

# Moral Hedging and Responding to Reasons

Amelia Hicks

**This paper is forthcoming in *Pacific Philosophical Quarterly*.  
Please cite the published version.**

**Abstract:** In this paper, I argue that the fetishism objection to moral hedging fails. The objection rests on a *reasons-responsiveness account of moral worth*, according to which an action has moral worth only if the agent is responsive to moral reasons. However, by adopting a plausible theory of non-ideal moral reasons, one can endorse a reasons-responsiveness account of moral worth while maintaining that moral hedging is sometimes an appropriate response to moral uncertainty. Thus, the theory of moral worth upon which the fetishism objection relies does not, in fact, support that objection.

## Introduction

Sometimes, people are motivated to exercise *moral caution* when they are morally uncertain. For example, someone who is uncertain about the moral status of animals might provisionally adopt a vegan diet, given the possibility that animals have a moral status that is incompatible with permissibly eating them for fun. Someone who exercises moral caution as a result of their moral uncertainty engages in a type of deliberation that we can call *moral hedging*.

Some philosophers<sup>1</sup> have objected that moral hedging is morally problematic, because it requires ‘fetishistic’ de dicto moral motivation—that is, moral hedging requires that one desire to do the right thing *as such*. According to one plausible interpretation of this objection, actions that result from moral hedging lack *moral worth*, because de dicto moral motivation prevents one from acting for the right sorts of reasons. In this paper, I will argue that when we pair a ‘reasons-responsiveness’ account (or ‘quality of the will’ account) of moral worth with a plausible theory of non-ideal moral reasons, it turns out that certain types of de dicto moral motivation are not fetishistic. Thus, because the reasons-responsiveness account of moral worth permits certain types of de dicto moral concern—including those types of de dicto moral concern that are at work in real-life cases of moral hedging—the account does not support the fetishism objection to moral hedging. My argument tells us something interesting about moral hedging, but also highlights important resources available to the reasons-responsiveness account of moral worth.

---

<sup>1</sup>Harman 2015; Hedden 2016; Weatherson 2014.

Outside of the literature on moral uncertainty, there are important criticisms of fetishism-based objections to motivational externalism.<sup>2</sup> However, those criticisms do not specifically address the objection that *moral hedging* is fetishistic.<sup>3</sup> The goal of this paper is to illustrate that even if one thinks that those earlier criticisms fail—that is, even if one still finds fetishism-based objections to motivational externalism compelling—one should not think that moral hedging necessarily (or even typically) involves fetishistic motivation.

In the literature on moral uncertainty, some philosophers have already replied to fetishism-based arguments against moral hedging, but I won't rehearse those replies here.<sup>4</sup> Those objections rely on rejecting the reasons-responsiveness account of the moral worth of actions;<sup>5</sup> and to the extent that they *don't* rely on rejecting a reasons-responsiveness account, they still leave unanswered the question of exactly how someone can endorse a reasons-responsiveness account of moral worth while also believing that moral hedging is sometimes morally appropriate. My goal, then, is to clarify how someone who accepts a reasons-responsiveness account of moral worth can at the same time recommend moral hedging under some circumstances.

Thus, my primary aim is to demonstrate that moral hedging can have moral worth. In section 1, I briefly clarify four related concepts that play important roles in this paper: *moral uncertainty*, *de dicto moral concern*, *de re moral concern*, and *moral hedging*. In section 2, I motivate my focus on a particular version of the fetishism objection, namely, a version of the objection that's rooted in a reasons-responsiveness account of moral worth. In section 3, I draw on work by Nomy Arpaly and Julia Markovits to more clearly characterize this version of the fetishism objection. In section 4, I argue that when we supplement the reasons-responsiveness account of moral worth with a plausible theory of *non-ideal moral reasons*, we get the result that moral hedging can yield morally worthy actions. In section 5, I respond to several objections to my argument. I conclude that the fetishism objection to moral hedging rests on an impoverished view of moral reasons.

## 1 Preliminaries

### 1.1 Moral Uncertainty

Moral uncertainty is uncertainty about moral propositions—we can say that one is morally uncertain when one's credences are split between mutually exclusive moral propositions. There are two types of moral uncertainty: *descriptively-based* and *morally-based*. When I experience descriptively-based moral uncertainty, my credences are split between mutually exclusive moral propositions because of prior uncertainty about descriptive propositions. (For example: I might be uncertain about whether it's morally permissible for me to shoot at a target, but only because I'm uncertain about whether anyone is dangerously close to the target.) When I experience morally-based moral uncertainty, my credences are split between mutually exclusive moral propositions *either* because of prior uncertainty about other moral propositions, *or* for no further reason at all. (For example, I might be uncertain about

---

<sup>2</sup>Carbonell 2013; Copp 1997; Lillehammer 1997; Olson 2002; Svavarsdottir 1999.

<sup>3</sup>With the exception of Carbonell, who briefly discusses cases of uncertainty in section 3.4 of her paper.

<sup>4</sup>Aboodi 2017; Sepielli 2016.

<sup>5</sup>See Section 5 of Sepielli 2016.

whether it's permissible for me to lie to someone because I'm uncertain about whether consequentialism is true. Alternatively, I might simply be uncertain about whether it's permissible for me to lie in this particular case.) This paper focuses on morally-based moral uncertainty.

## 1.2 De Re and De Dicto Moral Motivation

One desires to perform the right<sup>6</sup> action *de re* if and only if (1) one desires to perform some action, A, and (2) A is, in fact, the right action. That is, one desires to perform the right action *de re* when two conditions are met: when one has a certain kind of propositional attitude about an action, and at the same time that action has the property of being right. Notice that the fact that the action has that property need not be contained within the scope of the agent's propositional attitude.<sup>7</sup> I can desire to perform the right action *de re* without even considering whether the action I want to perform is right. We can say that one desires to perform the right action *de re* when it's true of the right action that one desires to perform it.<sup>8</sup>

One desires to perform the right action *de dicto* if and only if one desires *that* one performs the right action. That is, one desires to perform the right action *de dicto* when one desires to perform the right action *under the description* 'the right action.' In the case of *de dicto* moral desire, the property of the action—namely, *rightness*—is contained within the scope of the agent's propositional attitude. One can desire to perform the right action *de dicto* without knowing which action is, in fact, right. Consider these two examples:

**De Dicto:** Assume that one is morally obligated to give one's extra money to charity. Albert wants to satisfy his moral obligations, but isn't sure whether he's morally obligated to give his extra money to charity. He would prefer to keep his extra money and spend it on himself.

**De Re:** Assume that one is morally obligated to give one's extra money to charity. Alicia doesn't think much about her moral obligations (at least not in those terms), but she wants to give her extra money to charity. When asked why she's so charitable, Alicia says things like, 'This is just what I want to do.'

Albert provides us with an example of someone who wants to satisfy their moral obligations *only in the de dicto sense*; Alicia provides us with an example of someone who wants to satisfy their moral obligations *only in the de re sense*.

---

<sup>6</sup>I intend to use the word 'right' broadly, because one can have desires concerning many different kinds of moral properties.

<sup>7</sup>I speak loosely about properties and facts being contained within the scope of propositional attitudes; I do not mean to commit myself to any particular ontology of propositions.

<sup>8</sup>My definition of *de re* moral desire leaves open the possibility that an agent could desire to perform the right action *de re* when the agent (a) desires to perform the right action *de dicto* and (b) successfully identifies which action is right. If we want to ensure that *de re* moral desire never rests on a prior *de dicto* moral desire, then we must add a third condition: (3) the agent satisfies (1) and (2), but not because of a mere *de dicto* desire that successfully 'attaches' to the action that is, in fact, right. I thank John Brunero for helping me clarify this point.

### 1.3 Moral Hedging

When one engages in *moral hedging*<sup>9</sup> in response to moral uncertainty, one opts for what one takes to be a ‘morally safe’ course of action.<sup>10</sup> What’s distinctive about moral hedging is that one chooses what to do under conditions of morally-based moral uncertainty, while allowing one’s uncertainty to affect one’s choice. Moral hedging is distinct from simply acting on whatever moral propositions one has the most confidence in, and it’s distinct from acting without considering the fact that one is morally uncertain. Moral hedging could, on occasion, be characterized by dominance-style reasoning.<sup>11</sup> But one need not employ dominance reasoning in order to engage in moral hedging. For example, perhaps I think that neither omnivorism nor vegetarianism dominates the other with respect to moral permissibility; still, I could judge that the probability of vegetarianism being impermissible is lower than the probability of omnivorism being impermissible. Alternatively, I could judge that if vegetarianism were impermissible, it wouldn’t be as morally bad as omnivorism would be if omnivorism were impermissible.<sup>12</sup> In either case, I might reasonably conclude that vegetarianism is the more morally safe option, even if I remain uncertain about whether it is ‘objectively’ required.<sup>13</sup>

## 2 The Fetishism Objection

### 2.1 Weatherson’s Presentation of the Fetishism Objection

According to some philosophers, moral hedging is ‘fetishistic’ because it relies on de dicto moral motivation.<sup>14</sup> For example, in ‘Running Risks Morally,’ Brian Weatherson defends normative externalism—the view that ‘[t]he most important norms concerning the guidance and evaluation of action and belief are external to the agent being guided and evaluated,’<sup>15</sup> in the sense that those norms are unrelated to an agent’s evidence and credences—by responding to the objection that normative externalism cannot explain the wrong of ‘moral recklessness.’ To respond to this objection, Weatherson argues that there *is* no wrong of moral recklessness, and that there’s something morally problematic with the alternative to moral recklessness (namely, with moral hedging). Weatherson argues that in order for a hedging principle to be

---

<sup>9</sup>This isn’t a perfect term for what it’s being used to refer to. To ‘hedge one’s bets’ is to make multiple bets so that one gets some payoff in the event of different outcomes. As we’ll see, this is not quite the same as moral hedging.

<sup>10</sup>For an example of a hedging principle, see Guerrero 2007.

<sup>11</sup>For example: although I might be unsure of whether vegetarianism is actually required of me, I might still be convinced of vegetarianism’s moral permissibility while remaining unconvinced of omnivorism’s permissibility, and as a result opt for vegetarianism. This would be a case in which vegetarianism dominates omnivorism with respect to moral permissibility.

<sup>12</sup>I could even make judgments both about the probability of the practices being impermissible and about how morally bad the practices would be if they were impermissible, in which case I might attempt to maximize ‘expected moral value.’

<sup>13</sup>For further discussion of types of (and alternatives to) moral hedging, see Tarsney 2018; Hedden 2016; MacAskill 2016; Riedener 2015; Gustafsson and Torpman 2014; MacAskill 2013; Sepielli 2009; Ross 2006; Lockhart 2000.

<sup>14</sup>Weatherson 2014. Harman 2015 develops a similar objection to moral hedging, although hers does not appeal to fetishism in Weatherson’s sense. See footnote 23 of Harman 2015.

<sup>15</sup>Weatherson 2014, 141.

true, ‘moral fetishism’—an attitude that leads one to care about right action *as such*—would have to be acceptable.<sup>16</sup> However, moral fetishism isn’t acceptable, says Weatherson, because good people care about what’s morally valuable *de re*, not *de dicto*.

Weatherson begins by having us consider a case of prudential reasoning.<sup>17</sup> Relying on an extended example, he argues that it’s reasonable for one to care about the features that actually constitute one’s welfare, but *not* reasonable to care about one’s welfare *as such*, apart from its constituents. With this conclusion in place, Weatherson argues that the domain of moral reasoning is analogous; just as it’s perverse to care about one’s welfare as such, so too is it perverse to care about *right action* as such, apart from the constituents of right action. Weatherson writes,

A good person will dive into a river to rescue a drowning child. (Assuming that it is safe to do so; it’s wrong to create more rescue work for onlookers.) And she won’t do so because it’s the right thing to do. She’ll do it because there’s a child who needs to be rescued, and that child is valuable.<sup>18</sup>

The person who dives in *simply because it’s the right thing to do* is what Weatherson would call a ‘moral fetishist,’ someone who simply cares about doing the right thing without caring about the features of the action (and its context) that *make* the action right.

The person who takes moral uncertainty into account in the course of moral deliberation is, according to Weatherson, also a moral fetishist. Weatherson has us consider Martha, who prefers the steak, but would be satisfied with the tofu. Martha’s fairly confident that eating meat is permissible; however, she’s not certain. Moreover, in the imagined case, it’s in fact true that meat-eating is permissible. Weatherson writes,

Why should she turn down the steak? Not because she values the interests of the cow over her dining. She does not. And not because she should have that value. By hypothesis, she need not do so.... Rather, she has to care about morality as such. And that seems wrong.<sup>19</sup>

Thus, Weatherson appears to believe that one should not engage in moral hedging.<sup>20</sup> One can hedge when morally uncertain only by having *de dicto* moral motivation, and having *de dicto* moral motivation involves focusing inappropriately on a thin moral property<sup>21</sup> without any concern for the morally relevant features that determine the instantiation of that property.

---

<sup>16</sup>Weatherson is appealing to Smith’s notion of moral fetishism, which comes up in the context of the internalist/externalist debate in metaethics. See Smith 1994.

<sup>17</sup>Weatherson 2014, pp. 149-150.

<sup>18</sup>Weatherson 2014, 152.

<sup>19</sup>Weatherson 2014, 152.

<sup>20</sup>However, it isn’t obvious that we should attribute this view to Weatherson, because it isn’t clear what sorts of conclusions we can legitimately infer from the fetishism objection; see footnote 24.

<sup>21</sup>A thin moral concept is a moral concept that encodes no descriptive information; a thin moral property is a property that a thin moral concept expresses. *Right* and *good* are paradigmatic examples of thin moral concepts.

## 2.2 Interpreting the Fetishism Objection

We can interpret the fetishism objection in several ways. First, there are different ways in which we can characterize the ‘target’ of the objection:

- *choosing in accordance with* a hedging principle;
- *consulting* a hedging principle;
- *acting on the basis of a consultation of* a hedging principle.

A hedging principle is a principle that describes how one ought to engage in moral hedging.<sup>22</sup> One *chooses in accordance with a hedging principle* when one merely acts as a hedging principle would recommend, even if one doesn’t necessarily intend to follow the principle. However, we can safely ignore this first characterization of the target. Weatherson would not say that it’s fetishistic to  $\phi$  whenever a hedging principle recommends  $\phi$ -ing; if a hedging principle recommends that Martha order the tofu, but Martha orders the tofu simply because she prefers it, then her action isn’t fetishistic. Rather, Weatherson objects to ordering tofu out of a concern for rightness as such.

We can think of *the consultation of a hedging principle* as an act of deliberation, which involves applying a principle in order to see what result it delivers. For example, we can imagine Martha at the dinner table, trying to figure out what her preferred hedging principle would recommend.<sup>23</sup> One *acts on the basis of a consultation of a hedging principle* when one has already determined which action the principle recommends, and then performs that action. For example, we can imagine Martha has already figured out that her hedging principle recommends ordering the tofu, and thus she *orders the tofu*.

Weatherson clearly objects to acting on the basis of a consultation of a hedging principle; after all, the examples of fetishism that Weatherson provides are examples in which people perform actions (such as *diving into a river* or *ordering the tofu*) that are not mere acts of deliberation. However, Weatherson also clearly objects to having a concern for rightness as such, and it’s the consultation of a hedging principle that requires concern for rightness as such. So, Weatherson should *also* object to acts of deliberation in which one consults a hedging principle. Thus, Weatherson appears to think that there’s something morally problematic about (a) actions that are performed because they’re recommended by a hedging principle, but also about (b) acts of deliberation that reveal a concern for rightness as such, which will include cases in which one consults a hedging principle. When I use the phrase ‘moral hedging,’ I refer to a type of action that combines (a) and (b).

But what, according to Weatherson, is morally problematic about consulting a hedging principle and then acting on the basis of that consultation? Again, we have several possible interpretations available to us:

1. These sorts of actions require fetishistic moral motivation, and it’s *impermissible* to perform an action out of a fetishistic moral motivation. Thus, one should not hedge.

---

<sup>22</sup>For example, Guerrero’s ‘don’t know, don’t kill’ principle is a hedging principle. See Guerrero 2007.

<sup>23</sup>Note that in order to consult a hedging principle, one need not be aware that the principle one is consulting is a *hedging* principle. It simply needs to be true of a hedging principle that one consults it.

2. These sorts of actions require fetishistic moral motivation, and it's *vicious* (or at least *not virtuous*) to perform an action out of a fetishistic moral motivation. Thus, one should not hedge.
3. These sorts of actions require fetishistic moral motivation, fetishistic moral motivation necessarily involves a failure to properly respond to moral reasons, and actions that result from a failure to respond properly to moral reasons lack *moral worth*. Thus, one should not hedge.<sup>24</sup>

These three versions of the fetishism objection reflect the three main types of moral criticisms of actions; we can morally criticize an action by appealing to its *deontic status*, by appealing to the fact that it reflects a *vice* (or a lack of virtue), or by appealing to the fact that it lacks *moral worth*.<sup>25</sup> All three versions of the objection are interesting, and each has some claim to being Weatherson's intended criticism of fetishistic actions. However, I will examine only the third interpretation, for three reasons. First: I focus on the third interpretation instead of the second, because Ron Abodi has already persuasively addressed the second interpretation by arguing that it's sometimes vicious to *fail* to hedge.<sup>26</sup> Second: when discussing work that has influenced his thinking on this issue, Weatherson cites philosophers such as Nomy Arpaly and Elizabeth Harman, both of whom are concerned with the moral worth of actions (and both of whom endorse 'reasons-responsiveness' or 'quality of the will' views of moral worth). And third: the third version of the objection captures the intuitive idea that one should not make decisions by hedging, because hedging requires performing actions for the wrong sorts of reasons.<sup>27</sup>

The third interpretation of the objection requires an account of what the 'wrong sorts of reasons' are; fortunately, such an account is exactly what's supplied by the 'reasons-responsiveness' account (or 'quality of the will' account) of moral worth. In the following section, I describe the reasons-responsiveness account of moral worth in more detail, so that we can better understand this version of the fetishism objection.

---

<sup>24</sup>Notice that the fact that a particular type of motivation is 'fetishistic' doesn't immediately entail any action-guiding moral claims (such as 'One should not hedge'). However, either the fetishism objection is intended to support an action-guiding moral claim, or it is not. If it is, then these characterizations of the conclusion of the fetishism objection are accurate. If it isn't, then we have no practical reasons to take the objection seriously (that is, the objection is practically inert). So, I will assume that the fetishism objection is intended to support the claim that one should not hedge.

<sup>25</sup>Note that performing an action that lacks moral worth is distinct from being vicious, and from lacking virtue. See Section VI of Arpaly 2002a.

<sup>26</sup>Abodi 2017.

<sup>27</sup>Notice that the inference from 'performing that action would require a motivation that would rob the action of moral worth' to 'I won't perform that action' is itself somewhat fetishistic, since it reveals what seems like an excessive focus on the moral worth of one's actions. Analogous problems exist for every interpretation of the fetishism objection, which demonstrates that the objection either applies to itself or is practically inert. But I won't discuss this problem further. Thanks to an anonymous reviewer for mentioning this problem to me.

### 3 Moral Worth

Roughly, according to the reasons-responsiveness account of moral worth, an action has moral worth when it's performed with a good will, where a good will is one that responds properly to moral reasons. Moreover, the moral worth of actions tracks culpability; a morally worthy action is one for which the agent deserves praise, whereas an action that lacks moral worth is one for which the agent does not deserve praise (and perhaps deserves blame).

The *guiding idea* behind the reasons-responsiveness account of moral worth is that *in order for one's action to have moral worth, one's motivating reasons must coincide with the reasons that morally justify the action*. Julia Markovits endorses the guiding idea; the guiding idea is one direction of the bi-conditional in her 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.'<sup>28</sup> Nomy Arpaly also endorses the guiding idea; although Arpaly rejects the 'if' direction of Markovits' bi-conditional,<sup>29</sup> she accepts that an agent must be properly responsive to moral reasons in order for the agent's action to have moral worth.<sup>30</sup> I focus on the guiding idea because Arpaly and Markovits's work is widely accepted to be foundational in the literature on the reasons-responsiveness account of moral worth.<sup>31</sup> Thus, if we want to understand the fetishism objection to moral hedging in terms of the reasons-responsiveness view, we should focus on the guiding idea, because it's the commitment that Arpaly and Markovits share.

The guiding idea describes a necessary condition for an action to have moral worth; it says that an action lacks moral worth whenever one's motivating reasons *fail* to coincide with the reasons that morally justify the action. There are two ways in which one's motivating reasons can fail to coincide with the justifying reasons: (a) when one is unresponsive to the 'right sorts' of reasons, and (b) when one is responsive to the 'wrong sorts' of reasons.<sup>32</sup> Let's call the first type of failure *apathy*, and the second type of failure *zealotry*.

The 'moral worth' interpretation of the fetishism objection says that hedging involves either apathy or zealotry (and could even involve both at the same time). According to the objection, a hedger suffers from zealotry because they are motivated by the belief that an action is right;<sup>33</sup> focusing on the rightness of an action without any appreciation for the action's right-making features seems to involve an excessive focus on a mere *label*. But the hedger might also suffer from apathy—if I'm unsure of whether I should save a child in distress, and then I rescue the child only because I think that rescuing is 'most likely to be permissible,' then it looks as if I'm failing to respond to the reasons that actually support

---

<sup>28</sup>Markovits 2010, 205. This statement of the thesis is revised later in the essay, but that revision isn't important for our purposes.

<sup>29</sup>Arpaly's worry is that one's motivating reasons could coincide with the justifying reasons, and yet one might still perform the right action *accidentally*. See Markovits 2010, pp. 212-215.

<sup>30</sup>Arpaly, 2002a, 226.

<sup>31</sup>As well as work by other authors. See especially Ross 1930, and Stratton-Lake 2000.

<sup>32</sup>These two types of failures are suggested by Arpaly, who argues that being properly responsive to moral reasons requires (a) being motivated by whichever moral reasons are actually at play and (b) *not* being motivated by sinister reasons. Arpaly 2002a, 231-3.

<sup>33</sup>Or by the belief that an action is most likely to be right, or the belief that an action is most likely to be permissible, or the belief that an action maximizes expected moral value, etc.



rescuing (such as the child’s distress, or the importance of the child’s well-being.). My goal in the remainder of this paper is to show that moral hedging need not involve either type of failure. And by defending moral hedging from the charges of apathy and zealotry, I’ll in effect argue that certain actions performed out of de dicto moral motivation can have moral worth.<sup>34</sup>

To defend moral hedging from the charges of apathy and zealotry, I will develop a view about decision-making under moral uncertainty that’s compatible with the guiding idea of the reasons-responsiveness account of moral worth. Specifically, I will supplement the reasons-responsiveness account with an account of *non-ideal* moral reasons. The result will be that hedging need not involve zealotry or apathy, and is sometimes *necessary* in order to avoid apathy.

## 4 Responsiveness to Non-Ideal Reasons

To summarize so far: the fetishism objection to moral hedging alleges that when one hedges, one suffers from either *apathy* or *zealotry* (or both), and thus fails to be properly responsive to moral reasons. Since being properly responsive to moral reasons is necessary for one’s action to have moral worth (according to the *guiding idea*), moral hedging necessarily lacks moral worth.

In this section, I argue that we ought to adopt a more expansive view of moral reasons—a view that can ultimately vindicate moral hedging, even from a reasons-responsiveness perspective. On this view, moral uncertainty can itself provide one with a moral reason to hedge; in such cases, one responds properly to moral reasons by hedging, and thus satisfies the necessary condition expressed by the *guiding idea*.

My argument proceeds from a general characterization of *obstacles*. An obstacle, roughly, is a feature of the world that makes performing an action difficult or impossible. We can extract three important features from this rough characterization of obstacles. First, a feature of the world is only an obstacle *relative to some action*. Second, obstacles come in degrees, in the sense that different obstacles create different degrees of difficulty. And third, we can distinguish between different types of obstacles by looking at the *source* of the obstacle.

---

<sup>34</sup>One might wonder how my defense of the moral worth of some instances de dicto moral motivation is different from Markovits’ defense of certain types of de dicto moral motivation. According to Markovits, one cannot perform a morally worthy action on the basis of a mere true belief that the action is right, since such a belief might have an unreliable connection to the moral reasons that justify the action. However, if one has a *justified* true belief that an action is right, then that belief can itself be a reason that morally justifies some action (much like acting on the advice of a reliable authority figure) (Markovits 2010, pp. 218-223). According to Markovits, a justified true belief that my action is right can itself be a justifying moral reason, because our moral reasons are subjective, in the sense that they depend on what we have most epistemic reason to believe it would be morally best to do (Markovits 2010, p. 219). My account of the moral worth of some actions that proceed from de dicto moral motivation has two advantages over Markovits’. First, my account does not rest on the contentious claim that our moral reasons depend on what we have the most epistemic reason to believe it would be morally best to do. And second, my account can be extended to explain the moral worth of actions performed under *moral* uncertainty, not just descriptive uncertainty. Markovits only discusses cases in which one is descriptively uncertain and as a result performs an action out of de dicto moral motivation; it’s unclear how to extend her account so that it handles de dicto moral motivation under moral uncertainty.

For example, consider the following three obstacles to going to the grocery store: *roadwork*, *my lack of motivation to leave the house*, and *my inability to figure out where the grocery store is*. I will shortly discuss types of obstacles—in particular ‘agential’ and ‘non-agential’ obstacles—in more detail.

With this broad characterization of obstacles in place, I can argue that:

1. All types of obstacles to morally right action can provide us with *transitional* and *non-transitional* moral reasons to respond to those obstacles.

If my congressional representative makes it difficult for me (and other constituents) to lobby for more just laws, then I have a moral reason to change my representative’s mind, or to vote against the representative. If the hiring procedures adopted by the company I work for make it difficult to engage in fair hiring practices, then I have a moral reason to try to change those practices. In both of these examples, obstacles to right action—a corrupt representative, an unjust hiring structure—provide me with moral reasons to act in certain ways, moral reasons that I wouldn’t have were it not for those obstacles. But notice that there are two ways in which I can ‘respond to’ an obstacle, which suggests that there are two types of non-ideal moral reasons that are generated by obstacles. First, I can respond to an obstacle by *removing* it. For example, if the obstacle to right action is a corrupt representative, then I can try to remove that obstacle by trying to vote the representative out, or by trying to make them less corrupt. Second, I can respond to an obstacle by *changing tacks*. Perhaps there’s nothing that I can do about my corrupt representative; in that case, I can respond to that obstacle by looking for alternative strategies to lobby for more just laws. In the first type of case, the obstacle provides me with a ‘transitional’ moral reason; in the second sort of case, the obstacle provides me with a ‘non-transitional’ moral reason.<sup>35</sup> Transitional moral reasons are reasons to remove obstacles, so that we can pursue the best course of action. Non-transitional moral reasons, on the other hand, take for granted the continued existence of the obstacle and thus are reasons to *switch strategies* (and to instead pursue a ‘second-best’ strategy).

On this picture, when one aims to perform some action—which we can call the *target action*—there are sometimes different strategies available for performing that action, and those strategies can be ranked by their ideality. For example, let’s say that my target action is to promote more just laws. Perhaps *ideally* I would promote more just laws by voting out my corrupt representative (because, were I to successfully vote them out, I would have the best chance of promoting more just laws). We could then say that my corrupt representative provides me with a transitional moral reason to vote them out, because it would be best if they didn’t stand in the way of passing more just laws. But what if that strategy isn’t available to me, because I can’t do anything to vote out my corrupt representative? In that case, there might exist a second-best strategy for promoting more just laws, a strategy I can pursue even as my representative remains in office. A second-best strategy isn’t ideal, but it is nevertheless the best strategy available given the continued existence of the obstacle. So, an obstacle produces a transitional reason to remove that obstacle when the removal of the obstacle is the best available strategy for performing the target action; but the obstacle

---

<sup>35</sup>See Berg 2018 for an excellent discussion of transitional and non-transitional obligations. Transitional and non-transitional obligations are two types of ‘norms of compensation’ or ‘secondary norms’; for a nice discussion of secondary norms in ethics and epistemology, see DiPaolo, forthcoming.

produces only a non-transitional reason to work around that obstacle when the obstacle can't be removed.

It should now be clear that the way in which obstacles create transitional and non-transitional reasons is constrained by a version of *ought implies can*. Plausibly, one does not have moral reason to  $\phi$  unless it's possible for one to  $\phi$ . If we grant this version of *ought implies can*, it follows that an obstacle creates a transitional moral reason for one to remove the obstacle *only if the obstacle is removable*, and that an obstacle creates a non-transitional moral reason *only if there's a second-best strategy available*. One consequence of this constraint is that obstacles do not *always* create transitional or non-transitional reasons; sometimes obstacles can't be removed, and sometimes there's no way to work around them. For my purposes in this paper, it's important to notice that an obstacle's capacity to create transitional and non-transitional moral reasons doesn't directly depend on its *type*, that is, it doesn't depend on the source of the obstacle; rather, it depends on how difficult the obstacle makes the target action, and on whether the circumstances provide a second-best strategy.<sup>36</sup> So, just about any obstacle to right action can, in some context, create transitional and non-transitional moral reasons.

So far I've motivated my picture of target actions, obstacles, and non-ideal moral reasons by appealing to examples. But there are good theoretical reasons to think that obstacles give us transitional and non-transitional moral reasons. Moral reasons are reasons to *act*, and thus must be action-guiding (in some sense). In order for a reason to be action-guiding in any sense at all, it must be context-sensitive; that is, in order for moral reasons to support performing actions *that are actually available to one*, they must take into account one's circumstances, especially those features of one's circumstances that are relevant to the likely 'success' of one's action. Obstacles are features of one's circumstances that are relevant to the likely success of one's action. Thus, the moral reasons one has must depend on the obstacles one faces.

Some obstacles to right action do not come 'from outside,' but rather stem from one's own imperfections as a moral agent. We can call these *agential obstacles*. I hold that:

2. Agential obstacles—including moral uncertainty—are a type of obstacle to morally right action.

For example, vices—such as a tendency to behave selfishly—are obstacles to right action, because they sometimes prevent one from doing what one really ought to do. Similarly, a lack of knowledge (or at least a lack of true belief) can be an obstacle to right action; if I want to help my patient but do not know which type of medicine is safe for them, then my lack of knowledge makes it difficult for me to pursue the best course of action. So, character flaws and epistemic imperfections are examples of agential imperfections that create obstacles to right action. More importantly for my purposes, *moral uncertainty* can be an obstacle to

---

<sup>36</sup>In this paper, I don't want to commit myself to any particular view about the threshold of difficulty at which an obstacle creates non-transitional reasons instead of transitional ones. It might be that an obstacle creates a non-transitional reason when the obstacle is *impossible* (in some sense) to remove; but it might also be that an obstacle creates a non-transitional reason when the obstacle is *really difficult* (in some sense) to remove. I find the latter more plausible, but cannot pin down the exact degree of difficulty at which an obstacle creates a non-transitional rather than a transitional reason.

morally right action. Recall my characterization of an ‘obstacle’ from earlier: an obstacle is a feature of the world that makes performing an action difficult or impossible. Given this characterization, moral uncertainty counts as an obstacle to right action. For example, my uncertainty about whether it’s permissible for me to eat animals for fun can, under some circumstances, make it difficult for me to make the best choice about what to eat (it raises the likelihood that I will choose wrongly). One might object that moral uncertainty—such as my uncertainty about the permissibility of eating animals—does not create an obstacle to right action, because even if I’m uncertain about what it’s permissible to eat, it’s still true of the right action (whatever the right action turns out to be) that I’m perfectly capable of performing it. However, this is like saying that my uncertainty about the location of the grocery store isn’t an obstacle to getting to the grocery store, because it’s true of the path to the grocery store that I could drive along it. What my response to this objection illustrates is that obstacles—including agential obstacles such as moral uncertainty—to right action make it difficult to perform the right action *non-accidentally*. It’s true of the path to the grocery store that I could accidentally drive along it, and find myself at the store in spite of my uncertainty about the store’s location. Similarly, it’s true of the right action (such as *abstaining from meat*, or whatever the right action turns out to be) that I could perform it, and thereby perform the right action in spite of my uncertainty about its rightness. But both of these ‘successes’ would be accidental. One way to see this is to notice that in both cases (as I drive along the path, or as I abstain from meat) I perform the *target* action (going to the store, doing the right thing) without performing the target action under its description (namely, *going to the store*, and *doing the right thing*).<sup>37</sup> To summarize: agential obstacles, including moral uncertainty, count as obstacles to right action because they make it difficult (or impossible) to non-accidentally perform the right action. We might say that agential obstacles to right action make it difficult for one to perform the right action in the following sense: they make it difficult for one to perform the right action *through the exercise of one’s agency*.

Agential obstacles—like non-agential obstacles—can generate both transitional and non-transitional reasons for action. If I’m a selfish person, then I have a transitional reason to try to become less selfish; but if there’s nothing I can do about my selfishness (or if becoming less selfish will take a long time), then my selfishness can provide me with non-transitional reasons to act in certain ways. For example, if I know that my selfishness tends to manifest itself in a certain type of context, then my selfishness might provide me with a non-transitional reason to avoid that context, or with a non-transitional reason to ‘tie my own hands’ before I enter into that context. Moral uncertainty operates similarly; if I’m uncertain about the moral status of animals and as a result uncertain about whether I can permissibly eat them, then my uncertainty could provide me with a transitional moral reason to resolve my uncertainty. But if my uncertainty isn’t resolvable, then it provides me with a non-transitional moral reason to find some way of pursuing the target action (e.g., *acting permissibly*) in spite of my uncertainty.

If we grant that obstacles can generate non-transitional moral reasons, and that agential imperfections are a type of obstacle, then it follows that:

---

<sup>37</sup>For a full defense of the claim that one accidentally performs a target action of some type when one performs it under a description other than the target description, see Johnson King, forthcoming.

3. *Agential* imperfections (such as moral uncertainty) can generate non-transitional moral reasons, too.

Of course, these reasons will be ‘non-ideal’ in the sense that they are reasons that an agent wouldn’t have ideally; however, the reasons generated by *non*-agential obstacles are non-ideal in the same sense, and yet we’re inclined to treat them as genuine reason-generating features of the world.

So, we know that moral uncertainty is an obstacle that can generate non-ideal moral reasons, including non-transitional moral reasons. But now we must consider what moral uncertainty can provide us with a non-transitional moral reason *to do*. Although I can’t consider all possible actions that moral uncertainty might provide me with a non-transitional moral reason to do, I hold that:

4. If moral uncertainty can provide one with non-transitional moral reasons, then those reasons will sometimes be reasons *to hedge*.

In cases in which I want to act permissibly but my moral uncertainty can’t be resolved, there will sometimes be a second-best strategy available (that is, a strategy for trying to act permissibly that doesn’t require resolving my moral uncertainty); in such cases, my moral uncertainty provides me with a non-transitional moral reason to take up that strategy. Which strategy (and resulting action) is supported by a non-transitional moral reason depends on the options that the context makes available, and on which of those options is the best strategy for reaching the target action. As a result, non-transitional moral reasons generated by moral uncertainty will not *always* recommend a hedging strategy. (For example, let’s say that my moral uncertainty makes it difficult for me to decide whether to have an abortion. In this type of case, there might not be a ‘safe’ course of action that’s recommended by a hedging principle; perhaps, from my perspective, I run the risk of acting wrongly no matter what I choose. Still, my uncertainty provides me with a reason to pursue whatever strategy available to me that would make it most likely that I reach the target action.) But surely *sometimes* consulting and acting on a hedging principle is the best strategy available. In cases of unresolvable moral uncertainty, hedging will be the best strategy when the context makes available a ‘safe’ course of action.<sup>38</sup> Take the case in which I’m uncertain about whether it’s permissible to eat animals for fun, and then I must choose between eating them and abstaining from eating them. Assuming that I’m confident that abstaining from eating animals is permissible (which is a reasonable assumption in many contexts), adopting a hedging strategy is more likely to allow me to reach my target action (acting permissibly) than adopting a different strategy (such as choosing at random, or consulting a friend).

From the claims 3 and 4, we can infer that:

5. Moral uncertainty sometimes provides us with non-transitional moral reasons to hedge.

At this point, we must consider how this view—that moral uncertainty sometimes provides us with non-transitional moral reasons to hedge—is related to the account of *reasons-responsiveness* that theories of moral worth rely on. I hold that:

---

<sup>38</sup>Keep in mind that hedging can take different forms, and different forms of hedging will treat different courses of action as ‘safe.’ See Section 1.3.

6. If one's moral uncertainty provides one with a non-transitional moral reason to hedge and one hedges in response to that reason, then one is responsive to moral reasons.

One worry about claim 6 is that it asserts that I am responsive to moral reasons (in general) when I respond to one moral reason; but it's plausible that responding to a single moral reason isn't sufficient for being responsive to moral reasons in general. For example, *that it will make my child uncomfortable for a few months* is a moral reason to not seek surgery for my child; but of course that reason to avoid surgery could be outweighed by other moral reasons. So, in general, it isn't true that responding to one moral reason entails that one is responsive to moral reasons in general. However, a version of this claim that's restricted to the non-transitional reasons generated by moral uncertainty *is* true, and is true because of an interesting feature of non-transitional moral reasons: non-transitional moral reasons are *decisive, relative to a target action and a set of strategies*.

A first-order moral reason—such as the moral reason that supports abstaining from surgery—is not always decisive, because whether a first-order moral reason is decisive depends on all of the *other* moral reasons that are operative in the context.<sup>39</sup> We might even want to say that no first-order moral reason is itself 'decisive,' because what one ought to do is determined by all moral reasons that are operative in the context.

But non-transitional moral reasons, including non-transitional moral reasons to hedge, *are* decisive, relative to a target action along with a set of strategies for performing the target action. If I have to decide whether to eat meat, and I'm morally motivated but also uncertain about the moral status of animals, then my target action might be *to act permissibly*. I have a number of available strategies (for performing the target action) that I must choose between: *continue eating meat, beginning to abstain from meat, changing the types of meat I eat*, etc. Given the available strategies and that target action, there is a *ranking of strategies* for successfully performing the target action. So far, I've treated these strategies as being ranked by their likelihood of successfully leading to the target action; perhaps there is a different, better way of ranking them.<sup>40</sup> In any case, it's plausible that there is *some* ranking of strategies from best to worst; surely *consulting the most vicious person I know* is a worse strategy than *consulting the wisest person I know*. When one has a non-transitional moral reason to hedge, that's because a hedging strategy is the *best available* strategy for successfully performing the target action. But once we note that I have a non-transitional moral reason to hedge only when hedging is the best strategy available to me, it becomes much more plausible that that reason is *decisive*; after all, it's the type of reason I can have only if it supports the best available course of action (relative to a target action and

---

<sup>39</sup>I want to remain neutral on exactly how moral reasons combine to support particular actions; although I sometimes use a weighing metaphor to discuss the behavior of moral reasons, I do not want to commit myself to the view that reasons have 'weights' in a meaningful sense, or that we ought to perform that action that's supported by the 'balance' of moral reasons. But I will assume that moral reasons *somehow* collectively determine what we ought to do.

<sup>40</sup>I suspect that there is a better way of ranking them. If one simply ought to pursue whichever strategy has the highest subjective probability of producing the target action, then my view delivers a counterintuitive result in the miners' puzzle. However, this problem only shows that we need a more sophisticated way of ranking strategies for pursuing the target action, not that there is no ranking. See Kolodny and MacFarlane 2010 for a description of the miners' puzzle.

a set of strategies). We might say that it isn't the type of reason that can be 'outweighed' by other contextually salient reasons.

Now, notice that:

7. If one is responsive to moral reasons when acting, then one suffers from neither apathy nor zealotry.

This claim expresses the idea (from section 3) that avoiding both apathy and zealotry is necessary for one to be properly responsive to reasons. From claims 6 and 7, we can infer that:

8. If one's moral uncertainty provides one with a non-transitional moral reason to hedge and one hedges in response to that reason, then one suffers from neither apathy nor zealotry.

And from claims 5 and 8, we can reach the conclusion of my argument: *sometimes when one hedges in response to moral uncertainty, one suffers from neither apathy nor zealotry.*

Here is a summary of the foregoing argument:

1. All types of obstacles to morally right action can provide us with *transitional* and *non-transitional* moral reasons to respond to those obstacles.
  2. *Agential obstacles*—including moral uncertainty—are a type of obstacle to morally right action.
  3. So, agential obstacles—including moral uncertainty—can provide us with transitional and non-transitional moral reasons to respond to those obstacles. (1, 2)
  4. If moral uncertainty can provide one with non-transitional moral reasons, those reasons will sometimes be reasons *to hedge*.
  5. So, moral uncertainty sometimes provides us with non-transitional moral reasons to hedge. (3, 4)
  6. If one's moral uncertainty provides one with a non-transitional moral reason to hedge and one hedges in response to that reason, then one is responsive to moral reasons.
  7. If one is responsive to moral reasons when acting, then one suffers from neither apathy nor zealotry. (From the analysis of the *guiding idea* in Section 3)
  8. If one's moral uncertainty provides one with a non-transitional moral reason to hedge and one hedges in response to that reason, then one suffers from neither apathy nor zealotry. (6, 7)
- C. Therefore, sometimes when one hedges in response to moral uncertainty, one suffers from neither apathy nor zealotry. (5, 8)

The main result of this argument is that someone who hedges does not necessarily fail to satisfy an important necessary condition for their action to have moral worth. According to the fetishism objection, a hedger either suffers from apathy or zealotry, thereby robbing their action of moral worth; but I've argued that the hedger need not suffer from either. However, an interesting secondary result is that a morally uncertain person who has a non-transitional reason to hedge and *fails* to hedge in fact suffers from a form of apathy. I've argued that non-transitional moral reasons to hedge are genuine (albeit non-ideal) moral reasons. And if that's correct—if non-transitional moral reasons to hedge are genuine moral reasons—then *failing* to hedge can reflect an unresponsiveness to moral reasons, and thus can deprive one's actions of moral worth.

## 5 Objections

### 5.1 A hedger fails to respond to the first-order moral reasons that morally justify actions

One might object to premise six in the following way: when I respond to a non-transitional moral reason to hedge, I'm not responding to the *first-order* moral reasons that justify my action, and thus I'm not responding properly to moral reasons. Recall that the guiding idea behind the reasons-responsiveness account of moral worth is that one's action has moral worth only if one is responsive to the moral reasons that morally justify one's action. According to this objection, the only sorts of moral reasons that can morally justify actions are first-order moral reasons. For example, if I hedge and as a result choose to stop eating meat, the moral reasons that *actually* morally justify my action are reasons such as *the fact that animals are morally valuable and all else being equal, we shouldn't kill things that are morally valuable*. But when I hedge, I'm not responding to such reasons; instead, I'm responding to the non-transitional moral reason to hedge generated by my uncertainty. Thus, perhaps when one hedges (in response to a non-transitional moral reason), one nevertheless fails to be responsive to the reasons that in fact justify one's action, and thus it isn't accurate to describe one as 'responsive to moral reasons' in the way required by the reasons-responsiveness account of moral worth.

In response to this objection, I'd like to make three points. First, one can be responsive to a non-transitional moral reason to hedge and *at the same time* be responsive to first-order moral reasons. In fact, one might be responsive to a non-transitional moral reason to hedge *because* one is responsive to some first-order moral reasons. For example, when one hedges in response to one's uncertainty about the moral status of animals, one might still care about animals; this is because it's possible for one to care about animals without understanding *exactly* how valuable animals are (and without understanding one's *exact* obligations toward animals). Moreover, one's concern for animals can motivate one to hedge in response to one's uncertainty about animals' exact moral status. Thus, it looks as if one can hedge while remaining responsive to at least some first-order moral reasons.<sup>41</sup>

---

<sup>41</sup>This point is similar to a point in Zangwill 2003, pp. 148-9. Zangwill discusses the possibility of complex moral motivations: 'What Smith overlooks, quite generally, is the fact that when a person acts out of duty, the ordinary natural features of the act that impose the duty remain significant. Those natural properties



Second, this objection applies to one's use of *any* strategy for making a decision under moral uncertainty,<sup>42</sup> and thus seems to commit us to the view that unresolvable moral uncertainty makes morally worthy action impossible. According to the objection, I only manage to be responsive to moral reasons when I'm responsive to first-order moral reasons; but in cases of moral uncertainty, I'm uncertain of—and thus unresponsive to—at least some of those first-order moral reasons. (Although, as I just argued, I might still be responsive to *some* of them.) So, this objection does not say that there's anything especially bad about hedging; were I to adopt a different strategy for acting in light of my moral uncertainty—such as consulting a trusted friend and acting on their advice—I would be equally unresponsive to the first-order moral reasons. I find it implausible that I never have a moral reason to act on a friend's advice; and I find it equally implausible that I'm 'unresponsive' to moral reasons when I act on a friend's advice. What my response illustrates is that this objection entails that unresolvable moral uncertainty<sup>43</sup> makes morally worthy action impossible; the view expressed by the objection fails to make room for the possibility that it can be morally worthy for imperfect agents to acknowledge their imperfections and try to moderate their behavior in light of them.

Third, the objection assumes that only first-order moral reasons can actually morally justify actions; but that is an assumption I want to challenge in this paper. I hold that non-transitional moral reasons can morally justify actions. Non-transitional moral reasons are genuine moral reasons that apply to people who face immovable obstacles; they are the types of moral reasons that morally justify the actions of imperfect people in imperfect situations. One way to see this is to consider the ways in which non-transitional moral reasons are analogous to non-transitional *practical* reasons. Consider the person who learns that there's construction that makes it difficult to get to the grocery store, and who then decides to take a longer, alternate route to the store. Surely that person's action—taking a longer route to the grocery store—is practically justified because it's an appropriate response to the construction; the person has a non-transitional practical reason to take the long route, a reason that *practically justifies* taking the long route. Similarly, people can be morally justified in performing 'morally non-ideal' actions (such as hedging) because they have non-transitional moral reasons to perform those actions. Morally non-ideal actions are morally justified because they are the best actions available, given the obstacles faced by the agent.

## 5.2 Non-transitional moral reasons are not decisive

One might object to premise six in a different way: a non-transitional moral reason is not decisive, because it *can* be outweighed by other moral reasons. For example, perhaps instead

---

play a role in moral motivation because we take those natural properties to determine a moral property and we are motivated by the fact that the act, with its natural properties, has that moral property.... It is true that the agent is acting on the moral properties he possesses; but he possesses them in virtue of having an intrinsic concern with other's needs. Those needs do not motivationally disappear in the motive of duty. They are part of the motive of duty.' Sepielli 2016 also makes a similar point.

<sup>42</sup>The objection might even apply to cases in which one exercises caution in response to *descriptively*-based moral uncertainty, given a view of moral reasons on which descriptively-based moral uncertainty prevents one from being responsive to moral reasons.

<sup>43</sup>Or very-difficult-to-resolve moral uncertainty; see footnote 36.

of hedging in an attempt to act permissibly, I should stop thinking about whether to eat meat and instead focus on helping my child with her homework.

By way of response, notice that a non-transitional moral reason can be ‘outweighed’ by other reasons only when we shift the target action or the set of strategies I’m choosing between. Recall that a non-transitional moral reason is decisive *relative* to a target action and a set of available strategies. I might have a non-transitional moral reason to hedge relative (a) to the target action *acting permissibly* and (b) to the set of strategies *eating meat* and *abstaining from eating meat*. But if we increase the set of strategies to include *all* available actions (including actions that aren’t related to my moral uncertainty about the moral status of animals, actions such as *helping my child with her homework*), then I might not have a non-transitional moral reason to hedge relative to that new set of strategies. However, it would still remain the case that when it comes to the decision about whether or not to eat meat, I have a non-transitional moral reason to hedge, given that I aim to act permissibly.

### 5.3 The fetishism objection can be applied to the adoption of some target actions

So far, I’ve said very little about how the view I’ve developed handles de dicto moral motivation. As a result, one might object that I’ve failed to respond to the fetishism objection; after all, isn’t it fetishistic to adopt an action such as *acting permissibly* as one’s ‘target action’?

On my view, de dicto moral motivation—which can manifest itself in the adoption of target actions such as *acting permissibly* and *doing the right thing*—are compatible with morally worthy action. We can grant that under ideal circumstances an agent would not adopt a target action described using thin moral concepts; but in many cases of moral uncertainty, the agent’s target action must be described using thin moral concepts, because the agent’s moral uncertainty makes a more fine-grained description of the target action impossible. Nevertheless, de dicto moral motivation is compatible with performing morally worthy actions, for two reasons.

First, as I argued in 5.1, people are sometimes motivated to hedge *because they’re partially responsive to some first-order moral reasons*; a corollary of this observation is that people sometimes have a de dicto desire to *do the right thing* or to *act permissibly* because they’re responsive to some first-order moral reasons. For example, I can adopt the target action *acting permissibly (in choice of food)* because I care about animals, but am unsure of what I owe them. If responsiveness to some first-order moral reasons is sufficient for one’s action to have some degree of moral worth, and if one can be responsive to some first-order moral reasons while experiencing de dicto moral motivation, then it follows that de dicto moral motivation is compatible with one’s action having some degree of moral worth.

Second, one can have *non-ideal* moral reasons to be motivated by rightness de dicto. Recall that one can have a non-transitional moral reason to hedge, and recall that hedging usually requires one to adopt a target action described in terms of a thin moral concept. To adopt a target action described in terms of a thin moral concept is to experience de dicto moral motivation. Presumably, if one has a non-ideal moral reason to  $\phi$  and  $\phi$ -ing requires  $\psi$ -ing, then one has a non-ideal moral reason to  $\psi$ . Thus, since one sometimes has a non-ideal

moral reason to hedge and must be motivated by rightness de dicto in order to hedge, one sometimes has a non-ideal moral reason to be motivated by rightness de dicto. We can think of some non-ideal moral reasons to hedge as reasons to exercise epistemic humility when it comes to moral matters; given that de dicto moral motivation is often necessary to exercise some epistemic humility when it comes to moral matters, non-ideal moral reasons to hedge can also be reasons to be motivated by rightness de dicto.

One might object that even if I have a non-ideal moral reason to adopt a target action described in thin moral terms, the *action* that results from consulting a hedging principle still lacks moral worth because it (the action) doesn't reveal responsiveness to moral reasons. This objection raises the possibility that even if one's *consultation* of a hedging principle can be morally worthy, *acting* on that consultation cannot be. In response: when one deliberates and then comes to a conclusion about what to do, one's prior deliberation provides one with a reason to act on that conclusion. For example, if I abstain from meat as a result of consulting a hedging principle, I thereby respond to my reasons to act on my prior good faith deliberation. If I have good (non-ideal) moral reasons to deliberate in a certain type of way, then presumably those reasons support my *acting* on the conclusion of my deliberation.

And yet, one might still worry that there's something wrong with wanting to *do the right thing* when one doesn't know what the right thing is. Perhaps we can understand this worry by thinking through the following example. Imagine Anthony, who is trying to decide whether to set his cat on fire for fun, but he isn't sure whether it's the right thing to do. Anthony understands what setting his cat on fire will entail, non-morally speaking. Yet he still reports that he 'just wants to do the right thing,' whatever that turns out to be.

Clearly, there is something horribly wrong with Anthony's motivational structure. But what's wrong with Anthony isn't his de dicto moral motivation; rather, the problem is that there appear to be no constraints at all on Anthony's application of the concept *rightness*. To see this, compare Anthony to Fatima. Fatima is trying to decide whether to fail a student for plagiarism; she thinks failing the student would be fair, but might not be in the student's best interest. She reports that she 'just wants to do the right thing.' Anthony and Fatima do not have the same type of de dicto moral motivation; we might say that 'rightness' is not an empty concept for Fatima, even though it might be for Anthony. Anthony's concept of rightness could include anything, including setting his cat on fire for fun. Fatima's concept of rightness could not (I assume) include anything. Presumably, Fatima knows of some paradigm cases of right actions as well as some paradigm cases of wrong actions, and those paradigm cases place constraints on which actions could conceivably count as right. Although she doesn't have a full analysis of 'rightness,' it's not an empty concept for her.

Thus, we shouldn't conclude that there's always something inappropriate about de dicto moral motivation. De dicto moral motivation is compatible with being responsive to some first-order moral reasons, and is compatible with being responsive to non-ideal moral reasons (such as non-transitional moral reasons to hedge). Moreover, hypothetical cases in which agents exhibit an intuitively horrifying de dicto moral motivation at most show that there's something inappropriate about de dicto moral motivation *when it's paired with a completely unconstrained concept of rightness* (or an unconstrained concept of some other thin moral property). Real-life cases are not similar to such hypothetical cases, and thus we don't have good reason to think that actual instances of de dicto moral motivation are particularly worrisome.

## 5.4 Both the hedger and the non-hedger fail to respond to reasons, but the hedger's failure is worse

Imagine two people who are both uncertain about the moral status of animals; one hedges in response to their moral uncertainty, whereas the other person doesn't (they simply continue to eat what they've always eaten). One might think that *both* the hedger and the non-hedger suffer from some degree of apathy, because they're uncertain of (and thus unresponsive to) a number of first-order moral reasons. So, one might object that it's inaccurate for me to assert that a hedger can avoid apathy, because *any* failure of reasons-responsiveness (including failure to respond to first-order moral reasons) is sufficient for some degree of apathy.

Even if we grant that this objection is correct, I can revise my argument and instead conclude that the hedger's apathy, although admittedly not ideal, is nevertheless *better* than the non-hedger's apathy. But why, exactly, should we think that the hedger's apathy is better than the non-hedger's?

First, notice that the hedger is responsive to more moral reasons than the non-hedger. Both the hedger and the non-hedger fail to respond to a number of first-order moral reasons; however, the hedger manages to respond to a non-transitional moral reason to hedge. The hedger's responsiveness to a non-transitional moral reason is significant, because they respond to a moral reason that applies to them *given* their inability to respond to some first-order moral reasons. So, it seems that the hedger and non-hedger both fail to respond to a number of first-order moral reasons, while only the hedger manages to respond to some of the non-ideal moral reasons that apply to them because of that failure. The non-hedger, by contrast, must simply hope that they *accidentally* act in ways that the moral reasons support.

Perhaps one could object that the hedger's apathy is worse than the non-hedger's because the hedger's unresponsiveness to first-order moral reasons is accompanied by a de dicto desire to act permissibly. However, as I argued in 5.3, there's nothing intrinsically fetishistic about a de dicto concern for doing the right thing; de dicto moral concern is only intuitively horrifying when it's combined with an unconstrained concept of rightness.

Thus, the hedger responds to more reasons than the non-hedger, and the hedger does not necessarily have an objectionable de dicto moral desire. It is unclear, then, how one could maintain that the hedger's apathy is worse than the non-hedger's. In any case, my aim in this paper has been to advance an account of moral worth that is compatible with the hedger not suffering from any significant degree of apathy.

## Conclusion

I've argued that if we understand the fetishism objection to moral hedging as an objection that appeals to a reasons-responsiveness account of moral worth, then the objection says that moral hedging lacks moral worth because hedging involves (a) responding to the wrong sorts of reasons (zealotry) or (b) failing to respond to the right sorts of moral reasons (apathy). However, if we supplement the reasons-responsiveness account with a theory of non-ideal moral reasons—including non-transitional reasons that are generated by agential imperfections—then it turns out that hedging can in fact be an appropriate response to moral reasons. My argument demonstrates that this version of the fetishism objection relies

on an impoverished view about the types of moral reasons that we have. If we adopt a more plausible view of the types of moral reasons that we have, then it's no longer clear that there's anything inherently fetishistic about moral hedging.

However, I have not provided a full defense of the existence of non-ideal moral reasons (or of non-transitional moral reasons created by agential imperfections). We ought to do more work on this. Nevertheless, my argument shows that the most fundamental commitments of a reasons-responsiveness account of moral worth are in principle compatible with the existence of non-ideal moral reasons, and thus are compatible with morally worthy moral hedging.<sup>44</sup>

*Department of Philosophy*  
*Kansas State University*

## References

- Aboodi, R. (2017) 'One Thought Too Few: Where De Dicto Moral Motivation is Necessary,' *Ethical Theory and Moral Practice* 20, pp. 223-237.
- Arpaly, N. (2015) 'Huckleberry Finn Revisited: Inverse Akrasia and Moral Ignorance,' in R. Clarke, M. McKenna, and A. Smith (eds) *The Nature of Moral Responsibility: New Essays* (Oxford: Oxford University Press), pp. 141-156.
- Arpaly, N. (2002a) 'Moral Worth,' *The Journal of Moral Philosophy* 99, pp. 223-245.
- Arpaly, N. (2002b) *Unprincipled Virtue: An Inquiry Into Moral Agency* (Oxford: Oxford University Press).
- Berg, A. (2018) 'Ideal Theory and 'Ought Implies Can',' *Pacific Philosophical Quarterly* 99, pp. 869-890.
- Carbonell, V. (2013) 'De dicto desires and morality as fetish,' *Philosophical Studies* 163, pp. 459-477.
- Copp, D. (1997) 'Belief, Reason, and Motivation: Michael Smith's *The Moral Problem*,' *Ethics* 108, pp. 33-54.
- DiPaolo, J. (forthcoming) 'Second best epistemology: fallibility and normativity,' *Philosophical Studies*.
- Goldman, H. (1976) 'Dated Rightness and Moral Imperfection,' *Philosophical Review* 85, pp. 449-487.

---

<sup>44</sup>I've received excellent feedback on earlier versions of this paper from many people, including John Brunero, Josh DiPaolo, Amy Lara, Rosa Terlazzo, Jon Herington, Graham Leach-Krouse, Han Li, and Angela Curran. I also received valuable feedback from audiences at the Workshop on Population Ethics and Moral Uncertainty (at the Institute for Futures Studies), the 2016 Bled Ethics Conference, the 8th Annual Rocky Mountain Ethics Congress, and the Philosophy Department at the University of Kansas. I'd also like to extend special thanks to several anonymous reviewers for their insightful comments on earlier drafts of this paper.

- Greenspan, P.S. (1978) 'Oughts and Determinism: A Response to Goldman,' *Philosophical Review* 87, pp. 77-83.
- Guerrero, A. (2007) 'Don't Know, Don't Kill,' *Philosophical Studies* 136, pp. 59-97.
- Gustafsson, J. E. and O. Torpman, (2014) 'In Defence of My Favourite Theory,' *Pacific Philosophical Quarterly* 95, pp. 159-174.
- Harman, E. (2015) 'The Irrelevance of Moral Uncertainty,' in R. Shafer-Landau (ed) *Oxford Studies in Metaethics, Vol. 10* (New York: Oxford University Press).
- Harman, E. (2011) 'Does Moral Ignorance Exculpate?' *Ratio (new series)* 24, pp. 443-468.
- Hedden, B. (2016) 'Does MITE Make Right? On Decision-Making Under Normative Uncertainty,' in R. Shafer-Landau (ed) *Oxford Studies in Metaethics, Volume 11* (New York: Oxford University Press).
- Johnson King, Z. A. (forthcoming) 'Accidentally Doing the Right Thing,' *Philosophy and Phenomenological Research*.
- Kolodny, N. and MacFarlane, J. (2010) 'Ifs and Oughts,' *Journal of Philosophy* 107, pp. 115-143.
- Lillehammer, H. (1997) 'Smith on Moral Fetishism,' *Analysis* 57, pp. 187-195.
- Lockhart, T. (2000) *Moral Uncertainty and its Consequences* (New York: Oxford University Press).
- MacAskill, W. (2016) 'Normative Uncertainty as a Voting Problem,' *Mind* 125, pp. 967-1004.
- MacAskill, W. (2013) 'The Infectiousness of Nihilism,' *Ethics* 123, pp. 508-520.
- Markovits, J. (2010) 'Acting for the Right Reasons,' *Philosophical Review* 119, pp. 201-242.
- Riedener, S. (2015) *Maximizing Expected Value under Axiological Uncertainty*, Dissertation, University of Oxford.
- Ross, J. (2006) 'Rejecting Ethical Deflationism,' *Ethics* 116, pp. 742-68.
- Ross, W.D. (1930) *The Right and the Good* (Oxford: Oxford University Press).
- Sepielli, A. (2016) 'Moral Uncertainty and Fetishistic Motivation,' *Philosophical Studies* 173, pp. 2951-2968.
- Sepielli, A. (2009) 'What to Do When You Don't Know What to Do,' in R. Shafer-Landau (ed) *Oxford Studies in Metaethics, Volume 4* (New York: Oxford University Press).
- Smith, M. (1994) *The Moral Problem* (Malden, MA: Blackwell Publishing Ltd).
- Stratton-Lake, P. (2000) *Kant, Duty, and Moral Worth* (London: Routledge).

Svarsvarsdottir, S. (1999) 'Moral Cognitivism and Motivation,' *The Philosophical Review* 108, pp. 161-219.

Tarsney, C. (2018) 'Intertheoretic Value Comparisons: A Modest Proposal,' *Journal of Moral Philosophy* 13, pp. 324-344.

Weatherson, B. (2014) 'Running Risks Morally,' *Philosophical Studies* 167, pp. 141-163.

Zangwill, N. (2003) 'Externalist Moral Motivation,' *American Philosophical Quarterly* 40, pp. 143-154.