Praise without Perfection: A Dilemma for Right-Making Reasons [1]
Paulina Sliwa
University of Cambridge

0. Introduction

When you don't know what to do, you'd better find out. Sometimes the best way to find out is to ask for advice. And when you don't know what the *right* thing to do is, it's sometimes good to rely on *moral* advice. So much should be uncontroversial.

Yet this straightforward thought spells serious trouble for a popular and widespread approach to moral worth: on this approach, agents deserve moral praise for a right action only if they are acting on right-making reasons.

The first part of this paper argues that cases of moral advice present right-making reasons accounts with a dilemma: depending on how we make the right-making relation precise, we either have to deny that agents who seek out and follow moral advice are morally praiseworthy or we have to credit morally wrong actions by unsavory characters with moral worth. This casts doubt on the claim that acting on right-making reasons can be both necessary or sufficient for moral worth.

The second half of the paper explores an alternative proposal: what's required for moral worth is moral knowledge. This idea has been unpopular in recent literature. My aim is to show that it deserves serious consideration.

1. Moral Praise for Moral Advice

Not all right actions are morally praiseworthy. The politician delivering a huge check to a charity may well be doing something that is morally right – either because it's something that he is morally required to do or because, even if not required, it's something that's morally good. But we are not inclined to give him moral credit for it. Whether an action has moral worth depends not just on the moral status of the action but also on the agent's motivation. In the case of the politician we have reason to believe that his generosity stems from ulterior motives. His reasons for donating the money do not have anything to do with morality – they are purely self-interested. And so, it was a matter of luck that the politician did the right thing; morality and self-interest happened to align in this case. For an action to have moral worth, on the other hand, there must be a connection between the rightness of the action and the agent's motivation. An account of moral worth aims to pin down what such moral motivation consists in.

An account of moral worth tells us how an agent must be motivated to be good. Morally perfect agents may always perform morally worthy actions but it's not reasonable to think that *only* morally perfect agents perform morally worthy actions. We can be morally admirable for doing the right thing even when we fall short of moral perfection. This is something that a plausible account of moral worth should be able to accommodate. In this section, I present two examples of agents who compensate for their various moral limitations and who, as a result, succeed in doing the right thing. I argue that these agents are morally praiseworthy for their actions.

Consider the following:

> Ron is working for a charity. In response to a funding crisis, Ron needs to shut down one of two programs that the charity is running: X or Y. The two programs benefit different individuals in very different ways. Ron has spent the last couple of days researching the details and it's time for him to make a decision. But although he has a wealth of information, he's not sure what he should do. Different factors point in conflicting directions: X is slightly more efficient than Y but Y serves a more disadvantaged population, X focuses on education, while Y focuses on health, etc. Ron wants to do the right thing but he's feeling overwhelmed by the number of considerations and the time pressure. He decides to ask his senior colleague who has worked in the field for many, many years; Ron admires her good sense and competence. Having looked over the details, she rightly advises Ron to end program X rather than Y.

Ron relies on his colleague's moral advice because he is not in a position to know what the right thing to do is on his own. Ron does have access to all the relevant nonmoral facts, but given his lack of experience, he is not sure how to weigh them. The fact that he needs to arrive at a decision quickly makes him anxious and this adds to his uncertainty. It would, of course, be better if Ron was in a position to discern what the right thing to do is himself. But given his moral limitation, relying on the moral advice of his more experienced colleague is a good thing. Moreover, when, following his colleague's advice, he does the right thing and ends program X, it's not just an accident that Ron managed to do what's right. And so, it seems that Ron is morally praiseworthy for his right action.

Here is a second case:
> Johanna is a teacher in an inner-city school who cares deeply about social justice and equality. While reviewing her teaching logs over the summer, however, she notices a disturbing trend: she realizes that when punishing disruptive behavior, she tends to resort to harsh measures much more quickly for black students than for white students. Johanna is horrified. She resorts to the following strategy: over the next year, whenever she encounters disruptive behavior, rather than deciding on the punishment on the spot, she runs the facts by an experienced and trustworthy colleague, without mentioning the student's race. She then defers to the colleague's advice about what kind of punishment would be fair.

Johanna relies on her colleague's moral advice because she is not in a position to know how to respond to her students' behavior fairly. While she has access to all the relevant nonmoral facts, Johanna's implicit bias prevents her from evaluating them correctly. It would, of course, be better if Johanna was simply in a position to recognize herself what the right way to respond is to a particular instance of disruptive behavior. But given her moral limitation, it seems like a very good thing that Johanna relies on her colleague's advice rather than muddling through herself. Moreover, given that her colleague is sensible and reliable, it's not an accident that Johanna ends up doing the right thing and, let's suppose, send a particularly disruptive student to the principal's office. And so, we should give Johanna moral credit for her right actions.

Johanna and Ron deserve moral praise for both seeking out and acting on their moral advice. Given that they are relying on sound moral advice, it's not an accident that they succeed in doing what morality requires. In the next section I am going to argue that these cases present a challenge to accounts of moral worth on which moral worth is a matter of being motivated by right-making reasons.

2. A Dilemma for Right-Making Reasons

A popular approach to moral worth has it that an action has moral worth only if it is 'motivated by right-making reasons'. Most recently, variants of this approach have been defended by both Nomy Arpaly and Julia Markovits and my discussion will focus on their proposals. Arpaly argues:

> For an agent to be morally praiseworthy for doing the right thing is for her to have done the right thing for the relevant moral reasons – that is, the reasons for which she acts are identical to *the reasons for which the action is right*.[2]

Similarly, Markovits proposes:

> 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. [3]

Right-making reasons accounts (RMR accounts) of moral worth have two central elements. One is a thesis in the *philosophy of action* about moral motivation. The second is a *metaphysical* thesis about what right-making reasons are. The RMR accounts agree, to a large degree, on the former – they differ on the latter. Depending on what we take right-making reasons to be, then, I argue that cases of moral advice present RMR accounts with a dilemma: either actions on moral advice are ruled out from having moral worth, or actions that intuitively seem morally repugnant are ruled in.

As for the first, moral motivation is taken to require being motivated by right-making reasons. To be motivated by a right-making reason F, F needs to appear in a rationalizing explanation of the agent's action.[4] Crucially, to be motivated by a right-making reason F, F must appear *de re* in the rationalizing explanation. This is to say, if the fact that aspirin will relieve your pain makes it right to give you some, then it must be *the fact that aspirin will relieve your pain* that explains my action. This is different from being motivated by *the fact that it would be right* to relieve your pain or by *the fact that aspirin would relieve your pain makes it right to give you some*. It's only in the former case that I'm motivated by right-making reasons *de re*. In the latter case, I'm motivated by moral considerations *de dicto*.

But moral worth is not about being motivated by right-making reasons *de dicto*. It's about being motivated by right-making reasons *de re*. Thus Markovits, argues:

> The Kantian 'truly moral man' seems guilty of a kind of moral fetishism (to borrow a phrase from Michael Smith), or at best, of having 'one thought too many' (to borrow one from Bernard Williams), 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.[5]

Similarly, here is Arpaly:

> [M]oral worth is fundamentally about acting for moral reasons, not about acting for reasons believed or known to be such, and distinguishing the two is important in evaluating moral agents.[6]

This picture of moral motivation is widely accepted beyond the moral worth debate.[7] In defending this account of moral motivation, Arpaly and Markovits both appeal to Smith, who argues:

> Good people care non-derivatively about honesty, the weal and woe of their children and friends, the well-being of their fellows, [...] and not just one thing: doing what they believe to be

---

[2] Arpaly [2003], p. 72, highlighting mine.

[3] Markovits [2010], p. 205.

[4] For Arpaly this is because being motivated by a right-making reason involves a desire for that reason. Markovits allows that not all such motivation may involve belief-desire pairs. See Markovits [2010], p. 221-222.

[5] Markovits [2010], p. 203. See also Stratton-Lake [2000].

[6] Arpaly [2003], p. 73. See also p. 84.

[7] It is also influential in the debate about blame. For recent examples, see Harman [forthcoming] and Weatherson [forthcoming].

right, where this is read *de dicto* and not *de re*. Indeed, commonsense tells us that being so motivated is a fetish or moral vice not the one and only moral virtue.[8]

My own view is that this account of moral motivation is a nonstarter – moral motivation cannot just be about responding to right-making reasons *de re*. In this paper, however, I will not attack it directly.[9] Rather, I'm going to argue that there is no plausible of story of what right-making reasons are such that it yields a plausible account of moral worth. On the two most plausible ways of accounting for right-making reasons that have been defended in the literature, RMR accounts are hard pressed to avoid some very implausible consequences for which actions do and do not have moral worth.

2. The First Horn of the Dilemma: Right-Making Reasons as Explanation
Perhaps the most straightforward way to cash out "right-making" is as an explanatory relation. Right-making reasons then are those considerations that explain why an action is right. This is the account of right-making reasons that Arpaly seems to have in mind. For example, she argues that Kant's grocer, who deals honestly with his customers because he thinks that such a policy is most likely to benefit his business, is not morally praiseworthy for his action. And this is because the considerations that lead him to act are not those that explain why one should deal honestly with one's customers:

> In pricing fairly, the grocer acts for a reason that has nothing to do with morality or with the features of his action that make it morally right. The reasons for which he acts have to do only with his own welfare; and whatever it is that makes his action morally right, the fact that his action increases his welfare is certainly not what makes it morally right. His reasons for action do not correspond to the action's right-making features.[10]

Appeals to such a notion of right-making reasons are pervasive throughout moral theory and not just limited to the literature on moral worth. Rosen, for example suggests:

> If an act is wrong, there must be some feature of the act that makes it wrong. Any given act may be wrong for several reasons, and some of these reasons may be more fundamental than others. A breach of promise may be wrong because it is a breach of trust, and a breach of trust may be wrong because it is prohibited by principles for social cooperation that no one could reasonably reject.[11]

How to spell out the explanatory relation in detail has been the subject of recent debate. One promising line of thought is to understand the right-making relation as a specific instance of metaphysical grounding, where grounding is understood as a form of metaphysical dependence distinct from causation.[12] Much recent debate is concerned with spelling out the properties of the grounding relation; these details need not concern us here.[13] What matters is that there are metaphysical resources we can draw on to cash out the right-making relation as an explanatory relation along the lines I have suggested.

---

[8] Smith [1994], p. 75.

[9] I do so in Sliwa [ms].

[10] Arpaly [2003], p. 72. Note that to accept this need not commit us to an account on which *all* reasons are explanations, as defended by Broome [2004], pp. 28-55. This is because not all reasons need to be right-making reasons.

[11] Rosen [2010], p. 110. Other examples include Schaffer [2009], p. 375, Audi [2012] p. 106-7.

[12] See Väyrynen [2009] and [2013] for detailed discussion of the relationship between grounding and explanation.

[13] Properties that make grounding a plausible contender for underwriting the right-making relation is its asymmetry and necessitation. See, for example Rosen [2010].

If this is how we spell out right-making reasons, can we accommodate actions on moral advice? The answer is *No*. Let's focus Ron's action. Ron cuts off funding to charity X which is, we're assuming, the right thing to do. Ron makes his decision on the basis of his colleague's moral testimony. On this basis, he rightly takes ending program X to be *the right thing to do* and this is why he goes ahead with it. But the fact that his colleague tells Ron that the right thing to do is to cut off funding to charity X does not *make* it the right thing to do. The right-making reasons for ending program X are, plausibly, nonmoral facts about the program X: facts about its effectiveness, the services it offers, how indispensable these services are, etc. But Ron is not motivated by those facts; he does not know how to weigh them. Instead, he is motivated by his colleague's advice about what the right thing to do is. And so, Ron's action does not have moral worth. And, for similar reasons, neither does Johanna's.

One might try to resist this verdict as follows: one could argue that if Ron's or Johanna's advisors *only* told them what the right thing to do was and didn't give them any justification for why it was the right thing to do, then indeed their actions do not have moral worth. However, she might argue, typically this is not what advisers do. Plausibly, Ron's colleague will tell Ron that he needs to end program X because, say, it is less effective than the others. In this case, Ron will be motivated by right-making reasons, after all. And so their actions can have moral worth. Advisors generally tell us not just what the right thing to do is but also what makes it right.

I don't find this response convincing. By relying on his colleague, Ron may well come to believe not only what the right thing to do is but also what the right-making reason is. But this does not mean that he will be motivated *by* the right-making reason. In particular, he will not be motivated by the right-making reason *de re*. At most he will be motivated by *the fact that the program's effectiveness is a right-making reason*. He will be motivated by the right-making reason *de dicto*, not *de re*.

There are two ways of bringing this out. First, note that Ron knew that program X was less effective than the others even before asking his colleague for moral advice. And yet its being less effective did not motivate him to end it. What then explains that his uncertainty? How do we explain why he only became motivated to act on this consideration after hearing his colleague's advice? One possibility is that after hearing his colleague's advice, Ron suddenly acquires a strong non-instrumental *de re* desire for ending the less effective program. But this seems puzzling: why would his colleague's testimony lead to such a change in Ron's desires? It seems much more plausible that while Ron realized that program X was less effective even prior to his colleague's testimony, he didn't know that this was *the right-making reason* for ending the program. He was missing *de dicto* information. Based on his colleague's testimony, he comes to know that this is the right-making reason and this is what motivates him to act: he is motivated by the fact that program X is less effective *qua* right-making reason, i.e. *de dicto*, not *de re*.

Second, we can imagine asking Ron about why he ended program X. He might well respond that he ended it because it's less effective. But it does not follow that he was motivated by the fact that it's less effective *de re*. After all, suppose we follow up "But you knew it was less effective all along. Why did you have to ask your colleague for advice then?" Presumably, Ron would respond by saying "Well, I knew it was less effective than the other programs but I didn't realize that this was a good enough reason for ending it – there were so many other considerations at issue." That is, Ron himself would ultimately justify his motivation by pointing to *de dicto* considerations.

The verdict that moral advice is incompatible with moral worth should strike us as both counterintuitive and problematic. After all, Ron and Johanna, do not rely on moral advice for frivolous reasons. They are aware of their moral limitations and they care about doing what's right. When they succeed in doing what's right, it's not by lucky accident – as in the case of the Kantian grocer or the shrewd politician, for whom self-interest and morality aligned. They succeed in doing

what's right because they successfully compensate for their epistemic disadvantage. This is a morally admirable motivation, and an account of moral worth should give them credit for it.

3. The Second Horn of the Dilemma: Right-Making Reasons as Evidence

There is an alternative understanding of right-making reasons that looks better placed to accommodate actions on moral advice: perhaps, we should take right-making reasons to provide evidence for the moral status of an action, rather than grounding or explaining it. On this view, right-making reasons are whatever *justifies* our actions. Such an account has been defended by Markovits – she argues:

> 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.[14]

Such an evidential account of right-making reasons accommodates actions on moral advice as having moral worth. This is because testimony is a source of evidence. Thus, insofar as Ron's colleague is trustworthy and reliable and Ron has good reason to regard her as such, her testimony provides Ron with evidence that ending program X is the right thing to do. And so, his colleague's testimony is itself a right-making reason. As Markovits argues:

> If expert testimony gives us most reason to believe that some act would be best, then this testimony *is the reason* why we should perform that act [...] our agent's acting so may be *made* right by the fact that the agent advises him to do so.[15]

So far so good. But troubling consequences quickly emerge. If right-making reasons are evidence for an action's moral status, and acting on right-making reasons is sufficient for moral worth, then morally worthy actions are too easy to come by. Such an account is much *too inclusive*: it counts actions as morally praiseworthy that are both morally wrong and morally repugnant.

Imagine a situation in which the fact that an injection is more painful than the alternatives is evidence for its effectiveness. In such a scenario, the fact that injecting you with the drug will cause you more pain than injecting you with an alternative can be a right-making reason. And so, a sadistic doctor who chooses the painful injection because it's painful and because she non-instrumentally cares about causing you pain is, on an evidential account of right-making reasons, motivated by a right-making reason. Thus, her action has moral worth: she is morally praiseworthy for injecting you with the more painful medication. This looks like a deeply implausible consequence.

Second, imagine Linda who grows up in a small tight-knit rural community. Her friends and neighbors are good people: they are honest and hard-working, they give to charity and help each other out. But her community is also deeply conservative and so, from an early age, Linda is taught that gay marriage is a grave moral evil that it's her duty to oppose. And so, she donates money to various organizations that oppose gay marriage and takes part in a letter-writing campaign against its legalization. Linda's belief that homosexuality is a grave moral evil is based on the testimony of those around her whom Linda has both good reason to regard as reliable and trustworthy. Moreover, these people really are, for the most part, good people even if they harbor misguided moral views about gay marriage. And so, it seems that Linda's unfortunate moral belief, while false, may well be justified. As Markovits argues:

> [I]f 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.[16]

In donating money to anti-gay marriage organizations and taking part in letter-writing campaigns, Linda then acts on right-making reasons. And so, Linda is morally praiseworthy for her support of organizations that rally against gay marriage.

Again, this is deeply implausible. Given Linda's sheltered upbringing, we may be willing to concede that she is *not blameworthy* for her false moral views and for acting on them. But to insist that her action has moral worth – that she is *morally praiseworthy* for donating money to organizations whose aim is to perpetuate discriminatory treatment based on sexual orientation – is going decidedly too far.

Where does this leave us? I have argued that a widely shared assumption about moral motivation – namely that moral motivation involves being motivated by right-making reasons *de re* – leaves us in a tight spot with regards to a certain class of intuitively morally admirable actions: actions on moral advice. Depending on how we make right-making reasons precise, moral worth becomes either too hard – it excludes actions on moral advice – or too easy – we have to include the sadistic doctor's giving a painful injection and misguided Linda's supporting homophobic causes. If we take the first option, we have to conclude that acting on right-making reasons cannot be necessary for moral worth. If we opt for the second, we have to conclude that it cannot be sufficient.

I suggest that this gives us good reason to look for an alternative – one that departs from the underlying picture of moral motivation.

4. The Importance of Moral Knowledge

The aim of this section is to lay out an alternative approach to moral worth. Here is the proposal:

> *The Knowledge Principle*: An action has moral worth *if and only if* the agent acts from moral knowledge.

I take it that an agent acts from moral knowledge if and only if she is motivated by her knowledge that the action is the right thing to do. In this section, I provide some positive arguments for both the necessity and the sufficiency claim. In the next section I address the most pressing objection.

The Knowledge Principle as a *sufficient* condition for moral worth is motivated by the cases of moral advice that I discussed in Section 1. Recall Ron and Johanna. They are motivated to seek out moral advice because they recognize that (for different reasons) they are not in a position to know what the right thing to do is in their situation. They rely on moral advice because they have reason to think that their advisors are epistemically better placed – that they are trustworthy and reliable. When we are morally uncertain, then relying on moral advice is a morally good thing. Moral testimony can be a source of moral knowledge. Thus, it seems plausible that by relying on moral advice Ron and Johanna can come to know what the right thing to do is. And this is what motivates them to act: Ron, for example, ends program X because he knows this is the right thing to do.[17]

If *The Knowledge Principle* is a sufficient condition for moral worth, then this explains why we regard Ron and Johanna as morally praiseworthy for their right actions. They are morally praiseworthy

---

[16] Markovits [2010], p. 219.

[17] Even those who object to moral testimony (e.g. Hopkins [2007], McGrath [2011] and [2009], Hills [2009]) do not deny that it can be a source of moral knowledge. For a general defense of moral testimony, see Sliwa [2012].

because they are motivated by their knowledge of what the right thing to do is.[18] It also explains why they seem morally praiseworthy not just for acting on moral advice but also for seeking it out in the first place. When you are morally uncertain, resolving your moral uncertainty is the right thing to do. Plausibly, Ron and Johanna know this and this motivates them to look for reliable and trustworthy moral advice. And so, they are morally praiseworthy for it.

Thus, one reason to regard *The Knowledge Principle* as a plausible sufficient condition is that it gives us an intuitively compelling way to accommodate actions on moral advice as having moral worth. A second reason is that an analogous condition seems sufficient for nonmoral praise. Suppose a doctor correctly treated a patient for an ear infection. And suppose that she succeeded because she knew that the patient had an ear infection and she knew what the right way to treat it is. This seems sufficient for us to give the doctor credit for both her correct diagnosis and the correct treatment. Note that we give the doctor credit for it, even if she was initially stumped by the patient's symptoms and consulted a textbook for both the diagnosis and the correct treatment. Since it would be surprising if our practice of moral and nonmoral praise radically diverged, this gives us some reason to think that *The Knowledge Principle* may be sufficient for moral praise, also.

Are there reasons for thinking that *The Knowledge Principle* could be a *necessary* condition for moral worth? Yes – we see this if we step back and recall a central feature of morally worthy actions: an agent's action has moral worth only if it's not a fluke that she did the right thing. The rightness of her action must be connected in some reliable way to her motivation, so as to make the agent's doing the right thing counterfactually robust. For this reason Kant's shopkeeper and the shrewd politician are not morally praiseworthy even when they do act rightly. Their motivations are self-interested, and it's a matter of luck that self-interest and morality align in their particular situation – all too often they do not.

Moral knowledge provides such a reliable link between the rightness of an action and the agent's motivation: if the agent knows that an action is right and she is motivated by this knowledge, then it's not an accident that she acts rightly. This is because knowledge is counterfactually robust: if you know that p, then you couldn't easily have been wrong about it.

By requiring moral knowledge for moral worth we evade the counterintuitive consequences of the evidential RMR account. Misguided Linda is not morally praiseworthy for donating money and time in order to oppose gay marriage. Even if we allow that Linda's belief that gay marriage is a grave moral evil is epistemically justified, it is false. Linda does not know that gay marriage is a moral evil and so she doesn't know that it's right to oppose it. So, she cannot be motivated by her moral knowledge. And so, her action does not have moral worth.

The Knowledge Principle also rightly excludes the sadistic doctor from moral praise for giving you the painful injection. The fact that the injection is painful may well be evidence on which the doctor can come to know that giving it to you is the right thing to do. But to be morally praiseworthy, it's not enough that the sadistic doctor knows or is in a position to know that she ought to give you the injection. She needs to *act on* this moral knowledge: she needs to give you the painful objection *because* she knows this to be the right thing to do. That is, she needs to be motivated both by her knowledge of what the right thing to do is and by a desire to do what's right. As described, however, the sadistic doctor is motivated by the fact that the injection is painful.

---

[18] *The Knowledge Principle* does not imply that *any* action performed on reliable moral testimony has moral worth. This is because acquiring moral knowledge based on testimony is not trivial; the agent needs to be in an epistemic position to identify a reliable advisor. This presupposes moral competence on the part of the agent. Someone who is morally incompetent are not in a position to acquire moral knowledge by testimony even if they chance upon a reliable advisor. And so, their actions will not have moral worth.

The Knowledge Principle thus rejects the picture of moral motivation that underpins RMR accounts. On this picture, moral motivation was a matter of being motivated by right-making reasons *de re*. In contrast, *The Knowledge Principle* requires agents to be motivated by the moral status of the action: the agent needs to know that her action is the right thing to do and she must be motivated by a desire to do what's right.

Still, you might ask, why should knowledge be required? Would justified true belief not do? I don't think it would. Consider a modified scenario involving Linda. Linda goes to a lecture about the morality of gay marriage and she hears one of the members of her community state what sounds like passionate support of gay marriage. Because Linda arrived late and is leaving early, the fact that the speaker is merely quoting 'the liberal press' eludes her. Feeling inspired, Linda makes a donation to a pro-same-sex marriage group when she gets home. Linda's belief that same-sex marriage is morally unproblematic is true and let's assume that we can fill in the details so that it is also justified. Still, I don't think we would regard Linda as morally praiseworthy for her donation. While, in supporting marriage equality, she did the right thing, her success seems too fragile and accidental to merit moral praise. She could have very easily done the wrong thing and donated money to an organization opposing same-sex marriage, had she only listened to the lecture for a couple of minutes longer.

The Knowledge Principle is attractive as an account of moral worth for two reasons. First, it accommodates much of the insight and appeal of both the explanatory and the evidential RMR account. If moral knowledge is necessary for moral worth, then being sensitive to those features that ground the moral status of our actions is a very good thing indeed. That's because it's hard to see how an agent could be in a position to discriminate right from wrong without being sensitive to these features. Similarly, if color physicalism is correct and what makes something red are its reflectance properties, then it's hard to see how an agent could be in a position to reliably discriminate red from yellow without being sensitive to the underlying reflectance properties. *The Knowledge Principle* also accounts for the intuitive plausibility of evidential right-making reason accounts: moral knowledge, like all knowledge, must be based on evidence. And moral evidence may come in many kinds: plausibly, it includes the underlying nonmoral facts grounding its moral status, but it also includes moral testimony.

We said that morally worthy action is supposed to reflect *something good* about the agent. As Markovits notes:

> [M]orally worthy actions are the building blocks of virtue – a pattern of performing them makes up the life of a good person.[19]

A morally good person is someone who generally performs morally praiseworthy actions. This applies more widely than just in the moral case. Plausibly, a good doctor is someone who, in the context of treating her patients, generally performs actions that are medically praiseworthy. A good car mechanic is someone who, in the context of repairing cars, generally performs actions that are praiseworthy. And in these nonmoral cases, knowledge is a crucial ingredient to such praise: a good doctor is someone who is medically competent. She makes medical decisions that reflect her medical knowledge. The good car mechanic is someone who is knowledgeable about fixing cars; her decisions reflect her knowledge.

Being a good doctor or a good car mechanic thus involves competence: it requires being knowledgeable about one's subject matter. Why should the moral case be any different? It seems natural to think that someone who is morally good is someone who is morally competent: they know right from wrong. Just as a good doctor is someone whom we might ask for medical advice, it

---

[19] Markovits [2010], p. 203.

is natural to think that the morally good person is someone whom we can ask for moral advice. And just as the good doctor's medical decisions reflect her medical knowledge, it seems natural to suppose that the actions and decisions undertaken by a morally good person reflect her moral knowledge.

These arguments are hardly conclusive and much more needs to be said to fill them in.[20] But I hope what I have said so far shows that *The Knowledge Principle* deserves serious consideration as a candidate account of moral worth. In the next section, I defend it from an influential objection.

5. *The Knowledge Principle*, Defended

The claim that moral knowledge should be a necessary condition for moral worth is widely seen as a nonstarter. A central reason for this is the case of Huckleberry Finn. Huckleberry Finn makes up an elaborate lie and thereby protects his friend, the fugitive slave Jim. But Huckleberry, it is argued, does not know that this is the right thing to do – in fact, he falsely believes that morality requires him to turn Jim over to the authorities. Still, it's argued, Huckleberry is intuitively admirable for protecting the fugitive slave Jim. But then moral worth cannot require moral knowledge.[21]

If Huckleberry does not know that protecting a fugitive slave is the right thing to do, then *The Knowledge Principle* does commit us to the verdict that he is not morally praiseworthy for it. But, contrary, to received wisdom, I argue that this is not as implausible a verdict as it is generally taken to be: to say that Huckleberry's action is not *morally* praiseworthy is not to deny that it is admirable in other ways.

Let's consider Huckleberry's motivation in detail. Believing that helping a fugitive slave escape is akin to 'stealing', Huckleberry was prepared to turn Jim over to the authorities. Then he changes his mind on what to do, without changing his beliefs about slavery. What then happened? According to RMR accounts of moral worth, Huckleberry Finn, despite his false moral beliefs, responded to Jim's humanity.[22] Since Jim's humanity is a right-making reason, Huckleberry's case is taken to motivate RMR accounts of moral worth. Thus, Arpaly argues:

> [W]hen the opportunity comes to turn Jim in and Huckleberry experiences a strong reluctance to do so, his reluctance is to a large extent the result of the fact that he has come to see Jim as a person...[23]

It's true that Huckleberry experiences a strong reluctance to turn Jim in. But I'm not sure that the best explanation of this is Jim's personhood. This seems to leave out the role of Huckleberry's friendship with Jim. Thus, consider what Jim shouted out to Huckleberry, just before Huckleberry wavered:

> 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 forgit you, Huck; you's de bes' fren' Jim's ever had; en you's de only fren' ole Jim's got now.[24]

It's upon hearing Jim calling out to him that Huckleberry's resolve melts 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.[25]

---

[20] For more discussion, see Sliwa [ms].

[21] For versions of this argument, see Arpaly [2003], Markovits [2010], Driver [2000].

[22] For a similar argument, see Markovits, p. 208.

[23] Arpaly [2003], p. 77.

[24] Twain, p. 103.

[25] ibid.

Thus, it seems that what was crucial for Huckleberry's change of heart was a reminder of his friendship with Jim. Arpaly herself considers the importance of Huckleberry's and Jim's relationship:

> Twain makes it very easy for Huckleberry to perceive the similarity between himself and Jim: the two are equally ignorant, share the same language and superstitions [...] While Huckleberry never reflects on these facts, they do prompt him to act toward Jim, more and more, in the same way he would have acted toward any other friend.[26]

But then it seems very odd to explain Huckleberry's weakening resolve by appealing to a desire to *protect a person*. Rather, it seems likely that Huckleberry was motivated by considerations of friendship. After all, one of the central tenets of friendship is loyalty – you don't just let your friends down when they are in need. And plausibly Huckleberry knows that. Huckleberry's change of heart then is in part explained by his desire to be a good friend and his knowledge of what friendship requires. And so, it seems plausible that he is praiseworthy *for being a good friend*.

Does Huckleberry deserve *moral* praise for helping his friend in need?[27] It's plausible that in the situation, Huckleberry is morally required to help his friend.[28] It's less clear that Huckleberry *knows* that he is morally required to help his friend. But even if Huckleberry is not morally praiseworthy for helping his friend, this is compatible with his being praiseworthy for being a good friend. Being a good friend is something that we value. And so it's not surprising that Huckleberry's action should strike us as admirable. Even if Huckleberry is precluded from moral praise by his ignorance, *The Knowledge Principle* allows that he may be praiseworthy in other, nonmoral, ways.[29] And so, *The Knowledge Principle* can accommodate the intuition that there is something admirable about Huckleberry's action, even if he is not morally praiseworthy for protecting a fugitive slave.

A second reason for insisting that Huckleberry deserves moral praise for his action is the thought that to deny it has implausible implications for our evaluation of Huckleberry Finn. Arpaly worries that it commits us to seeing Huckleberry as a bad boy who just chanced upon the right action. But as Arpaly insists:

> Huckleberry Finn [...] is not a bad boy who has accidentally done something good, but a good boy.[30]

The Knowledge Principle tells us how to evaluate *actions*, not agents. However, it's plausible to think that an account of the former will have implications for the latter. The objection here is, then, that *The Knowledge Principle* has implausible implications for how we evaluate agents.

But to deny that Huckleberry is morally praiseworthy for helping a fugitive slave is not to say that he is a 'bad boy who has accidentally done something good'. The objection presents us with a false

---

[26] ibid.

[27] It's worth noting that Bennett [1974], who introduced this case to the literature did not think that Huckleberry Finn was morally praiseworthy for his action either, though on different grounds. He thought Huckleberry was not responding to moral reasons but just acting from sympathy.

[28] See Manne [forthcoming] for more discussion on this point.

[29] Similarly, consider a doctor who encounters a wounded villain and, after some hesitation, saves the villain's life. We can imagine that she has reached the wrong conclusion: the morally right thing would have been to let the villain die. On *The Knowledge Principle* she is thus not morally praiseworthy for saving the villain's life. But this is not to say that there is nothing admirable about what she did. Insofar as she was motivated by her knowledge of what the medically right thing to do was, she may well deserve nonmoral credit for being a good doctor. *The Knowledge Principle* only rules out her deserving moral praise in this situation; it is compatible with her deserving *medical* praise. Thanks to John Brunero and Jonathan Way for pressing me on this point.

[30] Arpaly [2003], p. 78. See also Driver [2000], p. xvi for a similar point.

dilemma. Whether an agent is morally good does not hinge on the moral worth of one individual action: a good doctor can sometimes be merely lucky to have made the correct medical decision – she rightly prescribed Tylenol for a high fever but she didn't know the child was running a high fever; the thermometer was malfunctioning. Similarly a morally good agent can sometimes perform a right action by mere luck: she did her best and it turned out all right even though she lacked moral knowledge of what to do in these circumstances.

While Huckleberry is not morally praiseworthy for protecting a fugitive slave, he plausibly is morally praiseworthy for plenty of other actions. Huckleberry has false beliefs about slavery but he shows sound moral judgment in plenty of other situations: he knows that one must keep one's promises, be loyal to one's friends, and apologize when having wronged someone. Huckleberry is not morally ignorant; and this is precisely what makes him a good boy. Arpaly then is right that Huckleberry is not a "bad boy who has accidentally done something good"; he is indeed a 'good boy.' But this is compatible with his having been lucky to avoid a serious moral mistake in this particular case.

The fact that *The Knowledge Principle* excludes Huckleberry's helping a fugitive slave from moral worth may have some intuitive cost. But I have argued that this cost has been overstated: denying that Huckleberry is morally praiseworthy for helping a fugitive slave does not commit us to the claim that there is nothing admirable about his action or that he cannot perform morally worthy actions on other occasions.

Second, as we have seen, RMR accounts of moral worth come at a significant price, themselves. One way of spelling out such an account commits us to denying that agents can be morally praiseworthy for acting on sound moral advice. But moral advice is a central and pervasive part of our moral lives; moral practice is not an individual pursuit. We rely on various advisors (friends, family, preachers, sometimes the internet) to navigate dicey professional situations, difficult life decisions, and tricky conflicts. This is a good thing: it enables us to compensate for our biases, lack of experience, or lack of perspective. An agent who is aware of her moral and epistemic limitations, as Ron and Johanna are, and compensates for them is morally admirable. To deny that such an agent deserves moral praise is deeply implausible. It's to deny that humility is a virtue. And on the second way of spelling out an RMR account, the intuitive costs are prohibitive: however implausible it is to exclude Huckleberry's action, it is certainly more implausible to maintain that the sadistic doctor and misguided Linda deserve moral praise.

6. Conclusion

Moral worth is widely taken to be a matter of acting "for the right reasons". This paper has argued that one popular way of cashing out this claim – taking moral worth to be a matter of being motivated by right-making reasons – fails as an account of moral worth. RMR accounts as defended in recent literature, I have argued, face a dilemma over actions on sound moral advice. Agents sometimes are morally praiseworthy for acting on moral advice – relying on moral advice allows them to compensate for their moral and epistemic limitations. RMR accounts have difficulties accommodating this. If right-making reasons are those considerations that ground the moral status of actions, then actions on moral advice cannot have moral worth. And so, acting on right-making reasons cannot be necessary for moral worth. If right-making reasons are considerations that provide evidence for the moral status of actions, then acting on right-making reasons cannot be sufficient for moral worth. I have then explored an alternative possibility: that moral worth is a matter of acting on moral knowledge. I have argued that this proposal – *The Knowledge Principle* – allows us to accommodate much of the insight behind RMR accounts, while avoiding its counterintuitive consequences. Most importantly, it sees the practice of moral worth and moral goodness as continuous with nonmoral praise and goodness. Knowledge is essential to successful actions in nonmoral domains. Being a good doctor or a good car mechanic is a matter of being

knowledgeable about one's subject matter and letting this knowledge guide one's actions. Why think it should be different in the case of morality?

7. Bibliography

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

Audi, Paul. 2012. "Grounding: Toward a Theory of the In-virtue-of Relation," *Journal of Philosophy* vol. 109, no. 12, pp. 685-711.

Bennett, Jonathan. 1974. "The Conscience of Huckleberry Finn," *Philosophy* vol. 49, no. 188, pp. 123-134.

Broome, John. 2004. "Reasons," in *Reason and Value: Themes from the Moral Philosophy of Joseph Raz*, eds. Jay Wallace, Michael Smith, Samuel Scheffler, and Philip Pettit (New York: Oxford University Press), pp. 28-55.

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

Harman, Elizabeth. Forthcoming. "The Irrelevance of Moral Uncertainty," *Oxford Studies in Metaethics.*

Hills, Alison. 2009. "Moral Testimony and Moral Epistemology," *Ethics,* vol. 120, no. 1, pp. 94-127.

Holton, Richard. Unpublished manuscript. "Inverse akrasia and weakness of will".

Hopkins, John. 2007. "What is Wrong with Moral Testimony?" *Philosophy and Phenomenological Research*, vol. 74, no. 3, pp. 611-634.

Jones, Karen. 1999. "Second-Hand Moral Knowledge," *The Journal of Philosophy,* vol. 96, no. 2, pp. 55-78.

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

Kearns, Stephen & Star, Daniel. 2011. "On good advice: a reply to McNaughton and Rawling," *Analysis,* vol. 71, no. 3, pp. 506-508.

------. 2009. "Reasons as Evidence," *Oxford Studies in Metaethics*, vol. 4, pp. 215-42.

------. 2008. "Reasons: Explanations or Evidence?" *Ethics*, vol. 119, no. 1, pp. 31-56.

Manne, Kate. Forthcoming. "On Being Social in Metaethics," *Oxford Studies in Metaethics*.

Markovits, Julia. 2012. "Saints, Heroes, Sages, and Villains," *Philosophical Studies,* vol. 158, no. 2, pp. 289-311.

-----. 2010. "Acting for the Right Reasons," *Philosophical Review,* vol. 119, no. 2, pp. 201-242.

McGrath, Sarah. 2011. "Skepticism about Moral Expertise as a Puzzle for Moral Realism," *Journal of Philosophy,* vol. 108, no. 3, pp. 111-137.

-----. 2009. "The Puzzle of Pure Moral Deference," *Philosophical Perspectives,* vol. 23, no. 1, pp. 321-344.

Rosen, Gideon. 2010. "Metaphysical Dependence: Grounding and Reduction," in *Modality: Metaphysics, Logic, and Epistemology,* eds. Bob Hale and Aviv Hoffmann (New York: Oxford University Press), pp. 109-136.

Schaffer, Jonathan. 2009. "On What Grounds What," in *Metametaphysics: New Essays on the Foundations of Ontology,* eds. David Manley, David J. Chalmers, and Ryan Wasserman (New York: Oxford University Press), pp. 347-383.

Sliwa, Paulina. 2012. "In Defense of Moral Testimony," *Philosophical Studies* 158 (2):175-195.

-----. Unpublished Manuscript. "Moral Knowledge and Moral Worth".

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

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

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

Väyrynen, Pekka. 2013. "Grounding and Normative Explanation," *Aristotelian Society Supplementary Volume*, vol. 87, no. 1, pp. 155-178.

Weatherson, Brian. Forthcoming. "Running Risks Morally," *Philosophical Studies*.