

Moral Worth, Credit, and Non-Accidentality

Keshav Singh

Penultimate draft - forthcoming in *Oxford Studies in Normative Ethics* (vol. 10)

1. Introduction

When someone does the right thing – what she morally ought to do – there is a certain kind of success involved. Her action complies with a moral standard. However, it is a further question whether this success is attributable to her – in other words, whether she is creditworthy for doing the right thing. This is the question of whether or not her action has *moral worth*. While extant accounts of moral worth have to some degree recognized the importance of the connection between moral worth and credit, they have been unable to fully capture it.¹ The goal of this paper is to develop an account that does.

The necessary connection between moral worth and credit can be given voice through the claim that one cannot perform a morally worthy action by accidentally doing the right thing. We don't deserve credit for complying with some standard if our compliance is a mere accident. Because morally worthy action consists in the agent's deserving credit for doing the right thing, one cannot perform a morally worthy action by accidentally doing the right thing. Call this the 'non-accidentality condition.'

Contemporary accounts of moral worth endorse the non-accidentality condition. They also agree that whether a token action is non-accidentally right in the relevant sense depends on its agent's motivations. In order for her action to be morally worthy, an agent must not only do the right thing, but do it for the right reasons. These accounts disagree, however, about which motivating reasons are the right reasons to satisfy the non-accidentality condition.

There are two main views here, which have gone by various names in the literature. I'll call them the Right-Making Features View and the Rightness Itself View. According to the Right-Making Features

¹ Some put this point in terms of praiseworthiness rather than creditworthiness. But creditworthiness is a better notion to appeal to because creditworthiness for different kinds of success is more closely tied to whether agents' success is attributable to them, whereas praiseworthiness is more closely tied to social and other factors that are orthogonal to moral worth. I follow Lord (2017) in this terminological choice.

View, morally worthy action is motivated by the features of the action that make it right. According to the Rightness Itself View, morally worthy action is motivated by the action's rightness itself.²

Defenders of each view think theirs is uniquely able to satisfy this non-accidentality condition. By contrast, much of what I'll argue in this paper is that *neither* view secures non-accidentality. Each view leaves room for cases where agents don't deserve credit for doing the right thing, despite their actions being motivated in the requisite way.

As I'll argue, each of these views makes a similar mistake. Each assumes morally worthy actions have as their motivational content exclusively either right-making features or rightness itself. In this paper, I'll develop an account of moral worth that rejects this assumption. On my account, a right action has moral worth if and only if it is motivated by sufficient moral reasons *as such*. I call this account the Guise of Moral Reasons account. The Guise of Moral Reasons account combines insights from both the Right-Making Features View and the Rightness Itself View in a way that satisfies the non-accidentality condition.

The paper proceeds as follows. In §2, I further motivate and explain the non-accidentality condition. §§3-4 argue that both the Right-Making Features View and the Rightness Itself View fail to satisfy it. In §5, I develop my own account and show that it succeeds at securing non-accidentality. §6 ties up some loose ends, including what my account says about the oft-discussed case of Huck Finn. §7 offers brief concluding remarks.

2. The Non-Accidentality Condition

If an agent performs a morally worthy action, she is creditworthy for doing the right thing. If an agent merely accidentally does the right thing, she is not creditworthy for doing the right thing. This makes it clear that if an agent performs a morally worthy action, she can't have merely accidentally done the

² The debate over the right kind of motivation for morally worthy action is sometimes framed in terms of *de dicto* versus *de re* moral motivation. I avoid this framing because I think it is mistaken for a variety of reasons. One such reason is that we can be motivated by rightness itself, or by right-making features, under various descriptions. Would this count as *de dicto* or *de re* moral motivation? The *de dicto/de re* distinction simply doesn't track what's at stake in the debate.

right thing. Some form of non-accidentality condition is endorsed by nearly everyone in the debate over moral worth.

The most obvious cases of accidentally doing the right thing are those in which someone does the right thing but is in no sense motivated by moral considerations. A classic example is Kant's prudent shopkeeper, who prices his goods fairly for the reason that doing so will increase his profits. The shopkeeper's action is morally right, but lacks moral worth because he is motivated by self-interest, not anything having to do with morality. In an important sense, it's a mere accident that he does the right thing.

As Kant puts it in the *Groundwork*, the prudent shopkeeper's actions lack moral worth because their connection to the moral law is "only very contingent and precarious" (4:390). But Kant's remark is potentially misleading, since the language of contingency and precariousness suggests that the relevant sense of non-accidentality is a kind of modal stability. This can't be right, for even if motivation by self-interest led the shopkeeper to do the right thing with extreme reliability, this wouldn't suffice for moral worth. This is why it's crucial to stress that what's really behind the non-accidentality condition is a point about creditworthiness. The sense of non-accidentality we're interested in here is the sense in which the appropriate connection obtains between the rightness of an action and its agent's motivation such that the agent is creditworthy for doing the right thing. And that connection is not a modal one.³

In order to understand moral worth, we must understand what this connection is, and what kind of motivation can secure it. The literature on moral worth often characterizes this motivation simply as motivation that reflects well on the agent.⁴ But in light of the above, it's clear that the requisite motivation can't be just any such motivation. There are lots of motivations that reflect well on agents, but do nothing to secure non-accidentality.

For example, consider the following case:

³ For further argument against modal conditions on moral worth, see Dishaw (ms). See also Faraci (forthcoming) for an argument against modal conceptions of non-accidentality in the epistemic domain.

⁴ See, e.g., Arpaly (2002, p. 224) and Markovits (2010, p. 203).

Promise. Dora’s motivation to attend the concert is that she believes the proceeds will go to charity. Actually, this is not the case – she was given misleading evidence. She has also forgotten that she promised a friend that she would go, making attending the concert the right thing to do. When she shows up at the concert, she discovers with dismay that the proceeds will not go to charity. However, upon seeing her friend, she remembers her promise and realizes that she has accidentally kept it.⁵

Dora’s motivation to attend the concert reflects well on her, as beneficent motives usually do. But as it turns out, her motivation has nothing to do with the rightness of her action. Despite Dora’s motivation reflecting well on her, this is still a case of accidentally doing the right thing. Importantly, it’s not that Dora doesn’t deserve credit for anything in this case; rather, it’s that she doesn’t deserve credit for the rightness of her action *per se*.⁶

Promise demonstrates that morally worthy action is not just right action that is laudably motivated. Rather, it’s action where the agent deserves credit *for doing the right thing* – that is, the agent must be creditworthy for the rightness of her action, not just for being laudably motivated. This connection between motivation and the rightness of action is the kind of non-accidentality that’s necessary for moral worth.

The above shows that there must be such a connection between the rightness of an agent’s action and its underlying motivation. But it doesn’t settle what kind of motivation secures this connection. The Right-Making Features View and Rightness Itself View offer competing answers. However, neither of them offers an answer that secures the right kind of connection. In the next two sections, I’ll illustrate the insufficiency of both views by giving examples of right action where the requisite motivation is present, but the agent nevertheless accidentally does the right thing.

3. The Right-Making Features View

According to the Right-Making Features View (hereafter RMF), morally worthy action is motivated by features of the action that make it right. Defenders of this view include Nomy Arpaly (2002) and Julia Markovits (2010). Both Arpaly and Markovits put things in terms of acting for moral reasons,

⁵ This case is inspired by a similar case from Johnson King (forthcoming a).

⁶ We might say that Dora’s action is accidentally right in much the same way a Gettierized belief is accidentally true. In Gettier cases, because the subject’s belief is justified, she still deserves credit, but not for the truth of her belief *per se*.

where moral reasons are understood as the right-making features of actions. As Arpaly puts it, morally worthy action consists in responsiveness to moral reasons. As Markovits puts it, it consists in a match between the agents' motivating reasons and her moral reasons.

So, for defenders of RMF, the contents of worth-conferring motivations are right-making features of those actions. This correctly excludes obvious counterexamples like Kant's self-interested shopkeeper because, although it's right for him to price his goods fairly, that it will increase his profits is not what makes it right.

There's also considerable intuitive pull to two ideas behind RMF. One is that morally worthy actions are performed on the basis of moral reasons – considerations that count in favor of them, morally speaking. The other is that right-making features are paradigmatic examples of moral reasons. In fact, on some views of normative reasons more generally, normative reasons for actions are just those facts that ground the overall deontic statuses of those actions.⁷

Arpaly and Markovits both claim that RMF satisfies the non-accidentality condition.⁸ For both of them, morally worthy action is motivated by considerations that in fact make the action right, whether or not the agent conceives of them as such. So, morally worthy action consists essentially in acting for (motivating) reasons that are moral reasons.

There's some sense in which actions so motivated are not merely accidentally right. However, it's not at all clear that they're non-accidentally right in the sense that entails that the agent is morally creditworthy for their rightness. This is because matches between the agent's motivating reasons and moral reasons can themselves be accidental.

This much is suggested by the unfortunately telling name of Markovits' thesis: The Coincident Reasons Principle. This principle says "my action is morally worthy if and only if my motivating reasons for acting coincide with the reasons morally justifying the action" (p. 205). Ironically, Markovits' use of the language of coincidence brings out precisely how RMF leaves room for an

⁷ See, e.g., Broome (2013).

⁸ See Arpaly (2000, p. 231) and Markovits (2010, p. 211).

objectionable form of accidentality. If, through mere coincidence, the reasons for which I act just *happen* to be features of the action that make it right, I don't deserve credit for the rightness of my action. In the relevant sense, I merely accidentally do the right thing.

Here's how to generate a case where someone acts on the basis of right-making features, but isn't creditworthy for doing the right thing: imagine someone who does the right thing on the basis of its right-making features, but has no idea that they are the right-making features, nor any particular concern for whether they are. Though such an agent happens to act on the basis of considerations that make the action right, his motivations are no more creditworthiness-conferring than those of the self-interested shopkeeper.

At this point, a clarification is in order. For RMF to have any plausibility, not just anything that plays a role in grounding the rightness of an action can count as a 'right-making feature' for the purposes of the view. Otherwise, we could get cases where the agent is motivated by some feature of her action that has no intrinsic moral significance, but just happens to be a partial ground of the action's rightness in a particular case. For example, even defenders of RMF would want to formulate their view to rule out the following as a case of morally worthy action:

Venom. Jack, a surgeon, is hiking when he sees a stranger get bitten by a venomous snake and faint. He immediately makes an incision near the bite so that the venom will drain out. Making the incision is the right thing to do, and Jack's reason for doing it (that it will allow the venom to drain out) is part of what makes it right. But Jack doesn't have any particular concern for doing the right thing in this case, nor does he conceive of his reason as one that makes his action right. He is simply intrinsically interested in draining venom out of wounds.⁹

The fact that it will allow the venom to drain out is a partial ground of the rightness of Jack's action. But it surely cannot count as a right-making feature in the relevant sense. This is worth clarifying

⁹ In an earlier version of this paper, I used **Venom** as a counterexample to RMF. As several people who read that version pointed out, while **Venom** may be a counterexample to the letter of extant formulations of RMF, it's not a counterexample to the spirit of the view.

because it isn't made obvious by Arpaly's and Markovits' formulations of RMF. However, Arpaly does address this issue in later work.¹⁰

On the assumption that defenders of RMF can draw some kind of distinction between right-makers that are intrinsically morally significant and those that are not, they can at least rule out Jack's motivation in **Venom** as conferring moral worth. But what about cases where someone is intrinsically interested in something intrinsically morally significant, yet fails to conceive of it as such? For example, imagine that Jack's fundamental motivating reason for making the incision is that it will save the stranger's life (call this **Venom***). But this is not because he sees anything particularly morally significant about saving lives. He has no concern for doing the right thing. Rather, he is simply intrinsically interested in saving lives in the same way he was intrinsically interested in draining venom in **Venom**.

Admittedly, **Venom*** doesn't elicit quite the same reaction as **Venom**. Acting from an intrinsic interest in draining venom strikes us as creepy, while acting from an intrinsic interest in saving lives does not. Crucially, however, the two cases are exactly morally equivalent from Jack's own point of view: in both cases, from his perspective, there is nothing morally significant about what he's doing. If Jack doesn't see what motivates him as morally significant in either version of the case, why should it be to his credit in **Venom*** that what he's motivated by just *happens* to be morally significant? The fact that the language of happenstance is no less apt in **Venom*** than in **Venom** suggests that this is still a case of accidentally doing the right thing. While (all else equal) we should prefer Jack's motivation in **Venom*** to his motivation in **Venom**, we should also be skeptical that Jack deserves any more credit for doing the right thing *per se* in **Venom*** than he does in **Venom**.

Though the two cases don't elicit the same reaction, this can be fully explained by the fact that intrinsic interest in draining venom is off-putting in a way intrinsic interest in saving lives is not. If Jack is an amoralist in both cases, then surely his action lacks moral worth in both cases. After all, if anyone's actions lack moral worth, it's those of the amoralist. The amoralist doesn't consider what motivates him morally significant and has no particular concern for whether it is. Without such features, his

¹⁰ See Arpaly and Schroeder (2014).

motivation by any particular consideration, whether or not it in fact happens to be morally significant, lacks the relevant connection to the rightness of his action.¹¹

Defenders of RMF might simply dig in their heels in response to **Venom***. But we can further bring out RMF's inadequacy by considering cases where there are multiple morally significant considerations at play. For example:

Hard Truth. Anna has a choice: either tell her friend a hard truth, or lie and spare her feelings. Anna cares deeply and intrinsically about telling the truth, but not at all about sparing her friend's feelings. So, without deliberating, she decides to tell her friend the hard truth. As it turns out, given the balance of moral reasons in this situation, this is the right thing to do.

In this case, it turns out that telling her friend something painful to hear is the right thing to do, and what makes it right is that it's the truth. Moreover, she tells her friend the hard truth for the reason that it's the truth. Nevertheless, Anna doesn't deserve credit for the rightness of her action. For while she does the right thing on the basis of a right-making feature, there is something strange about the fact that Anna has absolutely no interest in another clearly morally significant feature of the situation: that telling the truth will cause her friend significant emotional pain.

What's missing Anna's case is a concern for doing what the balance of moral reasons favors. While it's true that she ends up doing what's favored by the balance of moral reasons, this plays no role in the explanation of her action. If it did, she would have weighed the moral importance of truthfulness against her friend's well-being to figure out which choice was right. It's precisely the absence of such weighing that prevents her from deserving credit for doing the right thing in this case.

This looks like another case of accidentally doing the right thing. But in this case, the accident is that Anna's action just happened to be what the balance of morally reasons favored, despite her only being interested in one kind of morally relevant consideration. By contrast, if she were to weigh the morally significant considerations and then, having come to the correct conclusion, perform the very same

¹¹ Johnson King (forthcoming a) argues on similar grounds that Huck Finn's actions lack moral worth. As she and many others understand the case of Huck Finn, Huck just *happens* to be motivated by morally relevant considerations. However, I think there is a better interpretation of Huck's psychology that makes sense of why his actions are intuitively morally worthy. I discuss this in §6.3.

action, that action might have moral worth. But as it stands, despite satisfying RMF's conditions, this doesn't look like a case of morally worthy action.

Of course, it can't be the case that actions lack moral worth unless the agent weighs every single morally relevant consideration. If an agent fails to take into account very weak and/or massively outweighed considerations, her action can still be morally worthy. Nevertheless, it's clear that morally worthy action requires enough concern for doing what the balance of moral reasons favors that the agent will thereby take into account a wide range of weighty moral reasons.

The cases I've discussed in this section show that RMF leaves room for cases in which someone acts rightly on the basis of right-making features, but nevertheless fails to be creditworthy for doing the right thing. As such, RMF fails to guarantee the kind of non-accidentality required for moral worth.

4. The Rightness Itself View

According to the Rightness Itself View (hereafter RI), morally worthy action is motivated by rightness itself. Defenders of this view include Paulina Sliwa (2016) and Zoe Johnson King (forthcoming a). Different versions of the view differ on the sufficient conditions for moral worth, but all of them agree that motivation by rightness itself is a necessary condition. Furthermore, they all agree that motivation by right-making features is *not* a necessary condition for moral worth.

RI holds that the content of worth-conferring motivations is exclusively the rightness of the action itself, as opposed to its right-making features. So, like RMF, RI assumes that the relevant content must be exclusively either right-making features or rightness itself. Also like RMF, RI successfully rules out obvious counterexamples involving selfish motivation to do the right thing. The self-interested shopkeeper's actions lack moral worth because though it's right for him to price his goods fairly, he isn't at all motivated to do so by the consideration that it's right.

RI also makes sense of something that RMF struggles with. In the cases discussed in §3, part of what was missing was a kind of concern on the part of the agent for whether she was actually doing the right thing. RI captures the importance of this kind of concern by requiring that agents care about the

rightness of their actions itself.¹² Moreover, there's considerable intuitive pull to the idea that when agents perform morally worthy actions, the rightness of the action must at least partially explain why they performed it.

Defenders of RI take it to be uniquely well-placed to satisfy the non-accidentality condition. In the course of defending their respective versions of RI, both Sliwa and Johnson King argue that RMF cannot satisfy it. As a reminder, it seems to be a shared assumption in the debate between RMF and RI that the kind of motivation identified by each view is sufficient to secure non-accidentality. But even if Sliwa's and Johnson King's views secure non-accidentality, this can't be in virtue of being versions of RI. This is because the kind of motivation identified by RI fails to secure non-accidentality.

Consider a simple version of RI on which morally worthy action is simply right action motivated by the consideration that it's right. This simple version of RI requires only a match between the rightness of the action and the content of the agent's motivation. But this match can itself be merely accidental. To see this, consider cases where the agent is mistaken enough about the nature of moral rightness that the connection between the rightness of her action and her being motivated to perform is too tenuous for her to deserve moral credit, but not so mistaken that the property of rightness cannot figure in the content of her motivations. For example, consider the following case:

Moving. Simon's friend is in a tough spot and needs last minute help moving. Simon helps him, which is the right thing to do. And he's motivated to do so by the consideration that it's the right thing to do. According to Simon's conception of morality, what's right is what benefits one's friends and harms one's enemies. But the content of his motivation is simply that helping his friend is the right thing to do.

In this case, Simon does the right thing, and he's motivated by rightness itself in the sense that the content of his motivation is the consideration that his action is right. Nevertheless, due to his mistaken conception of morality, the connection between the content of Simon's motivation and the rightness of his action is too tenuous for him to deserve credit for its rightness. In the relevant sense, it's a mere accident that his motivation by rightness (as he conceives of it) results in his doing the right thing.

¹² Of course, some would see this as a vice of RI, not a virtue, because they take any form of so-called 'de dicto moral motivation' to be an objectionable kind of 'moral fetishism' (see e.g. Smith (1994)). I address moral fetishism in §6.2.

To further bring this out, it is worth noting that **Moving** can easily be modified into a version of **Hard Truth** that serves as a counterexample to RI. We can imagine a version of **Moving** in which there are other morally significant considerations that are outweighed, but which are not morally significant on Simon's conception of morality. In such a scenario, just like Anna's in **Hard Truth**, Simon's motivation would fail to manifest sensitivity to the balance of moral reasons. This would introduce an objectionable form of accidentality. So, insofar as RI allows for Simon's action to count as morally worthy in either version of **Moving**, it seems no less subject to counterexample than RMF.

The simple version of RI therefore fails to secure non-accidentality. This makes it clear that what RI says about motivational content isn't what does the work in satisfying the non-accidentality condition.¹³ Of course, neither Sliwa nor Johnson King endorses the simple version of RI. Each of them defends conditions for morally worthy action that outstrip RI. According to Sliwa, morally worthy action must be motivated by moral knowledge. According to Johnson King, morally worthy action must be an instance of deliberately doing the right thing. Both Sliwa and Johnson King write as if their respective conditions are compatible only with what RI says about motivational content, which explains why they take RI to be uniquely placed to secure non-accidentality.

I think this is a mistake. Nothing either of them says establishes any necessary connection between RI and moral knowledge or deliberateness. Instead, they each seem to argue against RMF on the grounds that it fails to satisfy the non-accidentality condition, then argue that their souped-up versions of RI succeed at satisfying it. But neither considers whether similar machinery is available to RMF.

First, consider moral knowledge. Sliwa's full view is that "[a] morally right action has moral worth if and only if it is motivated by concern for doing what's right (conative requirement) and by knowledge that it is the right thing to do (knowledge requirement)" (p. 394). She then argues that the knowledge requirement secures non-accidentality. But why couldn't defenders of RMF appeal to similar machinery? They could argue that a right action has moral worth if and only if it is motivated by (a)

¹³ Sliwa herself recognizes this when she gives the example of "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" (p. 396).

its right-making features, and (b) knowledge of their status as right-making features. It's not clear that this version of RMF would fare any worse than Sliwa's version of RI.

Now consider deliberateness. Johnson King's full view is that "[a]n act has moral worth just in case it is an instance of someone's deliberately doing the right thing" (p. 16). As she notes, "The terms 'deliberate' and 'accidental' are antonyms, and the categories to which they refer are logical contraries; if someone does something deliberately, then she does not do it accidentally, and vice versa" (pp. 16-17). This makes it trivially true that her view satisfies the non-accidentality condition.

Johnson King takes her view to be a version of RI. So, she must think that deliberately doing the right thing consists uniquely in being motivated by rightness itself. If she were to show this, she would thereby show that non-accidentality is connected to motivation by rightness itself and that RI uniquely satisfies the non-accidentality condition. But she doesn't show this, because she doesn't defend any particular account of deliberateness. This leaves it no clearer that deliberateness is uniquely connected to motivation by rightness itself than it was at the outset that non-accidentality is uniquely connected to motivation by rightness itself.

The present point is not about the merits of Sliwa's or Johnson King's accounts (I'll return to that subject in §6.1). Rather, it's this: the dialectic presupposes that motivational content provides the connection to rightness that satisfies the non-accidentality condition. In the context of this dialectic, our aim should be to figure out what motivational content could secure non-accidentality. This isn't an aim we can achieve by positing conditions that are independent of motivational content.

Moreover, regardless of the dialectic, positing such conditions doesn't help us answer what is independently a very important question about moral worth: what kind of motivation confers moral worth on an action? Each view claims that one motivational content exclusively confers moral worth. According to RMF, it's right-making features. According to RI, it's rightness itself.

What I hope to have shown so far is that neither content by itself secures non-accidentality. As such, we should reject the assumption in the debate that the right kind of motivation consists exclusively in either motivation by right-making features, or motivation by rightness itself. The failure of each of these options to provide a motivational structure that secures non-accidentality shows that this central

assumption gives rise to a false dilemma. In the following section, I'll discard this assumption and develop an alternative, intermediary view.

5. The Guise of Moral Reasons

Despite my criticisms of RMF and RI, I don't want to pretend that they haven't gotten anything right. In fact, both offer important insights about the nature of morally worthy action. I'll develop my account partly by drawing upon these insights.

The central insight of RMF is that morally worthy actions are performed on the basis of moral reasons – considerations that count in favor of them, morally speaking. In the clearest cases of morally worthy action, the agent appreciates what the balance of moral reasons favors, and does the right thing on the basis of those reasons. So, it seems right to hold that moral reasons, insofar as they serve as bases for morally worthy actions, are part of the content of worth-conferring motivations.

RMF gets this much right. But as shown in §3, it makes two mistakes. The first is thinking that these considerations are the whole of the content of such motivation. Relatedly, the second is thinking that the relevant kind of responsiveness to moral reasons can occur without the agent conceiving of them as such, or having any particular concern for their moral significance.

The central insight of RI, on the other hand, is that that morally worthy actions must be done out of concern for doing the right thing. Absent any kind of concern for whether their actions are right, agents simply don't deserve credit for doing the right thing *per se*, though their motivations might be otherwise laudable. Furthermore, in cases where there are multiple morally relevant considerations, the agent must appreciate the balance of such reasons and act accordingly for her action to have moral worth. It's difficult to make sense of how agents would be able to do this without weighing such considerations with regard to their moral significance. Some kind of concern for doing the right thing seems like a precondition for even engaging in moral deliberation.¹⁴

¹⁴ As such, I disagree with Arpaly (2015), who claims that moral deliberation doesn't require any concern for doing the right thing *per se*. For an extensive argument that moral deliberation does require such concern, see Johnson King (ms).

RI gets all of this right. But as §4 highlighted, concern for doing the right thing also can't be the entirety of the relevant motivational content. Besides, it's not at all clear that concern for doing the right thing ordinarily plays the same motivational role that right-making features do. I've suggested that concern for doing the right thing is what makes it possible for us to weigh and balance moral reasons with regard to their moral significance. But if this is the role that concern for doing the right thing plays, then it's a mistake to think that motivation by rightness itself is a competitor to motivation by right-making features. Instead, the former is a precondition of the latter's proper functioning.

The above point is of utmost importance, because it provides strong positive justification for exploring a motivational structure that has both right-making features *and* rightness itself as contents. If motivation by moral reasons and motivation by concern for doing the right thing are not only compatible, but necessary for each other's proper functioning, then the central insights of the two views can and should be combined. This is precisely what I'll do in formulating my own account.

Much of what I've argued so far suggests that morally worthy action requires motivation both by moral reasons *and* concern for doing the right thing.¹⁵ I've also shown that these two contents need to be related to each other in the right way. So, a simple conjunctive view won't do.¹⁶ Instead, we need a view on which the agent is motivated by moral reasons in virtue of her concern for doing the right thing. As I'll argue, such a structure is secured by a view on which the agent is motivated by moral reasons *as such*.

Here's a brief statement of my account:

GUISE OF MORAL REASONS: A right action has moral worth if and only if the agent performs it on the basis of sufficient moral reasons as such.

This requires some unpacking. What is it to be motivated by sufficient moral reasons *as such*? First, the agent must be motivated by sufficient moral reasons. Moral reasons are sufficient when they favor an

¹⁵ Toppinen (forthcoming) argues that this sort of view runs into problems unless we reject cognitivism about moral thought. However, his argument seems to rely on the assumption that both contents would need to play the role of motivating reasons. Since I reject this assumption, my account sidesteps the issue raised by Toppinen.

¹⁶ Though I don't have space to discuss this in detail, neither will a view involving primary and secondary motives, where secondary motives are limiting conditions that are activated only in cases where the agent's primary motives would lead to wrong action if left unchecked (see Baron (1995), Herman (1996), and Stratton-Lake (2000)).

action strongly enough to make it the case that the balance of moral reasons favors it. In other words, moral reasons are sufficient when they favor an action strongly enough to make it right. The language of sufficiency is important here because doing the right thing on the basis of insufficient moral reasons is a way of accidentally doing the right thing. For if the reasons for which the agent acted were the only moral reasons for her action, that action would have failed to be right.

Now we can tackle the ‘as such’ clause. This clause says that in order for an action to be morally worthy, the agent must not only act rightly on the basis of sufficient moral reasons, but act on the basis of those reasons under the guise of their being sufficient to make her action right. Importantly, to act for some reason under the guise of a moral reason is not simply to act on the basis of that reason and also to believe it to be a moral reason. Rather, to act for some reason under the guise of a moral reason is to be motivated by that reason in virtue of taking it to contribute to the overall moral status of the action. This is what it is to act for a moral reason as such.^{17,18}

Furthermore, in order to act on the basis of *sufficient* moral reasons as such, the agent must be motivated by these reasons in virtue of taking them to be sufficient – i.e. in virtue of taking them to favor the action strongly enough to make it the case that the balance of moral reasons favors it. This requires a kind of concern for what the balance of moral reasons favors. For without such a concern, an agent could not be motivated by some set of reasons in virtue of taking it to be sufficient to make it right.

It may be helpful to think of the agent’s being motivated by moral reasons in virtue of *attempting* to do the right thing. If I’m attempting to perform an action with a certain property, I must have some concern for whether my action actually has that property. So, if I’m attempting to do the right thing,

¹⁷ Acting for a moral reason as such thus entails that the agent is *not* motivated by the reason solely in virtue of taking it to favor the action in some other, nonmoral way – e.g. prudentially. So, my account correctly rules out the kind of instrumental motivation by moral reasons discussed by Markovits (2010).

¹⁸ It might be objected that the reason’s connection to the moral status of the action need not figure into the content of the agent’s motivation, because non-accidentally doing the right thing only requires manifesting a more general disposition to act on the basis of sufficient moral reasons. For examples of this sort of view, see Lord (2017) and Way (2018). I disagree with such views on the grounds that manifesting dispositions is insufficient for creditworthiness. For an argument that manifesting dispositions is insufficient for creditworthiness, see Singh (ms).

I must have some concern for what the balance of moral reasons favors – in other words, for what is morally right.

The motivational state identified by the Guise of Moral Reasons account thus includes both motivation by moral reasons *and* concern for doing the right thing. The former content captures the central insight of RMF: that morally worthy actions are performed on the basis of moral reasons. The latter content captures the central insight of RI: that morally worthy actions must be done out of concern for doing the right thing. But this latter content isn't itself a motivating reason. Instead, it's the guise under which the agent is motivated by her motivating reasons. Concern for doing the right thing thereby plays the role of *guiding* the agent's motivating reasons rather than being a motivating reason itself.

My account captures the central insights of both RMF and RI by making room in worth-conferring motivation for both of the contents they identify. In doing so, it provides the necessary resources to make sense of things that RMF and RI struggle with. One thing RMF struggles to make sense of is why actions lack moral worth when agents are only concerned with one of many morally relevant considerations. And while RI can explain how moral deliberation is possible, it also doesn't explain the importance of weighing and balancing different morally relevant considerations. If the consideration that an act is right plays the role of the agent's motivating reason instead of the higher-level guiding role I've suggested it should, it looks like she too fails to act out of concern for what the balance of moral reasons favors. My account makes sense of this. If someone is only concerned with one of many morally relevant considerations, she fails to count as acting for sufficient moral reasons *as such*.

The Guise of Moral Reasons account sits between the extremes of the two views that have come to dominate the debate over moral worth. Each of those views posits one kind of content as the sole content of worth-conferring motivations. But neither kind of content is sufficient. The Guise of Moral Reasons account solves this by carving out motivational roles for *both* kinds of content. The rightness of the action itself plays a motivational role by guiding the agent's motivating reasons, while the moral reasons that favor it play a role by constituting the agent's motivating reasons. In this way, the agent of a morally worthy action is motivated both by what makes the action right and the action's rightness itself.

This motivational structure gives the Guise of Moral Reasons account the resources to satisfy the non-accidentality condition. Here's how it deals with all of my counterexamples to RMF and RI. In cases like **Venom** and **Venom***, the agent isn't motivated by sufficient moral reasons as such, because he doesn't act under the guise of moral reasons at all. In cases like **Hard Truth**, the agent isn't motivated by sufficient moral reasons as such, because she lacks overall concern for what the balance of moral reasons favors. Finally, in cases like **Moving**, the agent isn't motivated by sufficient moral reasons at all; moreover, the rightness of his action plays the wrong role in his motivational structure.¹⁹

Unlike RMF and RI, the Guise of Moral Reasons account doesn't leave room for cases of accidentally doing the right thing. This is because in requiring that the agent be motivated by sufficient moral reasons under that very guise, it requires the agent's motivational structure to more closely mirror normative reality. The agent's concern for acting rightly mirrors the rightness of her action. The reasons for which she acts mirror the moral reasons that count in favor of it. And finally, the explanatory connection between her concern for acting rightly and the motivating force of moral reasons under that guise mirrors the explanatory connection between the rightness of her action and the moral reasons that favor it.

While this is a more complex motivational structure, I don't think this is a problem for the Guise of Moral Reasons account. Human motivation is complex, and we are constantly motivated in sophisticated ways. As such, there's nothing unrealistic about the idea that how we conceive of our actions plays a guiding role in which considerations motivate us to act. Less realistic are the kinds of simplistic motivation that RMF and RI take to be sufficient for morally worthy action. This explains why the counterexamples to these views involve characters with such odd psychologies. It's simply very strange that anyone would care only about whether an action is right, and not about the reasons why, or vice versa. By contrast, ordinary cases of moral motivation plausibly involve exactly the kind of complex structure I've identified. The complexity of the Guise of Moral Reasons account only makes it more plausible.

¹⁹ It might be objected that the fact that an action is right is a sufficient moral reason to perform that action. While this kind of view is often taken to lead to a variety of problems, Johnson King (forthcoming b) gives a plausible defense of it. I don't have the space to fully address such an objection here, but my response is basically that any way of being motivated by such a consideration that could confer moral worth would collapse into the kind of motivational structure identified by my account.

6. Loose Ends

In this section, I'll attempt to tie up some loose ends. First, I'll reconsider the details of Sliwa's and Johnson King's views in light of the Guise of Moral Reasons account. Then, I'll discuss what my account says about moral fetishism and the oft-discussed case of Huck Finn.

6.1. Moral Knowledge and Deliberateness

In §4, I temporarily set aside the details of Sliwa's and Johnson King's views in order to consider a simple version of RI. These details deserve further consideration. In particular, one might wonder what my account has to say about conditions on moral worth involving moral knowledge or deliberateness.

On Sliwa's account, morally worthy actions are motivated by knowledge that what we are doing is right. On my account, morally worthy actions are motivated by sufficient moral reasons as such. Does this require knowledge that what we are doing is right? Much depends on how we understand the knowledge condition. As Sliwa understands it, this condition "crucially relies on the assumption that knowledge is a mental state that can play a causal role in producing actions" (p. 395). Because this is a controversial assumption that Sliwa doesn't defend, it seems preferable not to have to take it on.

Furthermore, there is a separate issue with Sliwa's knowledge condition on the basis of which Johnson King rejects it. Johnson King's argument is roughly as follows. In order to have knowledge that the action we're performing is right, we need to know that it's the action we're performing. In cases where we successfully attempt to do something right, but where our success was highly uncertain, our actions can have moral worth despite our not knowing we were doing the right thing. If Johnson King is right about this, then we don't need to know that what we're doing is right for our actions to have moral worth.²⁰

²⁰ Of course, one might think that in cases of uncertainty, the right thing to do *just is* to try the best we can. If we accept this view, then the idea that we must know that what we're doing is right becomes more plausible, because we can know what we're trying to do despite being uncertain of our success.

On Johnson King's account, an action is morally worthy if and only if the agent deliberately does the right thing. And as she notes, deliberateness and accidentality are contraries. If morally worthy action is non-accidentally right action, then of course her account must be correct. However, it's a mistake to think that deliberately doing the right thing consists partly in being motivated by the reason that the action is right. As I argued in §5, if a right action is motivated by sufficient moral reasons as such, this guarantees non-accidentality. Thus, it must also guarantee deliberateness.

Johnson King also says that "that the best way to develop this [deliberateness-based] view is to take the performance of an act with moral worth to be a kind of achievement: the achievement of someone's trying to act rightly and succeeding" (p. 16). The Guise of Moral Reasons account makes good sense of this. To try to act rightly and succeed *just is* to do the right thing in virtue of being motivated by sufficient moral reasons as such. So, there is ultimately not much on which Johnson King and I disagree. Our only real point of disagreement is that she takes the deliberateness condition to entail RI, whereas I take it to be simply a restatement of the non-accidentality condition and so not by itself informative about the nature of worth-conferring motivation.

6.2. *Moral Fetishism*

A charge often levelled against RI is that motivation by rightness itself is a kind of 'moral fetishism.'²¹ Since motivation by rightness itself plays a role in the Guise of Moral Reasons account, does this mean it inherits the moral fetishism objection from RI?

No. While there's something to the moral fetishism objection, its significance is highly exaggerated. Moral fetishism is supposed to consist in a kind of monomaniacal obsession with doing the right thing, or a narrow-minded focus on the rightness of actions instead of the features that make them right. Insofar as RI allows for such motivational structures to be worth-conferring, this does look like a problem for RI.

However, the moral fetishism objection is often posed as an objection to *any* motivation by rightness itself. This is a mistake. As Sigrún Svavarsdóttir puts it, "a concern for being moral should not be

²¹ This objection to RI originally comes from Smith (1994).

confused with a rigorous obsession with morality or a resistance to examine hard reflective questions about morality” (1999, p. 200). If someone cares only about whether her actions are right, and not at all about why, that does seem strangely fetishistic. But one can be concerned with rightness itself without such concern manifesting itself so monomaniacally.²²

On the Guise of Moral Reasons account, concern for rightness itself is only one part of worth-conferring motivation. Agents who perform morally worthy actions are motivated by a range of morally relevant considerations in virtue of taking them to be morally relevant. There’s nothing narrow-minded or monomaniacal about concern for rightness itself playing this guiding role. So, there’s nothing fetishistic about the kind of moral motivation required by the Guise of Moral Reasons Account.

It’s also worth reiterating that the motivational structure required by the Guise of Moral Reasons account looks generally like the less odd form of moral motivation. It’s utterly ordinary for agents to be concerned with doing the right thing, and to express this concern by acting on the basis of what they take to be morally relevant considerations. By contrast, either a singular focus on rightness itself *or* a singular focus on right-making features (conceived of apart from their connection to rightness) should strike us as both fetishistic and psychologically abnormal.

6.3. *Huck Finn*

One of the most central cases in the literature on moral worth has been the case of Huck Finn.²³ Now that the Guise of Moral Reasons account is on the table, it’s worth looking at what it says about whether Huck’s action has moral worth. The action in question is helping Jim, a fugitive slave, escape.

Helping Jim escape is obviously the right thing to do. But because Huck believes Jim is the rightful property of Miss Watson, he believes he’s acting wrongly by helping Jim escape. According to Arpaly and Markovits, Huck’s action nevertheless has moral worth, making it a counterexample to RI. They interpret Huck as acting on the basis of the right-making features of his action, without conceiving of

²² Johnson King (ms) makes a similar point in discussing moral fetishism.

²³ This case originally appears in Twain (1884). It was introduced into discussions of moral worth by Bennett (1974), and has since been discussed by nearly all parties to the debate over moral worth.

them as such, or of his action as right, in any way. As Arpaly puts it, “He does not have the belief that what he does is right *anywhere* in his head – this moral insight is exactly what eludes him” (2002, p. 229). Nevertheless, because implicitly recognizes Jim’s personhood, he acts on the basis of a right-making feature.

Both Sliwa and Johnson King argue that if this is really Huck’s psychology, his action lacks moral worth because he merely accidentally does the right thing. I agree with this conditional verdict; the way Huck’s psychology is interpreted in the debate, it’s close enough to Jack’s in **Venom** and **Venom*** that it’s a counterexample to RMF. On the Guise of Moral Reasons account, Huck’s action would lack moral worth because although he acts for a sufficient moral reason, he fails to act for that reason as a moral reason.

Nevertheless, there’s quite a bit of intuitive pull to the idea that Huck’s action has moral worth. Does this mean that the Guise of Moral Reasons account has counterintuitive results in this case? No, because the debate operates with an overly simplistic interpretation of Huck’s psychology. As such, when we make intuitive judgments about the case, we implicitly fill in Huck’s psychology to make it more plausible, which prevents our intuitions from being probative. Given a more plausible reconstruction of Huck’s psychology, my account is well-placed to explain why his action is morally worthy.

Arpaly and Markovits interpret Huck as implicitly recognizing Jim’s personhood, and acting on the basis of this right-making feature. But because they insist that Huck doesn’t recognize this right-making feature as such, they leave it completely inexplicable *why* Huck is motivated by considerations of Jim’s personhood. In doing so, they make Huck into a kind of personhood fetishist, whose psychology seems no less strange than that of the oft-derided moral fetishist.

Instead, I think the best explanation of why Huck is motivated by considerations of Jim’s personhood is that he tacitly takes them to be morally significant. If he can tacitly recognize Jim’s personhood, why not think he also tacitly recognizes it as a moral reason, and is motivated by it in virtue of this recognition? My suggestion is that Huck, in deciding to help Jim, makes what Alex Worsnip (2017) calls a “cryptonormative judgment” that helping Jim is the right thing to do: “a judgment that is

presented by the agent who makes it as non-normative (either generally or in some particular respect), but that is in fact normative (either generally or in that particular respect)” (p. 3).²⁴

Unlike the amoralist, Huck clearly cares about doing the right thing. We should be wary of completely divorcing this general concern from his motivation to help Jim. Despite the fact that he doesn’t explicitly present himself in this way, I think we should interpret Huck as motivated by considerations of Jim’s personhood in virtue of tacitly taking them to constitute sufficient moral reason to help him. If we interpret his motivational structure in this way, his action counts as morally worthy on the Guise of Moral Reasons account.

This isn’t an ad hoc maneuver to save my account from counterexample, because my interpretation of Huck’s psychology is independently more plausible than the standard interpretation. As I’ve argued, motivation solely by right-making features (not conceived of as such) and motivation solely by rightness itself are each not only fetishistic, but fail to reflect the structure of ordinary moral motivation. If we give Huck the benefit of doubt that, despite how he presents his motivation to help Jim, it has the complex structure of ordinary moral motivation, then my account doesn’t just accommodate the verdict that Huck’s action has moral worth – it actually does a better job of explaining this verdict.

7. Concluding Remarks

Here’s what I’ve argued in this paper. Neither of the two dominant views about moral worth delivers the kind of motivational structure necessary to satisfy the non-accidentality condition. This is because it’s a mistake to assume that a simple motivational state (whether it be motivation by right-making features or by rightness itself) can guarantee non-accidentality. Instead, morally worthy action requires a motivational structure in which both contents are present and related in the right way. The account I have offered, The Guise of Moral Reasons account, delivers just such a structure.

²⁴ Of course, I have not shown that this is the only plausible interpretation of Huck’s psychology. But I have provided some support for my suggestion that it is a more plausible interpretation than the received interpretation. Moreover, to show that my account doesn’t have counterintuitive results, I only need to show that my interpretation is at least as plausible as the received one, which I think I have done.

The Guise of Moral Reasons account holds that morally worthy action consists in right action motivated by sufficient moral reason *as such*. When an agent is motivated by sufficient moral reasons as such, she is motivated by moral reasons in virtue of recognizing that they favor an action strongly enough to make it the case that the balance of moral reasons favors it. The agent's motivation by right-making features thereby expresses her concern for the action's rightness itself. This satisfies the condition that when an agent's action has moral worth, she deserves credit for acting rightly, because it is no accident that she has done the right thing.

References

- Arpaly, Nomy (2002). Moral Worth. *Journal of Philosophy* 99 (5):223-245.
- (2015). Huckleberry Finn Revisited: Inverse Akrasia and Moral Ignorance". In Randolph Clarke Michael Mckenna & Angela M. Smith (eds.), *The Nature of Moral Responsibility*. New York, USA: Oxford University Press. pp. 141-156.
- Arpaly, Nomy & Schroeder, Timothy (2013). *In Praise of Desire*. Oxford University Press.
- Baron, Marcia (1995). *Kantian Ethics Almost Without Apology*. Cornell University Press.
- Bennett, Jonathan (1974). The Conscience of Huckleberry Finn. *Philosophy* 49(188): 123–134.
- Broome, John (2013). *Rationality Through Reasoning*. Wiley-Blackwell.
- Dishaw, Samuel (ms). The Puzzle of Lucky Right Action.
- Faraci, David (2019). Groundwork for an Explanationist Account of Epistemic Coincidence. *Philosophers' Imprint* 19 (4).
- Herman, Barbara (1996). *The Practice of Moral Judgment*. Harvard University Press.
- Johnson King, Zoë (forthcoming a). Accidentally Doing the Right Thing. *Philosophy and Phenomenological Research*.
- (forthcoming b). We Can Have Our Buck and Pass It, Too. In Russ Shafer-Landau (ed.), *Oxford Studies in Metaethics, Volume 14*. Oxford University Press.
- (ms). How To be a Moral Fetishist.
- Kant, Immanuel (1785). *Groundwork of the Metaphysics of Morals*. Edited by Mary Gregor and Jens Timmerman (2012). Cambridge University Press.
- Lord, Errol (2017). On the Intellectual Conditions for Responsibility: Acting for the Right Reasons, Conceptualization, and Credit. *Philosophy and Phenomenological Research* 95 (2):436-464.
- Markovits, Julia (2010). Acting for the right reasons. *The Philosophical Review* 119 (2):201-242.
- Singh, Keshav (ms). Rationality and Correctly Responding to Reasons.
- Sliwa, Paulina (2016). Moral Worth and Moral Knowledge. *Philosophy and Phenomenological Research* 93 (2):393-418.

- Smith, Michael (1994). *The Moral Problem*. Blackwell.
- Stratton-Lake, Philip (2000). *Kant, Duty, and Moral Worth*. Routledge.
- Svavarsdóttir, Sigrún (1999). Moral cognitivism and motivation. *Philosophical Review* 108 (2):161-219.
- Toppinen, Teemu (forthcoming). From Duty for the Right Reasons. In Mark Timmons (ed.), *Oxford Studies in Normative Ethics, Volume 9*. Oxford University Press.
- Twain, Mark (1884). *The adventures of Huckleberry Finn (Tom Sawyer's comrade) by Mark Twain*. New York & London.
- Way, Jonathan (2018). Creditworthiness and Matching Principles. In Mark Timmons (ed.), *Oxford Studies in Normative Ethics, Volume 7*. Oxford University Press.
- Worsnip, Alex (2017). Cryptonormative Judgments. *European Journal of Philosophy* 25 (1):3-24.