reasons.

4   A Problem for Utilitarianism

But if the Coincident Reasons Thesis is right, and the Principle of Utility instead identified the correct criteria for the rightness of acts, the same could not be said of the agent who acts rightly on Kantian grounds, or who does so for some of the reasons picked out by common-sense morality. Let's say, for example, that someone is motivated to perform some right action – perhaps, keep an important appointment – by the fact that she promised a friend she would do so. If she is a Kantian, she may be motivated to keep her promise by the thought that doing so respects her friend's value as a rational being. Or she may, if she is an ordinary moral agent, take the fact that she promised as itself providing her with a fundamental moral reason for keeping the appointment. If utilitarian is right, however, neither of these facts will provide noninstrumental reasons for keeping the appointment. The appointment ought to be kept because keeping it maximizes utility – the balance of pleasure over pain. Keeping promises or respecting rational nature are only *instrumentally* valuable, as means of promoting this end. When Kantians or ordinary moral agents are motivated by such reasons, their noninstrumental motivating reasons fail, according to utilitarianism, to overlap with the noninstrumental normative reasons in virtue of which their actions are right. So utilitarianism, in conjunction with the Coincident Reasons Thesis, seems to entail that many actions performed by Kantians or common-sense moralists – actions that seem intuitively to be morally worthy – lack moral worth. And this seems to me to tell against the plausibility of utilitarianism.

The problem is exacerbated by the fact that utilitarianism is arguably *indirectly self-defeating*: if we *tried* to achieve the aims utilitarianism gives us – that is, if we *tried* to maximize net happiness – these aims would be, on the whole, worse achieved than if we were motivated differently.[55] Happiness is notoriously difficult to achieve by aiming at it. In order to be happy, we must care about things other than happiness as ends and not merely as means to achieving happiness. Many of our most valuable endeavors – writing, painting, playing music, excelling at sports, the pursuit of knowledge – bring us the satisfaction they do only because we see them as valuable independently of their ability to bring us happiness. If we viewed possible relationships with other people as mere means towards achieving happiness, it is unlikely that we would be able to form the kinds of bonds with them that could bring this happiness. We must care about family, friends, and lovers in their own right and not just as instruments for producing happiness if our relations with them are to produce happiness.[56] Other practices that seem essential to human happiness would be unlikely to survive among a people motivated entirely on utilitarian grounds. Promise-keeping, for example, could not survive among agents who lost their motivation to keep their promises as soon as they became convinced that utility would be better

---

[55] I take the term from Derek Parfit, who provides a sophisticated discussion of the ways in which utilitarianism, and other versions of consequentialism, might be self-defeating. See *Reasons and Persons* (Oxford: Oxford University Press, 1984), chap. 1, especially 5 and sec. 10. Parfit argues that the fact that a thesis is self-defeating may not tell against its plausibility. I will argue that, if the Coincident Reasons Thesis identifies the right necessary conditions for the moral worth of actions, the fact that a thesis is self-defeating may well tell against its plausibility.

[56] The worry described here targets a simple, hedonist version of utilitarianism according to which the only thing that is valuable as an end is pleasurable experience. One might hold a more complicated utilitarian view according to which certain pleasurable activities and relationships are themselves *components* of happiness, and so are valuable as ends. This kind of utilitarianism would be less vulnerable to this aspect of the self-defeatingness charge, and so to the kind of worry I'm raising. But the problems raised by valuable practices, by the demandingness of applying the principle, and by Williams' worry about a debasing of "the moral currency," which I go on to discuss, apply to this version of utilitarianism also. See also Roger Crisp, *Mill on Utilitarianism* (London: Routledge, 1997), 26-28, in which Crisp argues that Mill held the simpler utilitarian view not the more complicated one.

served by breaking them. "I promise I'll do it" would come to be no more reliable an indicator

of future intentions than "I'll do it if I judge it best."

It has also often been argued that the utilitarian principle is too difficult to apply because

it is too demanding, either intellectually or emotionally. Then there is the worry that, as Bernard

Williams put it, utilitarianism might "debase the moral currency": in a community in which

people are motivated by utilitarian ideals, an "escalation of pre-emptive activity" might ensue,

resulting in a society in which "the bad acts of bad men elicit from better men acts which, in

better circumstances, would also be bad." Williams concludes, "the total consequences of this,

by utilitarian standards, will be worse than if it had never started."[57] These are, of course, not

new insights, but have long been the subjects of the philosophical debate surrounding

utilitarianism.

I've presented these familiar arguments only very quickly, with no consideration of

possible rejoinders. Some utilitarians have of course rejected the charge that utilitarianism is

self-defeating,[58] although the range and variety of the arguments that support it lend it

considerable plausibility, even if some of those arguments can be undermined. However,

defenders of utilitarianism have often replied to such assertions not by disputing their truth, but

by denying that they provide objections to the theory. Thus Derek Parfit, from whom I take the

term "indirectly self-defeating," follows up his verdict that all theories that can be grouped as

---

[57] Williams, *Morality* (Cambridge: Cambridge University Press, 1972), 96.

[58] See, for example, Peter Singer, "Is Act-Utilitarianism Self-Defeating?", *Philosophical Review* 81 (1972): 94-104. My goal here is not to establish beyond doubt that utilitarianism is self-defeating, but rather to suggest that if it is (and there are some grounds for thinking it is), this is a problem for utilitarianism. As I note, some very prominent defenders of utilitarianism have responded to the self-defeatingness charge by arguing it is on target but provides no objection to the theory, and it is this line of response that I am addressing.

consequentialist (including utilitarianism) might plausibly be judged indirectly self-defeating

with the conclusion that "in the case of these theories, to be indirectly self-defeating is not to be

damagingly self-defeating. Nor do these facts provide independent objections to these theories.

These facts do not show that these theories are either false, or indefensible."[59] At most, they

show that if any such theory is true, we ought to form dispositions not to act directly on the

theory. John Stuart Mill, in his reply in *Utilitarianism* to the worry that the principle may prove

too demanding to apply, avails himself of a similar defense: he famously writes, "this is to

mistake the very meaning of a standard of morals and confound the rule of action with the

motive of it."[60] Other defenders of the doctrine have been quick to follow his lead, arguing that

utilitarianism provides a *criterion of rightness* and not a *decision procedure*. Roger Crisp, for

example, argues that a utilitarian agent who visits her sick friend in the hospital needn't think,

when doing so "'What really moves me is reflection on the fact that it maximizes welfare for

people to show special concern for their friends.' She can just make the visit, without thinking

of utilitarianism at all."[61]

Why should this worry a utilitarian? It should do so if, as I've argued, the Coincident

Reasons Thesis identifies the necessary conditions for an action's having moral worth. That

thesis claims, precisely, that in the case of morally worthy actions, decision criteria *are* the same

as criteria of rightness: the reasons why we do them are the same as the reasons why they are

right. We're now considering the probability, acknowledged by many prominent defenders of

utilitarianism, that doing the right thing, *by utilitarian standards*, requires cultivating in

---

[59] Parfit, *Reasons and Persons*, 51.

[60] John Stuart Mill, *Utilitarianism*, ed. by George Sher (Indianapolis, IN: Hackett Publishing Company, 1861/1979), 17.

[61] Crisp, *Mill on Utilitarianism*, 144-145.

ourselves the disposition *not* to be motivated by utilitarian considerations, but instead to learn to see many things and activities and relationships as valuable independently of their ability to promote happiness, and many moral rules as binding even at the expense of total utility. But in that case doing the right thing, *by utilitarian standards*, will require being motivated to act by reasons that are not (noninstrumentally) right making *by utilitarian standards*. That is, it will require acting unworthily. Failing to act rightly is also unworthy. So if the Coincident Reasons Thesis is correct, and if utilitarianism is indeed indirectly self-defeating, then even in a good world – perhaps even the best world accessible to us from our actual world – morally worthy actions will be very rare. This strikes me as a highly implausible conclusion. It seems to me much more plausible that good worlds, in which agents generally act rightly and are motivated to act by their caring for others and enthusiasm for valuable pursuits, are worlds in which there are many morally worthy actions. So the Coincident Reasons Thesis gives us some reason to doubt the utilitarian principle.

A utilitarian might respond that in cultivating in themselves, on utilitarian grounds, the disposition not to be motivated by utilitarian considerations, people act worthily (in response to right-making reasons), and so the actions motivated by the resulting dispositions are also worthy. But motives are not transitive in this way. Imagine, for example, a religious believer, who, from self-interested motives (she fears divine retribution) turns herself into someone who is motivated (noninstrumentally) by right-making reasons.[62] Or consider a violent heavy drinker who breaks his drinking habit because his wife threatens to leave him unless he gives up alcohol, and he fears loneliness. It seems right to say, of both these people, that they are morally better after their reformations than before, despite the self-interested motives driving those reformations.

---

[62] Although it's not clear to me that we can train ourselves to be good in this way.

Though they may not reform for the right reasons, their actions *after* their reformation *can* be

motivated noninstrumentally by the right reasons. Second-order motives are discharged when

the primary motives they target are formed – they needn't linger on to *co*motivate, so to speak,

the actions motivated by those primary motives.

That principle holds true for the case of the well-motivated formation of bad motives as

well. Consider the following example. Let's say I'm a medical student, who has as yet had no

hands-on practical experience, and in some emergency I need to perform a delicate operation on

a patient in order to save her life. The reason I ought to perform the operation is that she will

probably die unless I do it. But if this is also my motivating reason – if, that is, I perform the

operation in order to save her life – the knowledge of what is at stake will cause my hands to

shake so much that I will be likely to botch the operation. In fact, if I operate under these nerve-

wracking circumstances, her chances of survival will be even less than if I had not operated at

all. I ought not, in this case, to operate for this reason. If this is the only reason that could

motivate me to operate, I ought not to operate at all. But it may still be the case that I ought to

operate if there is some way in which I can bring myself not to care about her life. If I can

harden my heart, and think instead only of the accolades that a successful operation would earn

me, and of the bragging I can later indulge in with my classmates, and let a desire for these

motivate me instead, then I could operate successfully.

If I were hard-hearted, cocky, and selfish by nature, and had decided to become a doctor

not out of any (noninstrumental) interest in saving lives, but only because of the status I thought

it would earn me, it would be true of me that I ought operate to save the woman's life, but it

would also be true that my act, motivated on purely selfish grounds, would have no moral worth.

And if I'm not hard-hearted and selfish, but it is somehow in my power to make myself so,

temporarily, then I ought to do that, and I ought to operate – ought to do so for the wrong reasons – and again my act, motivated on purely selfish grounds, would have no moral worth. There is nothing worthy about saving someone's life just for the sake of the accolades, although there may be something worthy, in this case, in my bringing myself to care only about the accolades.

The example illustrates that if the Coincident Reasons Thesis is true, there could be some right acts – acts we'd be obligated to perform – which we could not perform worthily. This follows from the thesis regardless of which normative ethical theory is the right one.[63] But the example I gave is pretty artificial, and there will not be many real-life situations in which we ought to act unworthily. Such cases would be particularly rare according to Kantian ethical views, which generally condemn meddling with anyone's ability to assess and respond to reasons, including one's own, because the stakes would have to be very high for any such act of self-deception and self-corruption to be warranted, especially if it's lasting. It would be no objection to a normative theory if it entailed that we might, in rare circumstances, be obligated to act unworthily. But it is an objection to a normative theory if it entails, as I have argued

---

[63] Indeed, whatever normative moral theory turns out to be correct, there may be a *possible* world in which one cannot act rightly for the right (that is, right-making) reasons *at all* – that is, in which morally worthy action is impossible. Another of Stratton-Lake's example illustrates this:

> Consider a world in which there is an omnipotent, evil demon whose aim is to stop good people from doing what they should in the light of the normative reasons why they should so act…. He achieves this by making it the case that if a good person ever acts from the normative reasons why she should so act, he will make it such that this action is wrong, and he tells them this. Every good person knows, therefore, that she cannot do the right thing from the normative reasons why this is right. For they know that if they are motivated to act in this way, then their actions will be morally wrong. (*Kant, Duty, and Moral Worth*, 18)

What makes the truth of the thesis problematic from the utilitarian point of view is that if utilitarianism is correct, even the best world accessible to us from the actual world is like the world of the example, despite the absence of evil demons. The truth of utilitarianism combined with the truth of the Coincident Reasons Thesis would entail that we should, ideally, make ourselves, as far as is possible, into people whose actions lack moral worth – into people who don't care about or respond motivationally to what really matters.

utilitarianism does, that even in the *best* circumstances people will usually have to act unworthily to act rightly.

It might be responded, on behalf of utilitarianism, that the category of "morally worthy actions" is a Kantian notion to begin with and one that doesn't much interest the utilitarian. What matters from a utilitarian point of view is that we bring about good consequences not that we act virtuously. So, the suggestion is, the utilitarian might be quite happy to abandon the Coincident Reasons Thesis. Of course, many utilitarians will argue that only the consequences (happy or unhappy) resulting from an act and not the intentions with which it was performed are relevant to determining whether it is *right*. Contrary to the claims of some Kantians, they might say, we needn't look at an agent's motivating reasons to see whether she acted rightly. The rightness of her act is independent of her motives.

But I've treated *rightness* as independent of motives throughout the course of my argument. There remains an open question about whether an agent's motives tell us anything else of interest – a question answered by the Coincident Reasons Thesis. And it simply isn't true that the idea of moral worth is merely a manufactured Kantian artifact – on the contrary, it is an important part of our everyday practice of moral judgment. Utilitarians acknowledge this when they take Williams' "one thought too many" objection and related concerns about the morally unattractive character of the ideal utilitarian agent seriously.[64] What's more, the concept of a morally worthy act is closely bound up with other centrally important moral concepts, including

---

[64] According to some forms of utilitarianism, what we ought to do depends on our epistemic position: the right act is the one that maximizes *expected* utility not the one that actually comes out best. These "subjective" forms of utilitarianism seem to me to already to concede the significance of motivating reasons, since what makes them plausible is the thought that someone could be responding as well as possible to the importance of bringing about happy consequences while still, and *blamelessly*, failing to produce the utility-maximizing act.

those of praiseworthiness and blameworthiness, and, of course, that of the good, or virtuous,

person. A utilitarian cannot abandon the notion of moral worth without embracing an error

theory about much of our ordinary moral discourse.[65]

I've argued that the Coincident Reasons Thesis provides a plausible necessary condition

for the moral worth of actions and that it does not declare unworthy any action motivated on

grounds that seem intuitively to be worthy. However, taken in conjunction with utilitarianism, it

does preclude seemingly worthy acts from having moral worth. We should not accept both the

Coincident Reasons Thesis and utilitarianism. So the reasons I have given in support of the

Coincident Reasons Thesis are also reasons to reject utilitarianism.

---

[65] Thanks again to Daniel Markovits for pushing me to contend with this response on behalf of the utilitarian.

A worry may lurk here for the larger strategy behind my argument. That strategy was to defend an account of moral worth on the basis of considerations whose appeal was supposed to transcend disagreements about first-order ethics, and then to show that that account had implausible further implications when adopted along side one first-order ethical theory: utilitarianism. But I began by saying that a morally worthy action is one that *merits praise* in virtue of the positive light it sheds on the moral character of its agent. In explaining the notion of moral worth I had in mind, I emphasized (in n. 4) a distinction between an action that we have instrumental reason to praise because praising it will have good consequences and an action that *merits* praise in its own right and irrespective of any consideration of the consequences of actually praising it. But is this a distinction a utilitarian could recognize? Wouldn't a utilitarian *have* to understand the idea of "meriting praise" solely in terms of the consequences of praising? If so, it seems the notion of moral worth I defend is from the outset one a utilitarian could not accept, and my defense of it no longer appears to be neutral between competing first-order moral theories.

However, I don't think a utilitarian has to understand the idea of an action's "meriting praise" solely in terms of the consequences of praising it. Utilitarians can and surely would acknowledge that the question of whether or not someone is a good person is not settled by determining whether it would be useful to call him good, just as the question of whether or not someone is kind, or clever, or talented is not settled by determining whether it would be useful to think him so. (Indeed, it seems even less plausible to think that the appropriateness of applying such labels to agents depends only on the usefulness of *our applying them* than to think that the appropriateness depends only on the contributions *such agents* make to total utility.) The version of utilitarianism I consider here is a version of act-utilitarianism, and act-utilitarians, of course, understand the *rightness of actions* in terms of the value of their consequences. But they need not and should not understand *any notion of appropriateness* along these lines. It's worth noting that if they did understand appropriateness or merit only in this way, they would be forced into implausibly revisionary accounts not just of moral worth, but also, for example, of epistemic justification (is it really appropriate to believe a proposition only if doing so would be useful?), and of meaning (does a phenomenon really merit a certain description only if describing it that way would be useful?).

I was helpfully pressed on this point by an anonymous reader for the *Philosophical Review*.

5   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 my discussion 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 (noninstrumental)

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 morally justifying them? If we can identify different

noninstrumental 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 morally justifying the action. So the Coincident

Reasons Thesis can, I think, be fairly readily rephrased to accommodate degrees of moral worth:

right actions have moral worth *to the degree that* the noninstrumental motivations for their

performance coincide with noninstrumental moral justifications for their performance.[66]

---

[66] This actually isn't quite right because sometimes the fact that we *ought* to perform an action is
overdetermined by the normative reasons. And it may also be possible for the fact that we're *motivated* to
perform an action to be overdetermined by motivating reasons (a question I won't tackle here). In cases
of overdetermination, there needn't always be perfect overlap between normative and motivated reasons
for an act to be fully morally worthy. In cases of normative overdetermination, it may be okay not to be
motivated by *all* the normative reasons justifying an act, so long as we're motivated by normative reasons
sufficient to justify the act (we may in some cases, perfectly reasonably, stop weighing reasons once
we've determined the right way to act). And if there are cases of motivational overdetermination, it may
be okay to have some non-moral motivations for doing the right thing, so long as we're also fully
motivated by the actual normative reasons justifying the act.

My handling of the topic of moral worth entails that the predicate "worthy" is of the same class as the predicate "full." An action is either worthy or not worthy but nonetheless can be partly worthy; similarly, a glass is either full or not full, but nonetheless can be partly full. How worthy an act is depends on the extent to which the (noninstrumental) reasons motivating it coincide with the (noninstrumental) reasons morally justifying it. If this is right, then degrees of worthiness, like degrees of fullness, are not measurable in units, so-to-speak, but are *inherently dimensionless*. We might say one act is worthier than another, just as we might say one glass is fuller than another, but we shouldn't say this act has so-and-so much moral worth, any more than we'd say a glass has X units of fullness. And, just as one full glass cannot be more or less full than another full glass, one fully worthy action, as I am using the term, cannot be more or less worthy than another fully worthy action: in both cases, there will be perfect overlap between motivating and morally justifying reasons.

This may seem a somewhat restrictive use of the predicate "worthy": it might seem that there is another common use of the predicate according to which even if two acts are each as worthy as they could be, one might nonetheless be worthier than the other. That is, we might think asking how worthy an act is is not like asking how full a glass is but rather like asking how much liquid is in it. If acts could have *units* of moral worth in this sense, then there will be a kind of moral worth (analogous to the volume of liquid in a glass) that actions can have that is not captured by the Coincident Reasons Thesis. Heroic acts might intuitively feel like they fall in this category. Aren't heroic acts worthier than ordinary good deeds that we might be expected

---

Another clarification which has been to some extent implicit in my discussion above: perfect overlap between motivating and justifying reasons requires that a fact providing a reason does the same amount of work in motivating us to ϕ as it does in justifying ϕing. Thanks to Roger Crisp and Joey Fishkin for raising some questions that pushed me to be more precise on these points.

to perform, even if such ordinary good deeds are also performed entirely for right-making

reasons, and thus are, in the sense given by the thesis, completely worthy?

More careful thought about what might make one worthy act worthier than another in this

sense, however, leaves me skeptical about the suggestion that there is a kind of moral worth that

actions can have that can vary from worthy act to worthy act in this way. In virtue of what might

one morally worthy action be worthier than another? A plausible suggestion is that how morally

worthy a morally worthy action can be is a factor of how difficult the circumstances were in

which the agent performed it. We admire heroes in part because we think we would not have

had the strength to act as they did had we been faced with the same decision. That is, perhaps, in

assessing degrees of moral worth, we should not ask (as, for example, Nomy Arpaly does, in her

appeal to counterfactuals that I assess in section 2 above) whether the agent would have

performed the same actions had her circumstances been more difficult, but rather ask whether

other agents would have performed the same action in the circumstances in which our agent was

actually placed.

Does this, then, identify another factor contributing to the *degree* of moral worth of

actions – one not accounted for by the Coincident Reasons Thesis? Does the degree of moral

worth of actions depend not only on the degree of overlap between noninstrumental motivating

and morally justifying reasons, but also how many (or few) of us would have acted similarly well

in similar circumstances? On refection, I don't believe so. Arpaly appeals to the intuition that

we generally think more highly of people who do the right thing in difficult conditions ("die-

hard" or "foul-weather philanthropists") than of people who do good only when the going is

easy. But let's instead compare Arpaly's fair-weather philanthropist with someone who does

good deeds *only* when things are going badly for her (a somewhat rarer breed: the *exclusively*

foul-weather philanthropist). Would we still think more highly of the second agent's actions than of the first? Perhaps the first agent is easily distracted from the needs of others by her own troubles, and is therefore motivated to help others only when things are going well for her, whereas the second is oblivious to the needs of others when her own life is going well, and becomes aware of them, and motivated to help, only when her own troubles draw her attention to life's difficulties. There seems to me to be nothing to choose, in terms of moral worth, between the actions prompted by these two different dispositions. This suggests that while we are inclined to be more likely to judge agents virtuous on the basis of good deeds performed in difficult circumstances than in easy ones, or when doing so seems to require greater energy or sacrifice, this is largely because we infer, often (but not always) correctly, that agents who act well in hard times will do so in easy times as well, and so will tend to be more virtuous overall (because they will act worthily more consistently).[67] Doing good when the going is tough can sometimes be *evidence* of the greater moral virtue of the *agent* but is not constitutive of greater moral worth in the action.

Whether or not the reader is persuaded by this account of degrees of moral worth, the issue does not affect the truth of my central claim that the Coincident Reasons Thesis provides necessary and sufficient conditions for the moral worth of actions. And, even if the degree of moral worth of an action is not determined *solely* by the degree of coincidence between the reasons motivating and morally justifying it, a version of the thesis holds for partially worthy actions as well: an action is partially morally worthy if and only if the noninstrumental reasons motivating it partially coincide with the noninstrumental reasons morally justifying it.

---

[67] A similar point applies to the case of supererogatory actions.

The Coincident Reasons Thesis shows us that an agent can go wrong in acting rightly by failing to be motivated by some consideration that in fact morally justifies her action, and she can go wrong by being motivated by considerations that don't in fact morally justify her action. But, as the thesis demonstrates, 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 noninstrumental reason to perform it); or, (like a self-interested religious believer who gives alms for the sake of a heavenly reward) she can treat as a merely instrumental consideration a moral reason that in fact plays a noninstrumental justificatory role.

The Coincident Reasons Thesis can also help explain when *wrong* actions have (at least some) moral worth. This is because even wrong actions can exhibit partial overlap between motivating and justifying reasons. This most commonly happens when people are motivated to act by genuine normative moral reasons, but get the balance of reasons wrong. Perhaps, motivated by a desire to provide well for their children, they work so hard that they have little time to spend with them. The fact that working the extra hours will enable them to pay, say, for their children to attend a more prestigious university may be a genuine normative reason to do so, but it may be outweighed by other, weightier reasons they have to work fewer hours. In this case, the decision to work long hours is motivated in part by normative moral reasons, but the agent is not motivated, or rather, not motivated enough, by other important normative moral reasons. The agent's motivating reasons partly overlap with her normative moral reasons, and so her action, according to the Coincident Reasons Thesis, has some moral worth. Such a person

might be said to perform the wrong action, but for the right reasons (or rather, *some* of the right reasons).[68]

Conclusion

A plausible account of moral worth should explain how acting worthily is a matter of acting for the right reasons. *At least* those of us persuaded by some central elements of the Kantian approach to ethics should also want such an account to explain the relation between acting worthily and acting rationally (or in response to normative reasons). The account should explain why and how, in the case of worthy actions, the connection between the agent's motivations and the act's rightness is not merely accidental. And it should fit our most firmly held and confident intuitions about examples of morally worthy actions. In particular, it should allow for the possibility of our acting worthily even when (like Huck Finn) we have false beliefs about what morality requires of us. And it should allow for the possibility of our acting worthily even when we're (blamelessly) ignorant of some relevant facts and even if we're mistaken or agnostic about the right fundamental ethical theory – the right account of what makes actions

---

[68] Stephen Kearns pointed out to me the significance of such cases. He also notes a difficulty such cases pose for the way I have described worthy actions at various points in this essay. I can't say, in the case of wrong but partially worthy acts, that they're partly motivated by "right-making" reasons (a formulation I appeal to at some points in this essay), because in the case of wrong acts motivated by genuine normative reasons, those reasons won't, in the circumstances, *be* right-making. Nor can I say of wrong but partly worthy acts (as I have of other morally worthy acts) that the reasons why they are performed and the reasons why they *ought* to be performed are (partly) the same. Because, again, in this case, the reasons motivating the act, though genuinely normative, are *not* reasons why the act *ought* to be performed since it's not the case that the act ought to be performed.

As I hope my discussion in the text above bears out, I think these are merely issues of proper formulation, and do not seriously undermine the thesis, or the account it can provide of the partial moral worth of such actions. They still provide examples of actions in which motivating and normative reasons (partially) coincide. The normative reasons in this case are *pro tanto* and not "winning" reasons. It may not be correct to say that they are reasons the act ought to be performed (though this doesn't sound *too* terrible); but we might certainly say they are normative reasons to perform the act.

right. It should not identify acts performed out of partial concern for family and friends as necessarily unworthy, nor acts performed with the impartial intention to increase the general happiness. And is should give us some insight into why or how some acts, including wrong acts, are partially morally worthy: why they are worthier than they might have been, despite being less worthy than they might be.

I've argued that the Motive of Duty thesis fails to meet many of these benchmarks, and so should not be accepted. I've argued that the Coincident Reasons Thesis meets all of these benchmarks, and thus provides very plausible necessary and sufficient conditions for the moral worth of actions. A Kantian, I've said, should be particularly drawn to the thesis; but most of the reasons for accepting the Coincident Reasons Thesis that I give in this essay do not depend for their force on acceptance of or sympathy with a particular fundamental ethical theory.[69] So the arguments I've given in support of the thesis should be persuasive to people with a range of convictions about theoretical ethics. This is not to say, however, that the thesis sits equally well alongside any ethical theory. As I argued in section 4, combining the thesis with utilitarianism, or with any other ethical theory that is "self-defeating," and so tells people to be motivated by reasons other than those that in fact (according to the theory) morally justify their actions, leads to implausible consequences. So the arguments I've given in support of the Coincident Reasons Thesis also serve as reasons to doubt such theories.

---

[69] Even the force of the Kantian argument for the Coincident Reasons Thesis could, I think, be felt by some non-Kantians as well. The general "Kantian" thought that being good is a matter of being responsive to moral reasons, which the argument is intended to make precise, certainly has a broader appeal, not least because it sits so comfortably with the very plausible idea that the intellectual, or epistemic, worth is a matter of responsiveness to epistemic reasons.

Defenders of utilitarianism or other self-defeating theories might of course respond by rejecting the Coincident Reasons Thesis, despite the considerations I've appealed to in its support. But I am skeptical that they could replace the thesis with an account of worthy action that is as successful in meeting the requirements I've outlined for a plausible theory of moral worth. So rejecting the thesis begins to look a bit like biting the bullet.

References

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

Baron, Marcia. 1995. *Kantian Ethics Almost Without Apology*. Ithaca, NY: Cornell University Press.

Bennett, Jonathan. 1974. The Conscience of Huckleberry Finn. *Philosophy* 187: 123-134.

Broome, John. 2007. Does Rationality Consist in Responding Correctly to Reasons? *Journal of Moral Philosophy* 4: 349–74.

_____. 2007. Wide or Narrow Scope? *Mind* 116: 359-370.

_____. 1999. Normative Requirements. *Ratio* 12: 398-419.

Crisp, Roger. 1997. *Mill on Utilitarianism*. London: Routledge.

Henson, Richard. 1979. What Kant Might Have Said: Moral Worth and the Over-determination of Dutiful Action. *The Philosophical Review* 88: 39-54.

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

Hume, David. 1740/1978. *A Treatise of Human Nature*, edited by L.A. Selby-Bigge and P.H. Nidditch. Oxford: Oxford University Press.

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

Kolodny, Niko. 2005. Why be rational? *Mind* 114: 509-63.

Korsgaard, Christine. 2008. Acting for a Reason. In *The Constitution of Agency*, 207-229. Oxford: Oxford University Press.

_____. 1997. *Creating the Kingdom of Ends*. Cambridge: Cambridge University Press.

_____. 1996. *The Sources of Normativity*. Cambridge: Cambridge University Press.

Mele, Alfred R. 2003. *Motivation and Agency*. Oxford: Oxford University Press.

Mill, John Stuart. 1861/1979. *Utilitarianism*, ed. by George Sher. Indianapolis, IN: Hackett Publishing Company.

Nagel, Thomas. 1970. *The Possibility of Altruism*. Oxford: Oxford University Press.

Parfit, Derek. 1984. *Reasons and Persons*. Oxford: Oxford University Press.

Singer, Peter. 1972. Is Act-Utilitarianism Self-Defeating? *The Philosophical Review* 81: 94-104.

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

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

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

Williams, Bernard. 1981. Persons, Character, Morality. In *Moral Luck*, 1-19. Cambridge: Cambridge University Press.

_____. 1972. *Morality*. Cambridge: Cambridge University Press.