Moral Worth and Moral Knowledge[1]

0. Introduction

Not all right actions are morally praiseworthy. We're hesitant to praise a political candidate who advocates deep cuts to poverty-relief programs for washing pots in a soup kitchen on the campaign trail. There's no question that volunteering in soup kitchens is morally right – perhaps even morally required. But whether an action is morally praiseworthy depends not just on whether it conforms to the correct normative theory (whatever it is). It also needs to be *motivated* in the right way. An account of moral worth aims to identify what such good motivations consist in. My aim here is to develop and defend one particular answer to this question. On this answer, morally worthy actions are those that are motivated by the rightness of the action. That is, they are motivated by an agent's concern for doing what's right and her knowledge that her action is morally right.

> A morally right action has moral worth *if and only if* it is motivated by moral concern (conative requirement) *and* knowledge that it is the right thing to do (knowledge requirement).

Call this the *Rightness Condition*. The Rightness Condition thus makes both conative and cognitive demands on the agent's motivation.

In the first part of the paper, I defend the Rightness Condition as a *necessary* condition for moral worth. In doing so, I argue contra Smith, Arpaly, and others, that moral worth requires a conative attitude with *moral content*. The central idea is that an agent who acts in a morally admirable way is not only morally reliable but also disposed to experience certain reactive attitudes, such as guilt, blame, praise. To be disposed to experience such reactive attitudes, an agent's motivation must have moral content; she must conceptualize her action in moral terms.

In defending the knowledge requirement, I appeal to a central feature of morally worthy actions that has been suggested by Kant:

> In the case of what is to be morally good it is not enough that it [i.e. the action] *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.[2]

Morally worthy actions are motivated in a way that makes their rightness neither "contingent" nor "precarious" – they are counterfactually robust. By appealing to moral knowledge, we can spell out the nature of this counterfactual robustness in a plausible and principled way.

The second part of the paper argues that the Rightness Condition is *sufficient* for moral worth. This has particular implications for the recent debate about moral testimony. Moral testimony can be a source of moral knowledge, and so, on the Rightness Condition, it can give rise to morally worthy actions. I argue that we should welcome this implication.

The third part of the paper argues that the Rightness Condition gives us a nuanced framework for how to think about agents who lack moral knowledge. It accommodates the intuitive thought that there is often much to admire about such agents while remaining clear-sighted about their moral limitations.

1. Some Clarifications

Before we start, two clarifications will be helpful. The first concerns what exactly moral worth is about. I am interested in when it's appropriate to praise an agent in the sense that she is *praiseworthy* for the action. Praising agents is itself something we do and we may do it for all kinds of reasons other than that the action merits

_____

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

praise. If you hold a gun to my head, I'm going to praise your "exquisite" singing even if sense of key is questionable. In this case it may be *prudentially* appropriate that I should praise it. But you are not *praiseworthy* for your performance: you do not deserve credit for your playing. So, the sense of being praiseworthy that I'm interested in is what we mean when we say that an agent "deserves credit for her success". We often contrast the cases in which the agent deserves credit with cases in which it was a matter of chance or luck that the agent succeeded. Even if our actual practice of praising agents is somewhat messy, I believe we have a good enough grip on this notion of praiseworthiness to provide us with a starting point for a philosophical investigation.

Secondly, when I refer to right actions being morally praiseworthy, I use the term 'morally right' broadly to include not just actions that are morally required but also those that may not be required but are nevertheless morally good. Thus, if there are supererogatory actions, they are morally right, even if they are not morally required.

2. Is the Rightness Condition Necessary?

To satisfy the Rightness Condition, an agent must meet both a conative and a knowledge condition: the agent must be motivated by concern for doing what's right and she must know what the right thing to do is. In this section I argue that both of these conditions are necessary for morally worthy actions. Thus, I defend the claim:

> A morally right action has moral worth *only if* the agent is motivated by concern for doing what's right (conative requirement) *and* knowledge of what the right thing to do is (knowledge requirement).

My argument proceeds in two steps: section 2.1 motivates and defends the conative requirement, section 2.2 focusses on the knowledge requirement.

2.1 The Conative Requirement

One central feature of morally worthy actions is that they are not merely accidentally right. To see this, consider some examples of actions whose 'conformity with the moral law' clearly just is a fluke. An example is Kant's shopkeeper who deals honestly with his customers because he wants to keep his business profitable; he doesn't care that it is the right thing to do. Right actions that are motivated by selfish or ulterior desires strike us as unacceptably flukey and rightly so. Insofar as the shopkeeper is motivated solely by a desire for profit, it is a matter of luck that he chose the right course of action. Had it been profitable to cheat his customers, he may have done that. His doing the right thing is a consequence of a fortuitous alignment of what suits his desire for profit and what's morally right.

The Rightness Condition rightly excludes such actions from having moral worth since such actions fail to satisfy the conative condition:

A right action has moral worth only if it was motivated by moral concern.

Moral concern is taken as concern for doing what's right. It is meant to be understood broadly enough to include such conative states as a desire to do what's right, an intention to do what's right, etc. What's important is that the concern for doing what's right be *non-instrumental*: the agent must care about what's right for its own sake, and not because doing so would further some other goal. Thus, the shopkeeper's honest dealings with a customer are not praiseworthy because insofar as he is motivated by a desire to do what's right, this desire is merely instrumental.

It's uncontroversial that there should be some conative requirement for morally worthy actions; most accounts of moral worth subscribe to a version of the conative requirement in order to exclude cases such as the Kantian shopkeeper. As Barbara Herman has argued:

...when we say that an action has moral worth, we mean to indicate (at the very least) that the agent acted dutifully from an interest in the rightness of his action: an interest that therefore makes its being a right action the *nonaccidental* effect of the agent's concern.[3]

The Rightness Condition departs from recently popular proposals by insisting that moral worth requires a conative attitude with *moral content*: it requires that the agent be motivated by, for example, a desire to do the right thing. In contrast, Arpaly argues that while moral worth requires that an agent be motivated by moral concern, moral concern should be understood as concern for those *nonmoral* features of an action which make an action right:

Moral concern is to be understood as concern for what is in fact morally relevant and not as concern for what the agent takes to be morally relevant. [...] To say that a person acts out of moral concern is to say that a person acts out of an intrinsic (noninstrumental) desire to follow (that which in fact is) morality, or a noninstrumental desire to take the course of action that has those features that make actions morally right.[4]

Thus, an agent can be morally praiseworthy when she is motivated by a desire to relieve suffering, feed the hungry, help the needy – whether or not she conceives of those as right-making reasons. Similarly, Markovits argues:

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. My motivating reason for performing some

---

[3] Herman [1981], p. 366.

[4] Arpaly [2003], p. 84

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.[5]

Contra Arpaly, Markovits, Stratton-Lake then, I argue that being motivated by *de re* desires is *not* enough for moral worth. Moral worth requires a conative attitude with moral content.

Consider first an agent who does the right thing and is motivated by an individual *de re* desire for the relevant right-making reason. Jean's friend missed her bus to work and frets over being late to an important meeting; coming late would be a great embarrassment to her. Wanting to spare her friend a major embarrassment, Jean gives her a ride. Let's assume that giving her friend the ride is the right thing to do in these circumstances and the fact that it spares her friend a major embarrassment makes it right. Thus, Jean is acting from a *de re* desire for a right-making reason. Does Jean's action have moral worth?

A central feature of morally worthy actions is that they are not merely accidentally right. Given Jean's motivation, it's not a fluke that Jean spared her friend a major embarrassment. But it is a fluke that she did the right thing. There are plenty of circumstances in which sparing one's friend a major embarrassment is outweighed by other morally relevant factors. In these cases, Jean's motivating *de re* desire would lead her to do the wrong thing. For example, Jean ought not murder her friend's ex-boy friend, even if doing so would eliminate a major source of embarrassment in her friend's life. Hence, if what motivates Jean is solely a *de re* desire for the particular right-making reason, then it's a matter of luck that she acted rightly. And so, her doing the right thing seems to "precarious" for her action to have moral worth.

---

[5] Markovits [2010], p. 205. Similarly, Stratton-Lake [2000] argues:

> we must abandon the idea that a morally good person would be disposed to be motivated to do what she should by the thought of duty, and *a fortiori* the idea that she would be disposed to be motivated solely by this thought. (p. 27)

An isolated *de re* desire will not do. But suppose Jean's psychology is so constituted that whatever the right-making reasons are, Jean will be motivated by *de re* desires for those reasons. Since Jean's *de re* desires track the right-making reasons, it's not an accident that Jean performs the right action; her doing the right thing is counterfactually robust. But suppose that Jean does not conceptualize these right-making reasons *as* right-making reasons. When the right thing to do is to help her friend, Jean is simply moved by the *de re* desire to help her friend. Thus, it seems an agent's right action may well be non-accidentally right – after all, the agent reliably does the right thing – even though she is not motivated by a desire with moral content. Wouldn't this be enough for moral worth?

I don't think it would. Moral worth is about more than just about moral reliability. What matters for whether an agent is morally admirable for doing what's right is not just which other *actions* she would perform under various counterfactual scenarios.[6] It also matters which *reactive attitudes* she is disposed to experience. Suppose that for whatever reason, Jean is prevented from doing what's right and helping her friend. Would she feel guilty and remorseful? Or suppose that Jean witnessed someone else failing to help their friend when it's right to do so. Would she respond with moral indignation? I doubt that we would be inclined to admire Jean for helping her friend, if Jean was not disposed to feel any guilt and remorse, had she failed to do so. I also doubt that we would admire her right action if she was not disposed to react with blame and outrage if someone else refused to help their friend in similar circumstances. But to be disposed to experience such reactive attitudes, an agent must conceive of her action in moral terms: she must care about it *as* the right thing to do.

Why do reactive attitudes such as guilt, moral regret, moral praise and blame require the agent to conceptualize her action in moral terms? Consider what is involved in feeling guilty. To feel guilty is not just a matter of feeling bad

_____

[6] I qualify this requirement in the next section.

or unhappy. Nor is it a matter of feeling this way because one's desires have not been satisfied. I may feel frustrated when I didn't succeed in satisfying my desire to buy avocados because I left it until too late the store had run out. And I may feel unhappy because my desire to help my friend to get to work on time was frustrated by the massive traffic jam that we got stuck in. But this sense of frustration or unhappiness is not guilt. Guilt requires conceiving of one's actions as contrary to what one *should* have done, not merely as contrary to what one *wanted* to do.[7] It involves feeling bad *because* one considers oneself as failing with respect to a moral standard. Feeling guilty thus essentially involves both affective, cognitive, and conative components. This means that to be disposed to feel guilty for failing to perform an action, an agent must conceive of it as something that she morally ought to do and not just as something she wants or feels compelled to do. And she must care about doing it because it's the right thing to do, i.e. under the relevant moral description.

Similarly, to feel moral indignation or to blame another person is not just a matter of feeling angry or upset with someone. Nor is it a matter of feeling frustrated because others fail to act in ways we want them to act. I may feel frustrated when my favorite soccer player slips and bungles her penalty kick. And I may feel upset because the nurse vaccinating my child against tetanus thereby hurts her arm. But this feeling of frustration or upset is not the same as moral indignation or blame. Blame and moral indignation require the recognition that someone has acted as they *should* not have and not merely as we don't *want* them to act. It's a matter of feeling a certain way because one regards the other person as

---

[7] My argument here is that to experience guilt one takes oneself to have violated a standard that one has some desire to comply with. However, one may not endorse the standard and one may not endorse one's own desire to comply with it. As Wallace [1996], p. 39 argues: "It is notoriously the case that one can feel guilt without believing that one has really done anything wrong. We might express this by saying that one can feel guilt without believing that one has violated any demands that one accepts."

having fallen short of a moral standard – as having not just hurt but *wronged* someone. Just like guilt then, moral indignation and blame essentially involve cognitive, affective, and conative components. One cannot be disposed to blame others for breaking their promises unless one conceptualizes breaking promises as the wrong thing to do and one cares about doing what's right.[8]

Thus, there is something defective about an agent's motivation when she acts solely on a *de re* desires, even when they result in her doing the right thing. We can bring this out by considering an agent who is only ever so motivated, who has no desires with moral content in her motivational set at all. Such an agent would be incapable of experiencing any of those reactive attitudes that we regard as central to our practice of holding each other morally responsible. While reliably keeping promises herself, she would not be disposed to respond to evidence that she may have unwittingly broken her word with guilt and regret. She would be disposed to get frustrated when others don't stay true to their word but she would not be disposed to blame them or feel moral indignation. Such an agent would be like a robot who has been programmed to perfectly respond to whatever the circumstances might be. She would strike us as less than fully human. In fact, we may raise doubts whether such an agent would even *be* a moral agent – someone who can be held morally responsible and who is eligible for moral praise and blame. After all, moral agency is closely tied to being the *object* of reactive attitudes.[9] But plausibly to be the object of such attitude presupposes that one participates in the kind of relationships that give rise to such attitudes. Such relationships require reciprocity: blame may only appropriate towards someone who is at the very least capable of feeling guilt and remorse. Thus, there is some

---

[8] I develop this account of reactive attitudes in more detail in my [removed for anonymous review]. See Sher [2006] for an example of an account of blame on which blaming someone requires both beliefs and desires with normative content.

[9] See Strawson [1962] for the original discussion of the importance of reactive attitudes for moral agency. This idea has been developed by a number of authors, see, for example, Wallace [1996].

reason to think that being an object of reactive attitudes may well presupposes that one can be also their subject.

I have said that moral worth requires moral concern: conative attitudes with moral content. But it's important that moral concern should not be equated with "concern for what one believes to be right". Much resistance to the conative requirement rests on conflating the first with the second. Thus, Arpaly objects to Herman, who argues that to have moral worth an agent's motivation must involve "an interest in the rightness of his action" as follows:

> Herman is mistaken if "an interest in the rightness of his action" is interpreted in the most obvious way – that is, as an interest in doing the right thing or the moral thing under this description, in a *de dicto* sense: a concern for doing what one feels or believes, even as a background belief, that one morally ought to do.[10]

Similarly Smith objects to the view that a good person must be acting from a desire with *moral content*, a desire to do what's right on the grounds that "good people" do not just care about "doing what they *believe* to be right."[11]

But a desire to do what's right is *not* a desire to do what one believes is right. Of course, if Jean wants to do what's right, she will tend to act in accordance with her judgments and beliefs about what morality requires. But if it turns out that her judgment led her astray, Jean's desire will, by her own lights, not have been satisfied. Jean will experience some of the reactive attitudes discussed above – regret, guilt, remorse. If, on the other hand, Jean desired to do what she judged to be right, she would receive news of having made a mistaken judgment with perfect equanimity: "Who cares if I was mistaken about what is right. I succeeded in performing the action that I *judged* to be right; this is all I wanted." Such an agent would indeed be troubling. Even if she acted rightly, her actions would hardly

---

[10] Arpaly [2003], p. 73.

[11] Smith [1994], p. 75, highlight mine.

strike us as morally admirable. The Rightness Condition explains why: the conative condition requires an agent to act from a desire to do what's right, not from a desire to do what she judges to be right.

There is more to morally worthy actions than just moral reliability. The agent must be morally reliable *in the right way* – in a way which allows her to experience the range of reactive attitudes that we take to be essential to moral agency. But this, in turn, requires that the agent's motivation have moral content. *De re* desires, no matter how finely tuned, are not enough.

2.2 The Knowledge Requirement

One kind of "accidentality" that's incompatible with moral praiseworthiness is when an agent performs a right action but from an ulterior or selfish desire. But sometimes the conative condition is met and yet the agent's right action seems worryingly accidental. Consider, for example, a bureaucrat like Eichmann, who seems to care about doing what's right but has a deeply misguided conception of what morality requires: he believes that doing what's right requires unquestioning obedience in carrying out the orders of his superior. Suppose that, for once, this superior does order our bureaucrat to do something that's morally right: send some supplies to the needy, for example. Thus, the agent does what's right and he is motivated by concern for doing what's right. Yet, intuitively, he is not morally praiseworthy for his action.

The Rightness Condition rightly excludes such actions from moral worth. This is because it says:

> A right action has moral worth only if the agent knows that it is the right thing to do.

Even if we grant that the bureaucrat genuinely wants to do what's right, he doesn't know that the action he's performing is the morally right thing to do.

His superiors clearly aren't a reliable source of moral guidance. And so, it's a matter of luck that following their orders, the bureaucrat did the right thing.

Together with the conative requirement, the knowledge requirement ensures that there is a counterfactually stable link between an agent's *motivation* and her *right action*. Knowledge is by its nature counterfactually robust: if an agent knows that p, then her belief is safe: she could not have easily believed that p falsely.[12] And so, if an agent is motivated by concern for doing what's right and she knows what the right thing to do is, then it's hardly a fluke that she acts rightly. The Rightness Condition thus makes precise the intuitive thought that praiseworthiness is incompatible with some ways of getting it right by accident.

The knowledge requirement gives us a principled response to a general difficulty for alternative accounts that try to take the counterfactual robustness of morally worthy actions seriously. The challenge is to spell out just how much and what kind of counterfactual robustness right actions must manifest for moral worth. As Markovits rightly notes:

> 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 entails that virtually no acts at all would qualify.[13]

Markovits describes the case of a dog-lover, who risks his life to save a drowning stranger. However, had his dog been present, he would have been too

---

[12] The question how exactly to characterize the counterfactual robustness of knowledge is subject to on-going debate. But there is wide agreement that knowledge implies *some* counterfactual stability. Williamson [2000], in particular, argues for safety as a condition for knowledge. See also Hawley [2003] for a discussion of counterfactual stability of knowledge how.

[13] Markovis [2007], p. 212.

distracted or perhaps unwilling to abandon the dog. If we insist that what matters for moral worth is not just how an agent was in fact motivated but how she would have been motivated in different circumstances, then we must conclude that the dog-lover does not deserve any praise for his heroic deed. [14] But clearly, it is unreasonable to demand that to have moral worth, the agent need to have acted rightly no matter what. *Some* contingency clearly is compatible with moral praiseworthiness. The question is: where to draw the line? Markovits argues that there is no principled line to be drawn and concludes that we should give up on counterfactual robustness as a mark of moral worth altogether.

I think Markovits is right that the views she discusses – she targets in particular Arpaly's and Stratton-Lake's accounts – do not succeed in drawing such a principled distinction.[15] But her drastic conclusion is not warranted.

For one, on the Rightness Condition, morally worthy actions are non-accidentally right *because* they are motivated by moral knowledge and concern for doing what's right. They do not have moral worth *in virtue* of being counterfactually robust in certain ways. The Rightness Condition not only makes this feature of moral worth precise, it also explains it. The Rightness Condition rightly excludes odd and deviant cases, in which the agent couldn't have done but the right thing – for example, in which the agent correctly guessed what the right thing to do is but a scheming demon would have intervened if she hadn't guessed correctly.

_____

[14] Markovits [2010], p. 210

[15] Both accounts take a second-order dispositional motive to be necessary for moral worth. On Apaly's [2003] account it's important that the agent wasn't just actually motivated by the right-making reasons, but she would have been motivated by right-making reasons even if circumstances had been different. According to Stratton-Lake [2000], an agent must not only have been motivated by right-making reasons, but it must be true that she wouldn't have performed the action if she had judged it to be wrong.

Second, the Rightness Condition tells us in which way morally worthy actions are non-accidentally right: the counterfactuals that matter are simply those that come from our best account of knowledge.

On a plausible reading of the case, the Rightness Condition can accommodate the dog lover's admirable deed: insofar as the dog lover is morally competent, he knew that saving the stranger is the right thing to do. (After all, this is hardly moral rocket science.) And insofar as he was motivated by a desire to do what's right, rather than, for example, a desire for glory, then he is morally praiseworthy for his action. Given his motivation, it's not a fluke that he acted rightly. It doesn't matter that he may have been motivated differently had his dog been present.

To see that it's plausible that only actual motivation matters, consider a nonmoral case: a competent surgeon who performs a successful appendectomy. If she wanted to do what's medically right and she knew what that was, then it's hardly a fluke she succeeded in doing what's right. And so she deserves credit for her success. We do not give her any less credit just because had she stayed up all night with her sick child, her motivation would have been different: she may have been to tired to figure out how to deal with a complication or overcome by worries, she would not have cared much about doing what's right for her patient.

The Rightness Condition is less demanding than accounts of moral worth that require that the *agent* would have done the right thing in various counterfactual scenarios: it just requires that, as a matter of fact, she was motivated by moral concern and knowledge of what the right thing to do is. Nevertheless, the knowledge requirement has been criticized for making morally worthy action *too hard* to achieve. In the remainder of this section I hope to dispel these worries. I will first argue that nothing weaker than knowledge will do: to settle for justification, or justification and truth is to give up on the thought that morally praiseworthy actions are non-accidentally right. Second, I will argue that the

Rightness Condition accommodates the intuition that there is often plenty to admire in agents who, through no fault of their own, fall short of moral knowledge.

Consider the following example:

Peter read that spinach has a lot of iron. Based on this information, he concludes that it's good for his child's development and that feeding it to his child is the right thing to do. As it happens, the information is misleading: spinach contains very little iron.

It seems both implausible and ungenerous to deny that there is something morally admirable about Peter's giving his child spinach. But Peter does not know that this is the right thing to do. And so, it seems that moral knowledge cannot be a necessary requirement for moral worth.

One strategy to accommodate such cases of blameless moral ignorance is to appeal to the observation that while Peter may be mistaken about the iron content in spinach, he still seems to be acting on *good reasons*. After all, he is *justified* in believing that giving his children spinach is the right thing to do. Thus, Markovits argues:

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 [...].[16]

Peter does have good reasons for feeding his child spinach; he has expert testimony that it contains a lot of iron. Thus, we might argue that moral worth does not require knowing what the right thing to do is, it just requires having adequate grounds for taking something to be the right thing to do. Rather than being a matter of knowledge, it's a matter of *justification*. And whether one's

---

[16] Markovits [2010], p. 220

actions are justified is determined by the agent's epistemic situation – not by what it is in fact the right thing to do.

But *mere* justification cannot be enough for morally worthy actions. This is because an agent can be justified in believing herself to act rightly, when she is in fact acting *wrongly*. And in such cases, the agent is not morally praiseworthy for her actions. Imagine Ann who grows up in a very remote and tight-knit community. Her friends and neighbors are honest and kind people who help each other out and are welcoming to strangers, rare as they may be. Unfortunately, there is consensus in Ann's community that gay marriage is a great moral evil. Ann has never met a gay person and her belief is not based on hatred or dislike of gay people. Rather, she has acquired her belief that same-sex marriage is a moral evil along her other moral beliefs, such as that one should keep one's promises, that one should be kind to others, that one should not be cruel to animals. It strikes her as intuitively obvious that same-sex marriage is 'different' from 'regular' marriage and that this difference is morally relevant. Since her community is so remote, she has had the chance to encounter someone who would challenge her view on same-sex marriage. Were she to encounter such a person, it would prompt her to critically scrutinize her convictions, rather than holding on to them dogmatically.

It seems plausible that Ann could be blameless for her moral ignorance. Her moral ignorance is the result of epistemic bad luck: her sheltered upbringing. She has no reason to distrust the community consensus or her own intuitions on this particular issue and those around her are morally reliable on other moral questions. And so, it seems, Ann's belief that same-sex marriage is a moral evil, while misguided and unfortunate, may well be justified.

Now suppose that Ann signs a petition against same-sex marriage. By her own lights this is the right thing to do. And she is doing so for 'good' reasons – after all, she is justified in taking same-sex marriage to be morally impermissible. Thus, if justification suffices for moral worth, it seems that Ann is morally

praiseworthy for signing a petition against same-sex marriage. But this cannot be right. Signing a petition against same-sex marriage is morally wrong. Insofar as Ann's belief that she's doing the right thing is sincerely held and justified, it may be plausible that we should not blame her for her moral mistake. Surely, however, we should not give her *moral credit* for following her misguided conscience. The Rightness Condition offers a straightforward explanation for why Ann fails to be morally praiseworthy for signing a petition against same sex marriage. Signing the petition is morally wrong, and so Ann cannot know that it's the right thing to do. Insofar as her belief that same-sex marriage is morally impermissible is epistemically justified, she may well not be *blameworthy* for signing the petition. Doing the best you can, given your epistemic situation, may well excuse you from blame. But it's not enough for earn you moral praise when, despite your best effort, you get it wrong.

What if we require justification and truth? Unlike Ann, Peter's belief that spinach is nutritious is both justified *and* true. But if it's a central feature of morally worthy actions that they are not just right but non-accidentally so, justified true belief in the absence of knowledge will not do. Insofar as an agent's justified true belief that an action is right falls short of knowledge, her doing the right thing looks worrisomely accidental.

This is best illustrated with nonmoral cases. Consider a doctor, who based on the her patient's lab results diagnoses her with iron deficiency and prescribes a nutritional supplement, recommending a common brand of it to the patient. Unbeknownst to her doctor, the lab technician made a mistake. The patient is doing fine on her iron level but she suffers from a vitamin B12 deficiency. Fortunately, the particular brand of nutritional supplement includes both iron and vitamin B12. The doctor's belief that the particular supplement will help the patient is thus true. It is also justified – after all, the lab results provide the doctor with good reason for believing that the patient suffers from an iron deficiency. But

does the doctor deserve credit for *prescribing the correct nutritional supplement*? Intuitively, not. While she did get the treatment right, her success was an accident. This is because the doctor did not know that the patient had a B12 deficiency and hence did not know that the nutritional supplement was the correct treatment.[17]

A second kind of example that raises doubts that justified true belief can suffice for praiseworthiness involves lottery cases. Take, Sam who bought a ticket to a lottery, having good reason to believe that this lottery is rigged so as to make his winning inevitable. Unbeknownst to Sam, the relevant mechanism malfunctions on this particular occasion. Nevertheless, through a very unlikely turn of fortune, Sam's ticket is drawn! Does Sam deserve credit for winning the lottery? Intuitively, not. While it would be appropriate to congratulate Sam, it would not be appropriate to praise him. This is because Sam did not know that his was the winning ticket and he did not know how to win the lottery. Given that he did not know, his success was a lucky accident.

In nonmoral cases justified true belief does not seem to be enough to deserve credit for one's success. This is because without knowledge an agent's success seems worryingly accidental. Does this carry over to moral cases? It would be strange if things were different for cases involving moral beliefs. After all, our practice of attributing moral worth to particular actions is just a special case of our general practice of praising agents for success. We praise the detective for catching a murderer, the tennis player for nailing her shot, the surgeon for making the correct diagnosis. Moral praise is just a special case of this general practice. When we attribute moral worth to an action, we are praising agents for moral success: we are praising them for having done the morally right thing. And so, it would be

---

[17] My examples and discussion here follow Gibbons [2001]. Gibbons uses them to argue for a different conclusion: that knowledge is necessary for intentional action. I'm not committed to this conclusion here; my argument only appeals to the claim that, without knowledge, the action's success is accidental in a way that's incompatible with praiseworthiness. See also Hawthorne and Stanley [2008].

very odd if our practice of praising agents for nonmoral success required knowledge but our practice of moral praise did not.

Thus there are two reasons for why mere justification, in the absence of moral knowledge, cannot suffice for moral praise. On the one hand, such a view yields implausible verdicts for when agents act wrongly on a false but justified moral belief. On the other hand, justified true belief in the absence of knowledge is hard to square with a central feature of morally worthy actions: namely that they are not merely right but non-accidentally so. Insofar as a justified true belief fails to be knowledge this is generally precisely because the agent's epistemic circumstances are precarious. And so, the agent's motivation could have easily led her astray. Let us now return to the charge that it is ungenerous to refuse an agent moral credit when she did the best she could, didn't act wrongly, and yet fell short of knowledge. What do we do about Peter?

The Rightness Condition does rule out that Peter is morally praiseworthy *for feeding his child spinach*. He does not know that this is the right thing to do; it's a matter of luck that he got it right. But this in and of itself does not strike me as ungenerous – after all, it's just denying moral praise for *one particular action*. It would be ungenerous to say that there is *nothing* that Peter is morally praiseworthy for. But the Rightness Condition is compatible with Peter's being morally praiseworthy for many other actions that he also performed: he may, for example, be morally praiseworthy for seeking out advice about his children's nutrition. This is the right thing to do and plausibly Peter knows that and cares about doing what's right. In the same way, while the doctor does not deserve credit for prescribing the correct medication, this is not to deny that she deserves credit for a great many other things that she has done: for assiduously reviewing the lab reports and carefully examining the patient.

These considerations highlight that when considering whether an agent deserves moral praise for an action, we must take care in identifying exactly which

action we take them to be morally praiseworthy for. Individuating actions can sometimes be a tricky business. Thus, consider the following case:

> Susan knows that a child is drowning in either the right pond or the left pond. But she doesn't know in which one – it's dark and foggy. She only has time to jump into one pond and if she doesn't do anything the child will die. Susan simply picks the left pond and jumps in. She picked correctly and manages to save the child.

Saving the child involved Susan's jumping into the *left pond*. Susan didn't know that jumping into the left pond was the right thing to do – she didn't have any evidence for thinking the child was more likely to be left rather than right. So, on the Rightness Condition Susan is not morally praiseworthy for jumping into the left pond. This does not strike me as counterintuitive: Susan was lucky that she got it right. But Susan's saving the child involved many other actions of which she plausibly did know that they were the right thing to do. Prior to jumping into one of the ponds, she knew that she was morally required to do *something* – doing nothing would have meant the sure death of the child. And once she had jumped in and found the child, she knew that the right thing to do was to get the child out of the water. These actions were necessary for saving the child. And she knew that they were the right thing to do. On the Rightness Condition she is morally praiseworthy for them. Susan's saving the child picks out a cluster of actions; and most of those (albeit not all) on the Rightness Condition, do have moral worth.

The aim of this section was to argue that the Rightness Condition provides us with necessary conditions for moral worth. For an agent to deserve moral credit for her right action, she needs to be motivated both by moral concern – a conative attitude with moral content – *and* by moral knowledge. The conative requirement explains why right actions performed from ulterior motives do not have moral worth. It also explains why when an agent performs a morally worthy action, she is not only disposed to do the right thing reliably, but also to manifest a range of

reactive attitudes, including praise, blame, guilt, indignation. The knowledge requirement guarantees that morally worthy actions are not just accidentally right. On the Rightness Condition then, morally worthy actions are counterfactually robust along two independent dimensions: the agent's desire to do what's right means that she would not have performed the action if she had believed it to be wrong. The agent's moral knowledge means that she would not have been easily mistaken about what the right thing to do is. And so, there is a counterfactually robust link between the agent's motivation and her doing the right thing.

3. Is the Rightness Condition Sufficient?

I have argued that moral concern and moral knowledge are both *necessary* for moral worth. Does the Rightness Condition also give us a sufficient condition? While worries about whether the Rightness Condition is necessary focus on the question whether it's too demanding, concerns about whether it's sufficient have focussed on the question whether it's demanding enough. Thus, Hills has recently argued that the Rightness Condition makes moral worth too easy to come by. Consider the following case:

> Ron is an extremist, believing that killing a person is not generally immoral but that killing a fellow Jew is a grave sin. Ron would like to kill Tamara, but he refrains from doing so because he wants to do the right thing, and he knows (on the basis of his rabbi's testimony) that the right thing to do is to refrain from killing her.[18]

Hills argues that, intuitively, Ron is *not* morally praiseworthy for resisting his desire to kill Tamara. But, she argues, Ron does cares about doing what's right

---

18 Hills [2009], p. 115. The case appears originally in Arpaly [2003], who initially uses it to defend a weaker claim: that justified true moral belief along with a desire to do is right is not sufficient for moral worth. Arpaly's later discussion (particularly on p. 73) suggests that she endorses the stronger conclusion also: that even desire to do what's right along with moral knowledge are not sufficient for moral worth.

and based on the rabbi's testimony he *knows* what the right thing to do is. And so, the Rightness Condition cannot give us a sufficient condition for moral worth. The culprit, Hills argues, is the knowledge requirement. Ron may have moral knowledge but he fails to have *moral understanding* and it's the latter that matters for moral worth. To have moral understanding, an agent must have the abilities to give and follow explanations why her action is right and to recognize what the right thing to do is in similar circumstances. Moral understanding is hence considerably more demanding than knowing what the right thing to do is.[19]

I agree with Hills that Ron does not deserve moral credit for refraining to kill Tamara. But I do not think that this intuitive verdict gives us reason to abandon the Rightness Condition. The case that Hills rests her case on is very unusual. Why does the question whether he may kill Tamara even arise for Ron? Why does he think that killing infidels is morally permissible? Circumstances in which a minimally decent moral agent might genuinely wonder whether it's morally permissible to kill another person are rare.

Here's a natural way to fill in the story: Having grown up in a fundamentalist society, Ron has never encountered anyone who would seriously question that all infidels that must be exterminated. Ron is not a psychopath; he is simply not particularly reflective. With this background, Ron's belief that it's morally permissible to kill infidels may not be unreasonable. If he has been steeped in propaganda enough, we can see how he might think it to be his duty to kill any infidel he comes across. I believe that when we are invited to consult our intuitions about this case, this is how we are inclined to fill in the details. And then, Hills' assessment of the case seems right: intuitively, Ron isn't praiseworthy for refraining to kill Tamara. But on this way of filling in the details Ron cannot *know* that it would be wrong to kill Tamara. To come to know what the right thing to do is based on testimony, Ron must be in a position to identify a reliable

---

[19] Hills also argues that it's distinct from simply knowing why something is the right thing to do. See her [2009], pp. 103-106.

advisor. But if Ron is so morally incompetent that he needs help to decide whether he may kill another human, and if he has such misguided moral beliefs that he believes there is nothing wrong with killing an infidel, how could he possibly be in a position to identify an advisor who can be a reliable source of moral guidance?

Thus, insofar as Hills is *stipulating* that Ron acquires moral knowledge based on testimony despite his lack of basic moral competence, she is describing a case that is subtly incoherent. And if she is not stipulating that Ron, then the case is so unusual that it's too under-described to gauge our intuitions. Even if there is a way to fill in the story so that it's plausible that Ron is (1) believes that there is nothing morally objectionable with killing a fellow human and yet (2) morally competent enough to identify a reliable moral advisor, it's not at all clear that with the relevant details filled in, it will still yield the intuitive judgment that Ron is not morally praiseworthy for his action.

As described Ron's case doesn't impugn the Rightness Condition. On the contrary, it seems quite natural to appeal to Ron's lack of moral knowledge to explain why he isn't morally praiseworthy for doing the right thing. Given Ron's moral incompetence, he was lucky to come across a reliable advisor and to do the right thing. Ron's case then lends further support to the knowledge requirement as *necessary* for moral worth.

Ron's action lacks moral worth, not because he acts on moral testimony but because he fails to have moral knowledge. But sometimes agents do gain knowledge of what the right thing do is by relying on others. In such cases, the Rightness Condition implies that their right actions can have moral worth. This is a welcome consequence. Consider the following case:

Advice: Anna's older sister is struggling with alcohol addiction; she lost her job, blew through her savings and is several months behind on rent. She asks Anna to "loan" her some money. Anna is conflicted. On the one hand, she does not

want her sister to end up homeless. But she wonders whether her sister needs to feel the full consequences of her addiction to finally seek treatment. Moreover, Anna's financial circumstances are modest and she has her own family to look after. Anna is uncertain what the right thing to do is. She turns to a friend whom she knows to be trustworthy and to have good judgment for advice. Her friend tells Anna that she shouldn't give her sister the money. Anna's friend is right: the moral considerations against giving her sister the money do outweigh those in favor of it. Anna trusts her friend and acts on her advice. Although it's hard on her, she stays firm and resists her sister's rage and pleas for help.

Intuitively, Anna is morally praiseworthy for doing the right thing and refusing to give her sister money. Anna's case is very different from Ron's. She is morally uncertain because she is faced with a genuinely difficult moral situation and she has good reason to worry about being biased. Even if she is aware of what all the relevant moral considerations are, she may well be uncertain about how to balance them. In difficult situation like this, her seeking out advice is a sign of moral competence: it shows that she is aware of her own moral limitations. Unlike Ron, Anna is well-positioned to identify a reliable moral advisor. And so, it's plausible that Anna by relying on her testimony *comes to know* that refusing to give money to her sister is the right thing to do.[20] Since she knows what the right thing to do is and wants to do what's right, it's hardly accidental that she acts rightly.

While the Rightness Condition allows that some actions on moral testimony have moral worth, it doesn't follow that *all* such actions do. Hills' objection to the Rightness Condition rests on the assumption that moral

---

[20] This assumes that we can acquire moral knowledge on the basis of moral testimony. While a number of people have argued that moral testimony is problematic (see Hopkins [2007], McGrath [2011]), most don't think that failure to transmit knowledge is the problem. An exception to this is McGrath [2010]

knowledge, and *testimonial* moral knowledge in particular, is cheap. But this assumption is false. Generally, to be in a position to know that p, an agent must be epistemically competent, i.e. reliable, responsive to evidence, sober, non-delusional,...Testimonial knowledge isn't exempt from them.

While the Rightness Condition accommodates Anna' s action as having moral worth, such cases create problems for alternative accounts of moral worth, on which being motivated by right-making reasons *de re* is *necessary* for moral worth, as defended by Arpaly and Markovits.[21] The problem is that there isn't an account of right-making reasons which will yield the right verdict in cases like Anna's without committing one to unpalatable consequences in other cases.

According to Arpaly, right-making reasons are those features of an action that *explain* why it is the right thing to do and an agent needs to be motivated by concern for these features.[22] But reliable moral testimony doesn't explain why an action is right. What explains why Anna should not give her sister money are the non-moral facts about Anna's situation – the fact that doing so would make Anna's sister worse-off in the long run, for example. Moral testimony isn't a right-making reason; it's evidence about what right-making reasons there are and what they on balance support. Thus, when Anna does the right thing because she wants to do what's right and knows, based on her friend's testimony what that is, she is not acting on right-making reasons *de re*. She is acting from a desire with moral content along with moral knowledge: knowledge what's right and what the right-making reasons are. And so, on Arpaly's view, her action lacks moral worth.[23]

This is an implausible result. Given Anna's doubts about her own moral judgment, it is a good thing that she relied on moral testimony. More than that: it

---

[21] Markovits [2010] argues that her Coincident Reasons Thesis gives us both necessary and sufficient conditions for moral worth. Arpaly [2003] offers "Praiseworthiness as Response to Moral Reasons" (p. 84) as a necessary and sufficient condition for moral worth.

[22] Arpaly [2003], p. 84.

[23] See her discussion in Arpaly [2003], p. 73.

seems that she acted virtuously. After all, it's morally admirable to know one's own moral limitations and to take them into account when deciding how to act. We admire the surgeon who consults her textbook or her colleague when she reaches the end of her wisdom rather than trying to muddle through on her own, endangering her patient. Similarly, we should regard the moral agent who knows when to look for advice as morally admirable.

Markovits seeks to accommodate cases like Anna by adopting a more liberal account of right-making reasons. Markovits suggests that a feature of a situation is right-making reason if it provides *evidence* for an action's rightness.[24] While this account of right-making reasons allows Markovits to accommodate actions like Anna's, it has troubling consequences elsewhere. We can imagine cases in which the fact that an injection is more painful than the alternatives may be evidence for its effectiveness. Hence the fact that injecting you with the drug will cause you more pain can be a right-making reason. The sadistic doctor who prescribes the painful injection because it's painful and because she non-instrumentally cares about causing you pain is then, on Markovits' account, motivated by concern for a right-making reason. Thus, according to Markovits, her action has moral worth. This strikes me as an unacceptable consequence.

The Rightness Condition avoids these unpalatable consequences. At the same time it does not deny that being responsive to right-making reasons is very important for performing morally praiseworthy actions. On the contrary: morally relevant features – features that either explain the rightness of an action or that reliably correlate with an action's being right – comprise one's moral evidence. Being responsive to our moral evidence is crucial for moral knowledge.

---

[24] Markovits [2010], see particularly her discussion on p. 219. Markovits' account of right-making reasons is close to the one defended in Kearns & Star [2009].

I have argued that the Rightness Condition gives us a sufficient condition for morally worthy actions. This is because if the agent is motivated by concern for doing what's right and knowledge of what morality requires, then there is a counterfactually stable link between the agent's right action and her motivation. The Rightness Condition gives us a principled account of this counterfactual link. It also allows that agents in many different kinds of epistemic situations can act in morally worthy way; moral knowledge can come from many sources – deliberation, testimony, or first-hand evidence.

4. Being Good and Acting Well

We can evaluate whether actions have moral worth. But we can also morally evaluate agents: we can evaluate the extent to which someone is a morally good person. While accounts of moral worth tell us how to do the former, they plausibly have implications for the latter. A central objection to the Rightness Condition, the much-discussed case of Huckleberry Finn – argues that these implications are implausible. The aim of this section is to respond to this objection and, in doing so, to outline the relationship between moral worth of actions and moral goodness of agents.

Let's start with Huckleberry Finn. Huckleberry Finn escapes his abusive father. He meets Jim, a fugitive slave and together they embark on a trip down the Mississippi river on a raft. They make it through quite a few adventures together. Then, suddenly, it occurs to Huckleberry that helping a fugitive slave is like 'stealing'. In the grip of his ill-trained conscience, Huckleberry resolves to turn Jim over to the authorities. But when the crucial moment comes, Huckleberry finds that he cannot go through with his resolution. Instead, he makes up an elaborate story that protects Jim. In doing so, he clearly does the right thing.

However, he believes that he is acting wrongly: he continues to believe that he is complicit in 'stealing property' from Jim's 'rightful owner', Miss Watson.[25]

Arpaly's argues that Huckleberry presents the Rightness Condition with the following challenge: if the Rightness Condition is correct, then moral ignorance is incompatible with morally worthy actions. If we take moral ignorance to be incompatible with morally worthy action, then we have to regard it as incompatible with being a morally good person. And if we regard moral ignorance as incompatible with being a good person, then Huckleberry Finn turns out to be a bad person. But this is implausible:

> Huckleberry Finn [...] is not a bad boy who has accidentally done
> something good, but a good boy."[26]

Arpaly argues that many people hold profoundly misguided moral views and yet strike us as fundamentally decent and good:

> Huckleberry Finn is not an isolated case. [...] it is not dissimilar to a
> considerably more common type of behavioral inconsistency – the person
> whose explicit views with regard to morality and politics are terribly wrong
> but who in everyday life "cannot hurt a fly." We all have friends, family
> members or acquaintances of this sort. We can all recall the likes of a
> student who, waving his copy of Atlas Shrugged in one's face, preaches that
> one should be selfish and then proceeds to lose sleep generously helping his

---

[25] The episode appears in Twain [1884], Chapter XVI. In the context of the moral worth debate, the case is discussed first in Bennett [1974].

[26] Arpaly [2003], p. 78. Driver [2000], p. xvi, objects to regarding moral knowledge as central to virtue on similar grounds:

The Huckleberry Finn case...illustrates a person who has demonstrably false beliefs about the good. But if virtue theorists insist that Huckleberry lacks virtue because of his flawed cognitive state, then this is bad news for most of us, who, even in some small way, are likely to harbor false views of value. The psychological requirements placed on virtue in the classical tradition seem far too rigid and unrealistic.

peers. If philosophers were right in believing [...] that only actions derived by deliberation from one's moral principles are done for moral reasons, we would have to view these people as bad people who happen to have some fortunate inclinations in their makeup. More commonly, however, we treat these people as fundamentally good people who happen to be incompetent abstract thinkers.[27]

The Rightness Condition does not commit us to seeing Huckleberry as a "bad boy who has accidentally done something good". It does imply that moral knowledge is a central part of being a good person. And it also implies that Huckleberry's action of protecting Jim in the crucial moment does not have moral worth. But whether someone is fundamentally good does not hinge on whether a particular action of theirs has moral worth. Fundamentally good people sometimes fail to act in morally admirable ways and fundamentally bad people may occasionally perform morally admirable actions. The Rightness Condition suggests that Huckleberry is a "fundamentally good" but imperfect boy who, on this particular occasion, was lucky to act rightly. And this, I will argue, is very plausible.

What does the Rightness Condition say about evaluating *agents*? The following seems like a truism: morally good people generally act in morally admirable ways. If acting in morally admirable ways requires moral concern and moral knowledge, we can draw the conclusion that morally good people are generally motivated by moral concern and knowledge of what's right. Thus, the Rightness Condition suggests that a morally good cares deeply about doing what's right and is morally competent.

There is an important difference in what matters for the moral evaluation of an action and what matters for the moral evaluation of an agent. Whether a particular action has moral worth depends only on whether *it* was motivated in

---

[27] Arpaly [2003], p. 78.

the right way: by moral concern and knowledge that it's right. In contrast, whether and the extent to which an agent is morally good depends on the *pattern* of her motivation. *Being good* is something that admits of degrees. Agents can be more or less morally competent; they can know what's right in some situations while being prone to moral mistakes in others. They can more or less deeply about doing what's right. How virtuous an agent depends on *which* morally praiseworthy actions they are in a position to perform. A morally good person is in a position to know what's right even in tricky situations, in which those less virtuous are at a loss. (This is why morally good people make good moral advisors.) Second, the depth of moral concern enable her to perform right actions that would be difficult for others: such as risking one's life to save a stranger.[28] And so the Rightness Condition suggests:

> An agent is morally good to the extent to which she both cares about doing
> what's right and she is morally competent, i.e. she generally knows what the
> right thing to do is.

Let us first see what the Rightness Condition says about Huckleberry's *action*. According to the Rightness Condition, Huckleberry is morally praiseworthy for helping a fugitive slave if and only if he is motivated by his concern for doing what's right and his knowledge that helping a fugitive slave is the right thing to do. But, plausibly, Huckleberry does not know that helping a fugitive slave is the right thing to do: he believes that he is complicit in moral wrong-doing. Huckleberry's belief, while unfortunate, is not unreasonable; it is a widely accepted view held even by those who are morally reliable on many

---

[28] The Rightness Condition thus departs from the "battle citation" model of moral worth, discussed by Henson [1979], on which attributing moral worth to actions "acknowledges a moral victory against odds" (p. 50). The odds need not have been unfavorable for an agent to deserve moral credit for her right action. But of course there may be good pedagogical reasons for *giving* the agent more praise in such circumstances.

questions, other than the permissibility of slavery. Huckleberry is thus not morally praiseworthy for helping a fugitive slave.

What about the evaluation of Huckleberry as an *agent*? Whether Huckleberry Finn is a "good boy" does not hinge on whether a particular action of his has moral worth. It depends on the Huckleberry the overall pattern of his motivations and on whether he is generally morally competent and cares about doing what's right. Huckleberry may have false beliefs about the permissibility of slavery. But we regard Huckleberry as "fundamentally good" precisely because Mark Twain's story gives us plenty of evidence that, for the most part, he *does* know right from wrong: he generally knows to treat others fairly, he knows to keep his promises, he knows when to be loyal. Huckleberry also cares about doing what's right: after all, he *apologizes* after treating Jim badly. But a genuine apology requires both moral insight and moral concern – it involves recognizing that one has acted as one should not have and that for this reason one ought to make amends.[29]

The point that agents can be morally competent in some, perhaps most, situations while being systematically morally ignorant in others, generalizes. The fact that Arpaly's libertarian student is mistaken about what the right thing to do is when discussing politics or giving to charity is compatible with his knowing what the right thing to do is when confronted with a friend in need.[30] This phenomenon is not limited to moral cases: I know where my friend lives when I'm driving there but not when I have to explain the way to someone else. I may have false beliefs about gravitational force when solving physics equations but not when

_____

[29] I'm puzzled by Arpaly's take on this; she suggests that Huckleberry "finds himself" apologizing to Jim without having any moral insight or beliefs. See her [2003], p. 77. A sincere apology expresses remorse; hence it requires the insight that one has treated someone in a way that one should not have. See Martin [2010].

[30] This relates to my discussion in Section 2.1: if he was not disposed to feel guilt upon letting his friend down, I doubt that his helping his friend would strike us as admirable.

I'm standing on the baseball pitch, catching the ball. Recently, philosophers of mind and epistemologists have suggested that we should describe such agents as having divided or "fragmented" cognitive states. While their proposals differ in important details, the general strategy is to argue that all-out belief is too crude a tool to adequately describe the cognitive state of such agents; instead we should ascribe to them different beliefs (or aspects of belief) in different contexts.[31] The libertarian student may well believe that we should be selfish – and hence be morally ignorant – in the context of classroom discussions. It does not follow that he believes that we should be selfish when dealing with family and friends and others in need. In these situations he may well know that he is required to help and his actions may have moral worth. And if his moral ignorance is confined to a narrow class of situations, then he may well be a fundamentally admirable person.

The Rightness Condition is compatible with Huckleberry's being a fundamentally good boy. Still, it commits us to regard even his more limited moral ignorance as a moral flaw. This is exactly as it should be. Moral ignorance is a moral flaw because it makes us systematically vulnerable to moral mistakes: it leads to moral blind spots. While it may be that the libertarian "cannot hurt a fly" in everyday life, he is prone to reach disastrously wrong conclusions when deciding which political party to vote for.[32] And it was Huckleberry's moral ignorance – his racist beliefs – which led to his morally misguided decision to betray Jim. We

---

[31] For different ways of developing this strategy, see Schwitzgebel [2001] and [2002], Gendler [2008], and [redacted for anonymous review] [ms].

[32] This fits well a growing body of empirical research on this topic, which suggests that the ordinary cases of sexism and racism that we encounter among "fundamentally decent" people are not just a matter of poor abstract reasoning. These mistaken moral views systematically manifest themselves in a plethora of biases in their everyday moral decision-making. Even those who explicitly endorse liberal views but grew up in societies that are structured by race or gender are not immune from such biases. But agents who in addition hold mistaken moral views are in a particularly bad epistemic position: they are precluded from recognizing their biases as such and thus take compensatory measures. For an overview, see Banaji & Greenwald [2013].

cannot blame Huckleberry for having acquired these attitudes. Nevertheless, they represent a moral flaw; Huckleberry just as the libertarian student would be morally better people if they lacked their false moral convictions.

I want to end by returning to Huckleberry's action, when rather than turning Jim in he protects him instead. I have said that the Rightness Condition implies that Huckleberry is not morally praiseworthy for protecting Jim. But this is not to say that there is nothing to admire about Huckleberry's action. Let's look at the crucial moment that prompts Huckleberry to abandon his misguided plan occurs just as Huckleberry is departing from the raft that he and Jim share in order to tell on Jim. As he sets out, Jim calls after him:

> Pooty soon I'll be a shout'n for joy, en I'll say, it's all on account o' Huck; I's a free man, en I couldn't ever ben free ef it hadn't been for Huck; Huck done it. Jim won't ever forget you, Huck; you's de bes' fren' Jim's ever had; en you's de only fren' ole Jim's got now.[33]

When Huckleberry hears Jim's words, his resolve starts melting away:

> I was paddling off, all in a sweat to tell on him; but when he says this, it seemed to kind of take the tuck all out of me.[34]

Arpaly and Markovits both take Huckleberry to be responding to Jim's personhood. But that seems like a very sophisticated concern to attribute to someone like Huckleberry. Rather it seems that Huckleberry responded to Jim's reminder of their *friendship*. It's not surprising that this reminder should stop Huckleberry in his tracks. Huckleberry knows that loyalty is a central demand of friendship. If there is one norm of friendship, it's that friends stick together, don't tell on each other, and help each other in need.[35] Plausibly, Huckleberry knows what being a good friend requires and he cares about Jim's friendship. Insofar as

---

[33] Twain, p. 103.

[34] ibid.

[35] In fact, at a number of points in the novel Huckleberry explicitly reflects on these demands. My discussion here follows Manne [forthcoming].

this is what motivates him, it's not an accident that he *protects his friend Jim* and he is praiseworthy for *being a good friend*. Being a good friend is something that we value in others for its own sake. And so, it's not surprising, that Huckleberry's action strikes us as admirable.

At the same time Huckleberry is a morally lucky boy: Jim protects him as much as he protects Jim. By reminding him of their friendship, Jim prevents Huckleberry from committing a grave moral mistake.

5. Conclusion

I have defended the Rightness Condition as an account of moral worth, arguing that morally worthy actions require that the agent be motivated by both concern for doing what's right as well as knowledge of what the right thing to do is. The Rightness Condition both accommodates and explains a central feature of morally worthy actions: that their rightness is not just "contingent" and "precarious". It recognizes there is more to acting admirably than mere reliability. Moral motivation is also about being disposed to manifest a range of reactive attitudes – guilt, resentment, admiration. Finally, it makes moral worth neither too easy to come by nor too hard to attain for imperfect and limited agents like us.

6. Bibliography

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

Baron, Marcia (1984). The alleged moral repugnance of acting from duty. *Journal of Philosophy* 81 (4):197-220.

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

Dreier, James (2000). Dispositions and fetishes: Externalist models of moral motivation. *Philosophy and Phenomenological Research* 61 (3):619-638.

Driver, Julia (2001). *Uneasy Virtue*. Cambridge University Press.

Harman, Elizabeth (2011). Does moral ignorance exculpate? *Ratio* 24 (4): 443-468.

Harman, Elizabeth (forthcoming). The Moral Irrelevance of Moral Uncertainty.

Hawley, Katherine (2003). Success and Knowledge-How. *American Philosophical Quarterly* 40 (1):19 - 31.

Hawthorne, John & Stanley, Jason (2008). Knowledge and Action. *Journal of Philosophy* 105 (10):571-590.

Henson, Richard (1979). What Kant might have said: Moral worth and the overdetermination of dutiful action. *Philosophical Review* 88 (1):39-54.

Herman, Barbara (1981). On the value of acting from the motive of duty. *Philosophical Review* 90 (3):359-382.

Herman, Barbara (2007). *Moral Literacy*. Harvard University Press.

Hills, Alison (2009), Moral Testimony and Moral Epistemology. *Ethics* 120 (1): 94-127.

Holton, Richard (2010). Norms and the Knobe Effect. *Analysis* 70 (3):1-8.

Hopkins, John (2007). What is Wrong with Moral Testimony? *Philosophy and Phenomenological Research* 74 (3): 611-634.

Jones, Karen (1999). Second-Hand Moral Knowledge. *The Journal of Philosophy* 96 (2): 55-78.

Kant, Immanuel (1785/1996). *Groundwork of the Metaphysics of Morals*, in: Gregor, Mary (trans. and ed.) *Practical Philosophy*. Cambridge University Press.

Kearns, Stephen & Star, Daniel (2009). Reasons as Evidence. *Oxford Studies in Metaethics* 4:215-42.

Korsgaard, Christine M. (2008). *The Constitution of Agency: Essays on Practical Reason and Moral Psychology*. Oxford University Press.

Lackey, Jennifer (2008). *Learning From Words: Testimony as a Source of Knowledge*. Oxford University Press.

Lillehammer, Hallvard (1997). Smith on moral fetishism. *Analysis* 57 (3):187–195.

Manne, Kate (forthcoming). On Being Social in Metaethics. *Oxford Studies in Metaethics*.

Markovits, Julia (2012). Saints, Heroes, Sages, and Villains. *Philosophical Studies* 158 (2):289-311.

Markovits, Julia (2010). Acting for the Right Reasons. *Philosophical Review* 119 (2): 201-242.

Martin, Adrienne M. (2010). Owning up and lowering down: The power of apology. *Journal of Philosophy* 107 (10):534-553.

McGrath, Sarah (2011). Skepticism about Moral Expertise as a Puzzle for Moral Realism. *Journal of Philosophy* 108 (3).

McGrath, Sarah (2009). The puzzle of pure moral deference. *Philosophical Perspectives* 23 (1): 321-344.

Moody-Adams, Michele (1994). Culture, responsibility, and affected ignorance. *Ethics* 104 (2):291-309.

Schwitzgebel, Eric (2010). Acting contrary to our professed beliefs or the gulf between occurrent judgment and dispositional belief. *Pacific Philosophical Quarterly* 91 (4):531-553.

Schwitzgebel, Eric (2002). A phenomenal, dispositional account of belief. *Noûs* 36 (2):249-75.

Sher, George (2006). In Praise of Blame. Oxford University Press.

Smith, Michael (1994). *The Moral Problem*. Blackwell.

Stratton-Lake, Philip (2000). *Kant, Duty, and Moral Worth*. Routledge.

Peter F. Strawson (1962). Freedom and resentment. *Proceedings of the British Academy* 48:1-25.

Svavarsdottir, Sigrun (1999). Moral cognitivism and motivation. *Philosophical Review* 108 (2):161-219.

Twain, Mark (1884). *Huckleberry Finn*. Edited by Walter Blair and Victor Fisher (1988), University of California Press.

Wallace, Jay (1996). *Responsibility and the Moral Sentiments*. Harvard University Press.

Williams, Bernard (1981). Persons, Character, and Morality. In *Moral Luck*. Cambridge University Press.

Williamson, Timothy (2000). *Knowledge and its Limits*. Oxford University Press.