

# One Thought Too Few: Where *De Dicto* Moral Motivation is Necessary

Ron Aboodi<sup>1</sup> 

Accepted: 27 June 2016 / Published online: 24 August 2016  
© Springer Science+Business Media Dordrecht 2016

**Abstract** *De dicto moral motivation* is typically characterized by the agent’s conceiving of her goal in thin normative terms such as *to do what is right*. I argue that lacking an effective *de dicto* moral motivation (at least in a certain broad sense of this term) would put the agent in a bad position for responding in the morally-best manner (relative to her epistemic state) in a certain type of situations. Two central features of the relevant type of situations are (1) the appropriateness of the agent’s uncertainty concerning her underived moral values, and (2) the practical, moral importance of resolving this uncertainty. I argue that in some situations that are marked by these two features the most virtuous response is deciding to conduct a deep moral inquiry for a *de dicto* moral purpose. In such situations lacking an effective *de dicto* moral motivation would amount to a moral shortcoming. I show the implications for Michael Smith’s (1994) argument against Motivational Judgment Externalism and for Brian Weatherson’s (2014) argument against *avoiding moral recklessness*: both arguments rely on a depreciating view of *de dicto* moral motivation, and both fail; or so I argue.

**Keywords** *De dicto* desire · Fetishism · Moral motivation · Moral uncertainty · Moral inquiry · Thin concepts · Michael Smith · Brian Weatherson

Michael Smith (1994: 75), Brian Weatherson (2014: 152), and others have criticized the motivation *to do what is right* as such, where the person pursues this goal under this particular description. Their criticism extends to similar motivations where the person conceives of her goal in other *thin* normative terms, such as *to act in accordance with morality* or *to promote the good*. Following Smith and Weatherson, I use the term *de dicto moral motivation* to denote any motivation of this kind. According to Smith and Weatherson *de dicto* moral motivation is less morally virtuous than what Smith (1994: 75) calls “non-derivative *de re* moral motivation” — where the agent pursues a morally worthy goal under a non-normative

---

✉ Ron Aboodi  
ron.aboodi@mail.huji.ac.il

<sup>1</sup> Department of Philosophy and Federmann Center for the Study of Rationality, Hebrew University of Jerusalem, Mt. Scopus, 91905 Jerusalem, Israel

description, such as *to save this drowning baby*; or under a *thick* moral description, such as *to act honestly*.<sup>1</sup> Pursuing such goals constitutes *de re* moral motivation insofar as doing so is morally good or right or required (etc.), regardless of whether or not the agent recognizes or cares about the goodness or rightness (etc.) of doing so.<sup>2</sup> If the agent pursues such a goal for the reason that it is morally required, she *derives* (in Smith's terminology, as I understand it) this goal from her motivation to do what is morally required, which constitutes *de dicto* moral motivation. But *de re* moral motivation can be *non-derivative*, for instance in cases where the person wants to rescue the baby merely for the baby's sake, regardless of morality as such.

Surely, a morally virtuous<sup>3</sup> person does not have to think about morality before saving a drowning baby or helping her beloved. There is nothing wrong, in many cases, with aiming merely at saving the baby, or helping the beloved, under such a description. Non-derivative *de re* moral motivations suffice for performing many good deeds. Moreover, it is hard to think of a morally-required action that *cannot* be motivated merely by *some* non-derivative *de re* moral motivation. This raises the question: Is *de dicto* moral motivation necessary, in any sense, for morally virtuous behavior?

The existing literature lacks sufficiently developed arguments supporting the necessity of *de dicto* moral motivation. Most of its defenses seek to establish merely that in some cases it is forgivable, permissible or sufficiently virtuous to be motivated by *de dicto* moral motivation, not ruling out that it would have been equally or more virtuous in such cases to be motivated instead by *de re* moral motivation.<sup>4</sup> But I argue, in section 1, that certain situations call for deciding to conduct a deep moral inquiry<sup>5</sup> of a certain type, for a *de dicto* moral purpose (in a certain broad sense). In such situations, lacking an effective *de dicto* moral motivation would put the agent in a bad position for making the morally best choice (relative to her epistemic state), and would therefore prevent her from responding in the most virtuous manner.

Section 2 discusses some implications of this point, focusing on the debate concerning the nature of moral judgments (Smith 1994) and the debate concerning the rational way of coping with moral uncertainty (Weatherson 2014). Smith and Weatherson both advance arguments that rely on a depreciating view of *de dicto* moral motivation, and I argue that both of these arguments fail.

Section 3 briefly addresses the tension between the importance of *de dicto* moral motivation, for which I argue in section 1, and Smith's (1994, 1996) criticism of *de dicto* moral motivation as "moral fetishism."

<sup>1</sup> To illustrate the relevant distinction between *thin* and *thick* with some paradigmatic examples: *good*, *right*, and *obligatory* are *thin* moral concepts; *kind*, *honest* and even *just* are considered *thick* moral concepts. The exact nature of this distinction is a matter of dispute (see, for instance, Smith 2013).

<sup>2</sup> Sometimes Smith (1996: 177, for instance) allows using the term *de re moral motivation* to denote motivation to  $\phi$  in cases where the agent *falsely* believes that  $\phi$ -ing is right. I do not use the term *de re moral motivation* in this way.

<sup>3</sup> From this point on I use the term *virtuous* in the sense of *morally virtuous*, unless noted otherwise.

<sup>4</sup> For such defenses see for instance: Lillehammer (1997); Svavarsdóttir (1999: 199–215); Enoch (2011: 255) Hurka (2014) and Aboodi (2015). Olson (2002: 92–94) proposes two particular cases where he claims that an effective *de dicto* moral motivation seems more virtuous than particular effective *de re* moral motivations, but he fails to establish conclusively that effective *de dicto* moral motivation would be the most virtuous type of motivation under those circumstances. See also my references to Vanessa Carbonell and Arnon Keren in the beginning of the following section.

<sup>5</sup> Throughout this paper, the locution "moral inquiry" refers to inquiries into moral matters rather than referring to *morally virtuous inquiries*.

## 1 Where *De Dicto* Moral Motivation is Necessary

I am not the first to consider the necessity of *de dicto* moral motivation. Vanessa Carbonell (2013: 472–477) makes the conditional claim that if *rightness* (of actions) is not a summary concept that passes the buck to concrete “right-making features”, then *de dicto* moral motivation must be [sometimes] obligatory. (In contrast, my argument will not depend on whether or not this buck-passing account of rightness is ultimately correct.) Since Carbonell does not try to establish the denial of this buck-passing account, she is careful not to purport to have established the strong claim that lacking *de dicto* moral motivation sometimes amounts to a moral shortcoming. In an unpublished manuscript (which focuses on a different topic), Arnon Keren suggests a short argument for this strong claim:

Given our undeniable moral ignorance, we clearly ought to desire to know what our moral obligations are. But someone who does not desire to do what is right, whatever it turns out to be, either does not desire to know what his moral obligations are, or desires to know what his moral obligations are, but lacks the desire to act upon this knowledge. Either way, such a person would exhibit a moral failing.<sup>6</sup>

Keren’s important insight here is that there is an important connection between moral uncertainty (and ignorance) and the necessity of *de dicto* moral motivation. However, we can identify this connection more precisely. In order to do so, let us start by attending to some shortcomings in Keren’s argument.

First, it seems that Keren assumes that in order to act upon newly-acquired knowledge concerning one’s obligations, it is necessary to have a *de dicto* moral desire, such as the desire to *fulfill one’s obligations as such*. But it is possible that, without any *de dicto* moral motivation, the newly-acquired knowledge—or the process of acquiring it—can causally generate non-derivative *de re* motivations to fulfill the newly-understood obligations.<sup>7</sup>

Second, pace Keren, it is not clear that in every case where a person doesn’t know some of her obligations, it is morally important that she would desire to know them. Arguably, in some cases where this desire is not likely to lead her to better her conduct, she is not morally required to have it. Another related qualification is that in some cases a person can unproblematically derive her desire to know her obligations from a *de re* moral principle she believes in and wants to fulfill (non-derivatively). I am thinking of cases where a person justifiably believes that her obligations in a certain situation are grounded in a particular moral principle, such as the Kantian categorical imperative, yet doesn’t know which obligations follow from it exactly. Let me stipulatively define a person’s *derived* moral belief, obligation or value as one that the person bases (in some sense) on another normative principle.<sup>8</sup> In cases where the person

<sup>6</sup> Keren A, “The Risk of Wrongdoing: On the Moral Significance of Moral Uncertainty,” Unpublished Manuscript.

<sup>7</sup> For example, Joshua May (2013) argues that beliefs about what’s right can generate new desires on their own, without any antecedent motivation.

<sup>8</sup> Notice that my stipulated sense of *derived/underived* is psychological and agent-specific, rather than meta-physical or epistemological. (In this respect it is similar to the notion of *derivative* motivation, as Smith and his respondents—including myself—use this term.) For further clarification, here is another example: if the agent’s view is that we ought to respect human rights merely for the reason that it is a good means to maximizing universal happiness, then her belief in human rights is *derived*. In contrast, if an agent holds that respecting human rights is one of her fundamental duties, her belief in human rights is *underived*. If an agent does not base human rights on another normative principle in any way, I classify her belief in human rights as *underived* (and this probably renders most of the moral beliefs and values of non-philosophers as *underived*).

justifiably lacks knowledge or certainty concerning her *derived* moral beliefs (or values or obligations), there may be nothing wrong with wanting to acquire the relevant knowledge or resolve the relevant uncertainty *merely in order to fulfill the relevant principle* (which the person considers as more basic). This does not require *de dicto* moral motivation. To identify where *de dicto* moral motivation is irreplaceable, we have to start by focusing on uncertainty (and ignorance) concerning *underived* moral beliefs or values. This type of uncertainty is also particularly relevant for Smith's and Weatherson's arguments that I discuss in section 2 of this paper.

My argument for the importance of *de dicto* moral motivation will be laid out in the following subsections, 1.1–1.4. To clarify where I am headed, let me outline the following subsections in reverse order: Section 1.4 argues that there are situations in which the morally-best response is deciding to conduct a deep moral inquiry which seeks an appropriate resolution of a conflict between underived moral values, for a *de dicto* moral purpose. Section 1.3 argues that there are situations in which the morally-best response is deciding to reevaluate an underived moral belief, for a *de dicto* moral purpose. Each of these two subsections constitutes an independent argument for the necessity of *de dicto* moral motivation. To better understand why a non-derivative *de re* moral purpose will not suffice in the two aforementioned types of situations, we have to first appreciate the thin epistemic concern that is inherent to the two aforementioned types of deep moral inquiries, as I shall explain in section 1.2. Section 1.1 defends a more basic assumption which serves to extend the variety of epistemic states in which *de dicto* moral motivation has an important role to play: moral uncertainty concerning underived moral beliefs (or values) is sometimes justified for human beings like us, even when we happen to have the right moral views.

### 1.1 Justified Moral Uncertainty Concerning Underived Moral Beliefs

There are some (not very rare) situations where moral uncertainty concerning underived moral beliefs is justified for human beings like us, who are fallible even in fundamental moral judgments.<sup>9</sup> (I focus only on such persons. I doubt that we have reliable intuitions about the psychological structure of beings that are infallible in their fundamental moral judgments.<sup>10</sup>) Most clearly, a person with a false underived moral belief could be justified in being uncertain about it, especially when she has information that suggests that it is false.

But even when a person's underived moral beliefs are correct, she might obtain new information that *misleadingly* suggests that they are wrong, and this could justify uncertainty about them.<sup>11</sup> This information might consist of: counter-arguments; facts that elicit counter-intuitions to the relevant belief; unexpected implications of the relevant belief; opinions of others whom the agent has good reason to trust; signs that there is something wrong with the experts or the tradition upon which the agent relied in forming her view; signs that the causal process that has led the agent to the relevant belief is unreliable; signs that there is something wrong with the agent's own memory or deliberative abilities, or the simple fact that a very long

<sup>9</sup> Smith and Weatherson would probably agree to this claim, as it seems to lie in the background of their arguments that I discuss in section 2.

<sup>10</sup> See also the discussion regarding morally perfect people in Svavarsdóttir 1999: 214–215.

<sup>11</sup> David Christensen makes a similar point regarding what he calls "ideally rational agents" (Christensen 2007: 10–16).

time has passed since the agent has last reflected on the matter. Under some such circumstances, full confidence in the relevant belief would not be justified.<sup>12</sup>

## 1.2 The Necessity of a Thin Epistemic Concern

As I shall argue in sections 1.3–1.4, there are situations where the most virtuous way of coping with uncertainty concerning underived moral beliefs is by deciding to conduct a deep moral inquiry into the matter. One type of deep moral inquiry that may be called for consists in reevaluating an underived moral belief. Intentional reevaluation must consist of intentional evaluation, and genuine intentional evaluation necessitates a concern (at least an implicit one) for the relevant truth of the matter, or some similarly *thin* epistemic concern: a concern for the right answer, for a justified answer, for a good enough answer, etc. Action-theoretic accounts for such intentional reevaluation should be careful not to explain away this inherent thin epistemic concern, or to reduce it to some thicker concern.

To see this, consider an attempt to account for an intentional reevaluation of an underived value as motivated merely by a concern for another underived *thick* value: Isabel wants to minimize suffering in the world, and also to keep promises, and considers each of these values as a true moral principle that is not grounded in some other moral principle. Now Isabel faces a situation where she can prevent suffering by breaking a promise. (This situation does not necessarily *call for* reevaluation, unlike the type of situations I will characterize in the following subsection. Here I am only examining attempts to account for reevaluations of an underived value without attributing a *thin* epistemic concern. Conflicts such as Isabel's seem particularly friendly to such attempts.) In this situation, Isabel's non-derivative concern for minimizing suffering might motivate her to reevaluate keeping promises.

However, the concern for minimizing suffering (or any other *de re* moral value) cannot be the *only* concern in a genuine intentional reevaluation of keeping promises. Indeed, this concern could lead Isabel to think, "oh no, keeping that promise would lead to suffering, I must break it." But if this captures all of Isabel's thoughts, it does not seem that she is genuinely reevaluating keeping promises (intentionally), because she is simply ignoring the value of keeping promises. Suppose we add to her thoughts, "but breaking a promise is bad, at least as I used to think. Well, I am free to change my mind. From now on I shall hold that breaking a promise is the right thing to do when it minimizes suffering. This way I can allow myself to break the promise now, and thereby prevent much suffering." These thoughts, as formulated, do not yet convey genuine intentional reevaluation.<sup>13</sup> What's missing is an effort to evaluate to what extent or under which conditions breaking promises would *really* be wrong, or some similar thin epistemic concern.<sup>14</sup> Without a thin epistemic concern Isabel would not count as genuinely reevaluating her belief (intentionally). Such a concern might be

<sup>12</sup> It may even be argued, along those lines, that for every underived belief in a principle that could serve as the object of a non-derivative *de re* moral motivation, there are some possible circumstances under which full confidence in it would not be justified; but this would not be necessary for my argument.

<sup>13</sup> Compare: "oh no, keeping that promise requires me to get up from bed so early in the morning. I must break it. But breaking a promise is bad, at least as I used to think. Well I am free to change my mind. From now on I shall hold that breaking a promise is the right thing to do if the promise requires you to get up early in the morning. This way I can allow myself to break the promise now, and stay in bed."

<sup>14</sup> Note that if Isabel changes her mind about the moral value of keeping promises without such a concern, merely in order to minimize suffering, then she seems guilty of wishful thinking. This applies also to cases where Isabel's desire to minimize suffering would lead her (by itself) to start believing that the value of keeping promises is grounded in minimizing suffering, without genuinely trying to determine whether this is right.

occasionally *triggered* by non-derivative *de re* moral concerns or motivations—as it could be by watching a movie or by many other experiences—but it cannot be reduced to (or explained away by) such motivations.

### 1.3 The Deliberative Role that only *De Dicto* Moral Motivation Can Play

Now let us proceed from thin epistemic concern to practical *de dicto* moral motivation.

It seems that in some cases there would be nothing wrong with reevaluating an underived moral belief merely for a non-derivative thin epistemic purpose, such as figuring out the truth. But having solely an epistemic concern is inappropriate when the relevant moral issue has important and urgent practical implications. Consider the following case: Simona is a leader who foresees a morally important choice-situation next week, which might have morally significant effects on many people. She realizes that the right choice in this upcoming choice situation depends on whether equality is indeed of intrinsic value, as she used to think. Simona also realizes that reevaluating<sup>15</sup> her underived moral belief in equality is likely to help her make a better choice next week.<sup>16</sup> To highlight the insufficiency of non-deliberative dispositions to reevaluate, as well as mere epistemic motivations, let us add two further stipulations to Simona's situation: (1) According to Simona's evidence, the circumstances will not be favorable for deep moral inquiries on any occasion between now and the anticipated choice situation *except for* tomorrow afternoon, on condition that she schedules the reevaluation for tomorrow afternoon in advance. (2) Simona is more curious and less certain about a different moral matter—say, whether states can have property rights on the moon—which has no practical implications for what Simona has to do in the near future.

It seems very plausible that Simona's morally-best possible response in this situation is deciding to reevaluate equality tomorrow afternoon. Furthermore, it would be less virtuous for Simona to make this scheduling decision merely for a non-derivative epistemic purpose, than to make it (also) for the purpose of making a better choice, or some other *de dicto* moral purpose. Fans of *de dicto* moral motivation would probably accept this claim as obvious. But I think that others should also accept this claim, for the following reason: only *de dicto* moral motivation can put the person in a good position to make the morally best choice in such cases (at least if we limit ourselves to realistic psychological constitutions). Note that throughout this paper my notion of *best choice* or *best response* is meant to be understood as relativized to the person's epistemic state, or mentally accessible evidence, in contrast to the perspective of an omniscient being.<sup>17</sup> I take it that a necessary condition for *responding in the most virtuous manner* is responding in the morally best manner—in the aforementioned relativized sense—out of a motivation or disposition that would quite reliably generate such responses when they are called for. (The type of responses my argument focuses on are mental: conducting moral inquiries or deciding to do so.)

<sup>15</sup> Reevaluating an underived belief in equality, as I use the term *reevaluating*, may sometimes consist in trying to figure out whether equality is of intrinsic value or whether it is valuable merely as a means to something else.

<sup>16</sup> Perhaps some think that reevaluation does not help to improve views concerning basic moral values. Responding to such a view would take away from the focus of this paper. Note that I am not claiming that there cannot be good people whose observable behavior is morally exemplary, who are *not* constituted to reevaluate their underived moral beliefs. (I thank an anonymous reviewer for raising this possibility.) However, given that such people are fallible in basic moral matters, I think they would be more virtuous if they also had the capacity and tendency to reevaluate underived moral beliefs when this is called for.

<sup>17</sup> This notion corresponds to the prospectivist/perspectival sense of moral obligation (or of “ought”) which is explained and defended (as more relevant than more objective senses) in Zimmerman 2008; Lord 2015.

Even if a person could be psychologically constituted so that once she changes her underived moral views she comes to have non-derivative *de re* moral motivations that accord with them,<sup>18</sup> a purely epistemic motivation would not reliably lead her to make the best choice in situations like Simona's.<sup>19</sup> To see this more clearly, compare two possible versions of Simona in the situation described above: Simona<sub>1</sub>, who has a *de dicto* moral motivation, vs. Simona<sub>2</sub> who instead (i) has an epistemic motivation to better her moral views, and (ii) is psychologically constituted so that once she changes her underived moral views she comes to have non-derivative *de re* moral motivations that accord with them. Now focus on the moment when each Simona is called upon to schedule the reevaluation to tomorrow afternoon. Simona<sub>1</sub>'s *de dicto* moral motivation could lead her to care about the foreseen choice-situation and its timing, to take these and the rest of the relevant practical facts into account, and to decide to schedule the reevaluation of equality for tomorrow afternoon, while postponing the inquiry into property rights on the moon. But how could Simona<sub>2</sub> be in a good position to make this decision rather than to reevaluate keeping promises right away, or worse — spend all her free time throughout the week reevaluating her beliefs about property rights on the moon? I do not see how epistemic motivations such as Simona<sub>2</sub>'s could discriminate—as is morally appropriate—between different possible reevaluations on the basis of how the issue impacts what the agent has to do, how important and how urgent the relevant practical choices are, and similar practical considerations. I cannot imagine a psychologically realistic account that explains how Simona<sub>2</sub> could be constituted to somehow reliably conduct the right reevaluations at the right times merely out of a purely epistemic motivation (or out of dispositions that do not count as motivations).<sup>20</sup>

The insufficiency of a purely epistemic purpose in Simona's situation is most apparent if we focus (as we did in the previous paragraph) on the moment where the decision to schedule the moral inquiry for tomorrow is called for. Such moments suffice for my argument, and for establishing the implications that I identify in section 2. But notice that it is usually more virtuous when the practical motivation for deciding to conduct a moral inquiry (assuming it is a virtuous motivation) persists throughout the inquiry, rather than vanishes. Surely, if we fill in the details in a certain way, this applies also to Simona's case: it would be most virtuous if her purpose for *deciding to conduct* the inquiry would also be her purpose for *actually conducting* the inquiry. This means that the thin epistemic motivation inherent to Simona's inquiry (such as *to figure out what's right*) could be described as a derivative of the same practical purpose she had in deciding to conduct the inquiry.

<sup>18</sup> See, for instance, Smith's (2002) description of how "the capacity to have coherent psychological states" (which I interpret as a type of disposition which Smith does not count as a motivation) can change the relative strength of some desires in correlation with changes in the person's relative certitude in some of her values, without any *de dicto* moral motivation.

<sup>19</sup> I thank two anonymous reviewers for pressing me here.

<sup>20</sup> Notice also how difficult it is to come up with a non-derivative *de re* moral goal that provides a motive for reflecting—for the purpose of deciding what to do—on whether this goal itself should be pursued and to what extent. None of the goals that appear in Smith's (1994: 75) sample list of *de re* moral motivations fits this bill. Wanting to promote equality, for instance, does not provide any motive to reevaluate equality. This point seems particularly important if it turns out—due to considerations that were raised in section 1.1—that for every non-derivative *de re* moral goal there are some possible circumstances under which it is appropriate to reevaluate it for the purpose of deciding what to do. (Does this apply also to *de dicto* moral goals? It probably does. But such goals can unproblematically motivate the reevaluation of their own appropriateness. I discuss this asymmetry and its implications further in Aboodi MS).

Could this practical purpose be a non-derivative *de re* moral motivation? Suppose that, at the moment when Simona is called upon to schedule the reevaluation, she has in mind only non-derivative *de re* moral purposes. It does not seem that the thin epistemic concern inherent to reevaluating an underived moral belief (such as the concern for the moral truth) can be properly derived from any kind of practical purpose other than a *de dicto* moral purpose (such as *to act in accordance with moral truth*). How could there be a virtuous, sound deliberative route from non-derivative *de re* moral purposes to the decision to schedule an open-minded inquiry into whether equality is of intrinsic value? Simona does not know what the results of the inquiry will be; so why should she anticipate that it will advance her *de re* moral purposes? How could it advance them better than a different kind of moral inquiry (which does not constitute deep reevaluation), one that seeks merely whether equality is instrumental for (or properly derives from) her *de re* moral purposes?

Perhaps the most virtuous goal in situations like Simona's does not have to be as abstract as *to act in accordance with moral truth*. For instance, the agent may want *to do what best accords with* a particular set of values, where this set includes all the values she already subscribes to. Let us name such motivations, which aim at practically fulfilling such coherentist ideals (under such descriptions), *coherentist* motivations. Another interesting example is the goal of acting according to the conclusion that a good (or perfect or ideal) deliberative process would lead to, whatever this conclusion turns out to be. I refer to motivations that aim at practically fulfilling such procedural ideals (under such descriptions) as *procedural*. I classify coherentist, procedural, and similar types of motivation as *de dicto* moral motivations, because the person pursues the relevant ideal under a normative description that is anchored in a *thin* normative concept (such as "*what best accords with*"). In regard to the possibility that Smith's and Weatherson's depreciating view is not meant to target procedural, coherentist and similar motivations, see section 2. While I doubt that such motivations suffice for responding in the most virtuous manner in the whole scope of relevant situations, I do not rule out this possibility in this paper.

I conclude that being in a (psychologically realistic) good position to decide to reevaluate an underived moral belief at the right times requires *de dicto* moral motivation; at least in the broad sense that includes procedural, coherentist, and similar types of motivation (such as wanting to act rationally or justifiably). In this broad sense, *de dicto* moral motivation is necessary for the most virtuous responses at least in situations like Simona's.<sup>21</sup>

#### 1.4 Resolving the Conflict between Underived Moral Values

Simona's example could be easily reformulated to illustrate a situation where the agent has two underived moral values (or beliefs), each recommending a different course of action (which is

<sup>21</sup> I include within this broad sense of *de dicto* moral motivation goals that the agent conceives of as normative but not moral, such as "to do what I have most reason to do" ("rightness" could be interpreted in this way as well). I also include thin normative goals that the agent conceives of as restricted to a particular context or domain, such as "to vote for the candidate that is best overall." Finally, I include also the psychological state of being committed to a thin normative constraint such as "not to act immorally" or "not to choose wrongly" (I thank David Heyd for raising this).



incompatible with the other's recommendation), and the best decision available to her is to schedule an inquiry that aims at appropriately resolving this practical conflict.<sup>22</sup> For instance: Simona\*'s situation is identical to Simona's except that her uncertainty only concerns how to appropriately resolve the practical conflict between her underived belief in equality and her underived belief in maximizing happiness, because each recommends a different choice in the anticipated choice situation next week. In the type of inquiries that are called-for in such situations, the agent attempts to answer questions such as: Should I be guided in this case by the principle that  $\varphi$ -ing is good (or required) or by the principle that  $\psi$ -ing is good? Is  $\varphi$ -ing more important than  $\psi$ -ing (under such circumstances)? Would it be right (or reasonable or permissible) to opt for some practical compromise between these principles? Like reevaluations of underived moral beliefs, inherent to such inquiries is a *thin* epistemic concern, such as a concern for whether keeping a promise is more important than minimizing harm under the given circumstances.<sup>23</sup> In situations like Simona\*'s—where it is of practical moral importance to appropriately resolve the conflict—it seems more virtuous when this concern is a derivative of *de dicto* moral motivation. Those who remain doubtful of this claim face the challenge of providing an alternative, a psychologically realistic explanation for how an agent could appropriately conduct such inquiries at the right times. It seems implausible that the best explanation will not appeal to deliberative planning which takes into account practical considerations that are relevant to the timing of the inquiry. And it seems very difficult to envision a virtuous, sound deliberative route from non-derivative *de re* moral motivations to Simona\*'s decision to conduct this type of inquiry tomorrow afternoon: When this decision is called for, Simona\* does not know what the results of the inquiry will be; so why should she anticipate that it will advance her *de re* moral purposes? I do not deny that one could conjure some strange version of Simona\*'s case which allows for a good answer to this question; but this would not help to offer an account of a psychological constitution that puts fallible human beings in a good position to schedule—or to actually conduct—the right deep moral inquiries at the right times without any kind of *de dicto* moral motivation. If we restrict ourselves to plausible, psychologically realistic accounts, we can conclude that *de dicto* moral motivation (at least in the broad sense) is necessary for coping virtuously with some situations where appropriately resolving a practical conflict between underived moral values is morally important, and the agent is uncertain about which resolution would be appropriate.

Note that the appropriate resolutions in the relevant type of inquiries may be of various sorts. I do not rule out, for example, discovering a general normative principle that determines the right course of action under circumstances of a certain type; or, without discovering any general principle, finding a unique course of action that strikes the right balance between the competing values under the particular given circumstances; or even flipping a coin between two compromises that seem legitimate to the best of the agent's estimates. There may be more than one optimal resolution, and perhaps it suffices to choose just a *good enough* resolution.

<sup>22</sup> This part of my argument might be less exciting for those utilitarians and other monists who also hold that the most virtuous human beings believe in the one fundamental principle that exhausts all of morality, and recognize its special status. Such agents would never experience a conflict between two underived values because, by definition, they have only one underived value. But, going back to my argument in the previous subsection, notice that it would sometimes be irresponsible for such agents not to reevaluate their monistic view in light of the many known counter-intuitions and counter-arguments to such monistic views.

<sup>23</sup> Note that merely attributing to an agent independent non-derivative motivations to fulfill the relevant conflicting values does not explain why the agent is conducting a moral inquiry rather than just being led to action unreflectively by the strongest motivation. The discussion of Isabel's case in section 1.2 helps to clarify this, even though it focuses on reevaluations of underived moral beliefs.

The minimal assumptions that my argument requires is that some possible resolutions in the given situation might be morally less preferable, that the agent might be able to avoid them by conducting a moral inquiry, and that this would be the agent's best response (in the sense that is relativized to the agent's epistemic state or mentally accessible evidence). These assumptions could hold even if the best possible resolutions are incommensurable or *on a par*, in Ruth Chang's (2002) sense (neither is better than the other, nor are they of equal value, and yet they are evaluatively comparable). If moral incommensurability and parity are possible, surely sometimes it takes a deep moral inquiry to determine whether one of these relations indeed holds between two particular resolutions, or rather, one of these resolutions is better than the other.

## Section 1 Overview

I conclude section 1 with a schematic presentation of one of its central arguments for the necessity of *de dicto* moral motivation:

**P1.** Intuitively, the morally-best response (in the relativized sense) for Simona and Simona\* involves conducting—tomorrow afternoon in particular—the deep moral inquiry which has more urgent practical implications.

**C1.** Having an effective *de dicto* moral motivation in Simona's and Simona\*'s situations is psychologically necessary for being in a good position to respond in the morally-best manner: to conduct the right deep moral inquiries at the right times. (From P1, and in light of the thin epistemic concern that is inherent to the relevant deep moral inquiries.)

**P2.** One necessary condition for *responding in the most virtuous manner* is responding in the morally-best manner out of a motivation or disposition that would quite-reliably generate such responses when they are called for.

**C2.** Effective *de dicto* moral motivation is necessary for the most virtuous response in a certain type of situations. (From C1 and P2.)

## 2 Implications

My argument in the previous section seems to imply that *de dicto* moral motivation should be cultivated (by educators, for instance), at least to a certain extent. It also seems to have implications for the debate on *moral deference*, since a certain kind of moral deference is arguably tied to *de dicto* moral motivation.<sup>24</sup> But I do not elaborate on these implications in this paper. Instead, I focus (in the following two subsections) on the implications for two contemporary arguments that share a similar structure. Both rely on what I call *the depreciating view of de dicto moral motivation*, according to which every *effective* (in the sense of actually affecting behavior) instantiation of *de dicto* moral motivation is non-virtuous; or at least less virtuous than an effective instantiation of some non-derivative *de re* moral motivation. My argument in the previous section undermines this depreciating view, for it establishes that an effective *de dicto* moral motivation (understood broadly) is necessary for putting fallible human beings in a good position to make the right choice in a certain type of cases. In these cases it would not be more virtuous to be motivated effectively instead by any non-derivative *de re* moral motivation.

<sup>24</sup> See McGrath (2011: 135), Enoch (2014: 28).

## 2.1 Smith's Argument against Motivational Judgment Externalism

According to Motivational Judgment Externalism, making a sincere moral judgment of the form “I ought to  $\phi$ ” does not necessitate having any motivation to act accordingly. Smith (1994: 71–76) argues against a certain kind of Motivational Judgment Externalism,<sup>25</sup> challenging it to account for “the striking fact” that when “good and strong-willed people” change their moral views—particularly with regard to their fundamental values—their motivation changes accordingly. Smith (1994: 71) illustrates this phenomenon with the following example:

Suppose I am engaged in an argument with you about [...] whether we should vote for the libertarian party at some election as opposed to the social democrats. In order to make matters vivid, we will suppose that I come to the argument already judging that we should vote for the libertarians, and already motivated to do so as well. During the course of the argument, let's suppose you convince me that I am fundamentally wrong. I should vote for the social democrats, and not just because the social democrats will better promote the values I thought would be promoted by the libertarians, but rather because the values I thought should and would be promoted by the libertarians are fundamentally mistaken. You get me to change my most fundamental values.

One version of Smith's argument proceeds as follows: Externalism (of the relevant kind) can only account for such cases of motivational change—which exhibits “the striking fact”—by attributing to the agent an effective *de dicto* moral motivation: claiming that the agent wants to act in accordance with his new moral view *because* he wants to act in accordance with morality as such (or some similar *de dicto* moral purpose). But, on the basis of Smith's depreciating view of *de dicto* moral motivation, such motivation is problematic and its attribution cannot plausibly account for virtuous behavior. Thus, Externalism is implausible.<sup>26</sup>

However, while Smith (1994, 1996) and his respondents have been focusing on explaining the correlation between the change in underived beliefs and the change in motivation that follows, they seem to have overlooked the need to account for the deliberative process that leads virtuous people to change their underived beliefs. Notice that usually, changing underived moral beliefs without any genuine attempt to figure out the relevant moral truth seems morally irresponsible, especially in cases where the person has an important practical decision to make on the basis of the relevant moral matter. This could be illustrated with Smith's (1994: 71) own example quoted above, that of changing one's belief concerning fundamental libertarian values before upcoming elections.<sup>27</sup> Since this example fits so well with Simona's scenario that I described in section 1.3, let us name the protagonist in this

<sup>25</sup> The particular version of Externalism that Smith attacks denies the existence of the type of conceptual or necessary connection between moral judgment and motivation that certain Motivational Judgment Internalists assume; but the details need not concern us here.

<sup>26</sup> In conversation, Michael Smith seemed to confirm that his argument was meant to rely on something like the depreciating view as I formulated it (which does not square with the narrow interpretation of Smith's moral fetishism charge that Sepielli 2016 offers). I find that this version of his argument has two advantages: (1) It is not committed to the implausible claim—criticized for instance by Copp (1997) and Svavarsdóttir (1999: 199)—that Externalists must assume that the only primary moral motivation that good people have is *de dicto* moral motivation; and (2) It is not immediately refuted by the fact that virtuous behaviors can be motivated simultaneously by both *de dicto* and *de re* moral motivations (Olson 2002: 91; Enoch 2011: 255).

<sup>27</sup> This could also be illustrated by Smith's (1996: 180) example of a Utilitarian who at some point comes to believe that “it is sometimes right to give extra benefits to his family and friends, even when doing so cannot be given a utilitarian justification.” (Smith argues that this case presents an explanatory problem for Externalists of the relevant kind.)

example (who changes his view concerning libertarian values) *Simon*. It seems that Simon, like *Simona*, foresees a morally important choice situation. (Those who think that the effect of a single citizen's vote is too insignificant to be morally important should imagine Simon as a leader whose political support is likely to influence many other voters.) If we fill in the details in a certain way, Simon's situation calls for deciding to conduct a moral inquiry that aims at figuring out the relevant moral truth—or some similarly thin epistemic purpose—partly *for the purpose* of voting in accordance with the relevant moral truth, or some similar *de dicto* moral purpose. It seems pretty obvious here that without some *de dicto* moral purpose (in the broad sense) Simon would be less virtuous, even if we do not stipulate the need to schedule the inquiry to a particular future moment; but the skeptics are invited to add this stipulation so that Simon's case matches *Simona's* case. The existence of such cases suffices to refute the depreciating view.

In order to avoid this result, Smith might claim that his depreciating view applies only to *de dicto* moral motivation in a certain narrow sense which excludes *procedural*, *coherentist*, or similar motivations. This may allow Smith to account for the types of responses on which my argument focuses by attributing one of these motivations. But if Smith goes down this road, allowing that one of these motivations is virtuous, then this virtuous motivation could be used by Externalists to explain *the striking fact*. Let me illustrate this with regard to *coherentist* motivation. Suppose that Simon reevaluates the relevant libertarian value in order *to do what accords overall best with all his existing values*. Or suppose that, after Simon becomes convinced by Smith's (2004: 45) analysis of *rightness*, Simon aims at doing *whatever his ideally rational self, who is fully informed and whose set of desires is maximally coherent, would advise (or want) his actual self to do*. Such motivation could lead Simon to implement his new moral view by voting for the social democrats.

Arguably, if *procedural* motivation is not fetishistic or problematic, and can reliably generate the responses discussed in the previous section, it could also account for the striking fact. Thus, no precisification of *de dicto moral motivation* allows Smith's argument to be sound: Whereas the broad precisification (which includes procedural and coherentist motivations) falsifies one premise (the depreciating view), the narrow falsifies the other central premise — that Externalism can only account for such cases of motivational change by attributing to the agent an effective *de dicto* moral motivation.<sup>28</sup>

## 2.2 Weatherson's Argument against Avoiding Moral Recklessness

What Weatherson (2014: 147–152) calls *avoiding moral recklessness* amounts to refraining from an action that you deem probably permissible, in order to avoid the risk that—contrary to your best estimates—this action is actually impermissible.<sup>29</sup> To illustrate:

<sup>28</sup> In response to my argument (and I thank an anonymous reviewer for raising this possible way out), Smith could concede that throughout the relevant types of moral inquiries *de dicto* moral motivation is not fetishistic, but insist on a restricted version of what I called his *depreciating view* of *de dicto* moral motivation that claims it is always inferior in non-deliberative activities. I think that such a view would be hard to defend: For instance, given that Simon's effective motivation for reexamining the relevant moral issue was *de dicto* moral motivation, and that it was virtuous for him to be so motivated up until the point of changing his views, it is not intuitively less virtuous for Simon to vote according to his new views—at least partly—out of *de dicto* moral motivation, than to do so purely out of new non-derivative *de re* moral motivations. But there's room for further discussion about this. This relates to the general question of when *de dicto* moral motivation is less virtuous, which I briefly address in section 3, referring to my 2015 article.

<sup>29</sup> This way of coping with moral uncertainty may be considered as a kind of *moral hedging*.

Martha is deciding whether to have steak or tofu for dinner. She prefers steak, but knows there are ethical questions around meat-eating. She has studied the relevant biological and philosophical literature, and concluded that it is not wrong to eat steak. But she is not completely certain of this; as with any other philosophical conclusion, she has doubts. [Weatherson 2014: 143.]

If Martha refrains from eating the meat due to the risk that it is—contrary to her best estimates—impermissible, this constitutes *avoiding moral recklessness*.

Weatherson (2014: 141) attacks *avoiding moral recklessness* by claiming that it always involves “an unpleasant sort of motivation, what Michael Smith calls ‘moral fetishism’.” Weatherson argues, referring to Smith (1994), that [effective] *de dicto* moral motivation is always a moral vice, and that it is necessary for *avoiding moral recklessness*. For instance, if Martha *avoids moral recklessness*, Weatherson (2014: 152) claims that “she has to care about morality as such. And that seems wrong.”

It is possible that when Weatherson (2014: 147–152) attacks *de dicto* moral motivation he has in mind only a particular kind of *de dicto* moral motivation: His (Weatherson 2014, 160) formulation “trying to figure out something about this magical thing, the good” hints, perhaps, at some platonic metaethical assumptions that are connected to the kind of *de dicto* moral motivation he opposes. Weatherson’s (2014: 160) proposed account of practical moral inquiries as an attempt to reconcile conflicting values, suggests that he (implicitly) attributes to the virtuous inquirer what I called coherentist moral motivation.

However, notice that Martha may *avoid moral recklessness* (and turn down the steak) out of coherentist motivation. In light of her uncertainty regarding the permissibility of eating meat, Martha might seek what accords best with all her existing moral beliefs (for example: the belief that all-things-equal it is better that animals do not suffer, and the belief that in certain respects human beings are more important than other animals). Martha may estimate that even though the permissibility of eating meat is most likely what accords best with all her beliefs, there is a significant chance that it is actually the *impermissibility* of eating meat that accords best with all her beliefs. In such a state of moral uncertainty, she may choose to *avoid moral recklessness*, deciding to refrain from eating meat (at least until she becomes more certain) due to the significant risk mentioned above. In this case *avoiding moral recklessness* does not require any other kind of *de dicto* moral motivation.

This shows that Weatherson’s claim that *avoiding moral recklessness* necessitates *de dicto* moral motivation cannot be true unless we count Martha’s coherentist motivation in the above scenario as *de dicto* moral motivation. Similarly, it is easy to see that other kinds of coherentist motivation, as well as procedural and similar types of motivations, can also suffice for *avoiding moral recklessness*. (The example in the previous paragraph can be easily reformulated to show that.) So Weatherson’s claim that *avoiding moral recklessness* necessitates *de dicto* moral motivation is false unless *de dicto moral motivation* is understood in the broad sense (which includes procedural, coherentist, and similar motivations). But it would be unreasonable to claim that *de dicto* moral motivation—in this broad sense—amounts to a moral vice, in light of my argument in the previous section. In particular, Weatherson should appreciate the necessity of *de dicto* moral motivation for the most virtuous response to some situations where appropriately resolving a practical conflict between underived moral values is morally important, as explained in section 1.4. So no precisification of *de dicto moral motivation* allows Weatherson’s (2014: 147–152) argument to be sound: Whereas the broad precisification falsifies one premise (the depreciating view), the narrow falsifies the other central premise — the premise that *avoiding moral recklessness* necessitates *de dicto* moral motivation.<sup>30</sup>

<sup>30</sup> For a different yet somewhat similar critical response to Weatherson 2014 see Sepielli 2016.

### 3 Necessary or Fetishistic?

One may wonder how my argument for the necessity of *de dicto* moral motivation squares with criticisms such as Smith's (1994: 75; 1996: 183), according to which this motivation amounts to *moral fetishism*. One modest version of this criticism does not conflict with my argument: sometimes doing what's right merely out of an effective *de dicto* moral motivation is less virtuous than doing it out of a non-derivative *de re* moral motivation.<sup>31</sup> For instance, as in the famous case with "one thought too many" in Bernard Williams (1981: 17–18), to which Smith's (1994: 75) appeals, it seems more virtuous to jump to rescue one's drowning wife out of a non-derivative motivation to save her life (or some other *de re* moral motivation).<sup>32</sup> My argument does show, however, that it would be wrong to generalize from such cases to the depreciating view that Smith and Weatherston rely upon, according to which every effective instantiation of *de dicto* moral motivation is less virtuous than an effective instantiation of some non-derivative *de re* moral motivation.<sup>33</sup> Such a generalization would fail to account for cases where lacking an effective *de dicto* moral motivation would result in "one thought too few": for instance, Simona's thought that in order to make a better practical choice next week she should reevaluate an underived moral belief tomorrow afternoon. This supports preferring the alternative generalization that I proposed in Aboodi 2015: there are only two contingent factors that can render an instantiation of *de dicto* moral motivation less virtuous than some alternative motivation that would lead to the same (right) action: (1) the circumstances are such that it would be more virtuous to be moved directly by certain non-deliberative dispositions (such as an emotional attachment to one's spouse),<sup>34</sup> or (2) the circumstances are such that *de dicto* moral motivation has sufficiently significant practical disadvantages (such as generating unnecessary moral reflections that waste precious time).

A different version of the fetishism charge alludes to an inherent problem in *de dicto* moral motivation. For instance, Smith (1996: 183) claims that *de dicto* moral motivation inherently involves non-virtuous alienation from the concrete things that are of primary moral value or importance. To take another example: Certain kinds of *de dicto* moral motivation, such as wanting to be virtuous, or wanting that one's acts will have a high moral status, may be criticized as self-centered.

The claim that all kinds of *de dicto* moral motivation are inherently fetishistic or inferior cannot be squared with the argument I presented in section 1. But I'd like to conclude by raising the possibility that, whereas some kinds of *de dicto* moral motivation are indeed vulnerable to criticisms like the ones in the preceding paragraph, there may be other, more virtuous kinds—which suffice for coping in the most virtuous manner with the situations that I discussed in section 1—that are invulnerable to any such criticism. For example, it does not seem that *wanting that the good as such will be promoted* is vulnerable to the self-centeredness charge. And perhaps the kind of *de dicto* moral motivation where the agent wants to do what's right *for the sake of* "the concrete things that are of primary moral value or importance" (understood *de dicto*) does not constitute any problematic alienation. Especially so if the

<sup>31</sup> Note that this modest criticism of *de dicto* moral motivation does not suffice for Smith's (1994: 71–76) and Weatherston's (2014: 147–152) arguments which I discussed in section 2.

<sup>32</sup> I thank an anonymous reviewer for encouraging me to address such cases in this paper.

<sup>33</sup> Thomas Hurka (2014) also argues against this view, responding to Nomy Arpaly and Timothy Schroeder (2014).

<sup>34</sup> This formulation might seem circular at first glance, but notice that not every non-derivative *de re* moral motivation is a *non-deliberative disposition*.

person is constituted to have the right affective responses to the things that are of primary moral value *de re*.<sup>35</sup> Exploring such possibilities (further)<sup>36</sup> seems to be the appropriate response to Smith's (1994, 1996) and Nomy Arpaly's (2015) criticisms of *de dicto* moral motivation. In light of the important deliberative role of *de dicto* moral motivation, which we have been discussing, such criticisms should not persuade us to dismiss all kinds of *de dicto* moral motivation as non-virtuous moral fetishism.

**Acknowledgments** This work was conducted as part of a research project of Nikkonna, registered Amuta. Many thanks to all those who have provided valuable feedback: Dan Baras, David Enoch, David Heyd, Alex Hyun, Tom Hurka, Hasko von Kriegstein, Nethanel Lipshitz, Ofer Malcai, Joshua May, Eli Pitcovski, Ittay Nissan-Rosen, Saul Smilansky, Etye Steinberg, Preston Werner, and two anonymous reviewers for this journal. I thank Sarah Paul for the title.

## References

- Aboodi R (2015) The wrong time to aim at what's right: when is *de dicto* moral motivation less virtuous? *Proceedings of the Aristotelian Society* 115.3:307–314
- Aboodi R (MS) *De dicto* moral motivation and moral uncertainty. Doctoral dissertation, Hebrew University of Jerusalem
- Arpaly N (2015) Huckleberry Finn revisited: Inverse akrasia and moral ignorance. In: Clark R, McKenna M, Smith A (eds) *The Nature of Moral Responsibility*, Oxford, pp. 141–156
- Arpaly N, Schroeder T (2014) *In praise of desire*. New York: Oxford University Press
- Carbonell V (2013) *De dicto* desires and morality as fetish. *Philos Stud* 163(2):459–477
- Chang R (2002) The possibility of parity. *Ethics* 112.4:659–688.
- Christensen D (2007) Does Murphy's law apply in epistemology? *Self-doubt and rational ideals*. *Oxford studies in Epistemology* 2:2–31.
- Copp D (1997) Belief, reason, and motivation: Michael Smith's the moral problem. *Ethics* 108(1):33–54
- Enoch D (2011) *Taking morality seriously: a defense of robust realism*. Oxford University Press, Oxford
- Enoch D (2014) A defense of moral deference. *J Philos* 111(5):229–258
- Hurka T (2014) Many faces of virtue. *Philos Phenomenol Res* 89(2):496–503
- Lillehammer H (1997) Smith on moral fetishism. *Analysis* 57:187–195
- Lord E (2015) Acting for the right reasons, abilities, and obligation. *Oxford studies in metaethics* 10:26–52
- May J (2013) Because I believe it's the right thing to do. *Ethical Theory Moral Pract* 16(4):791–808
- McGrath S (2011) Skepticism about moral expertise as a puzzle for moral realism. *J Philos* 108(3):111–137
- Olson J (2002) Are desires *de dicto* fetishistic? *Inquiry* 45:89–96
- Sepielli A (2016) Moral uncertainty and fetishistic motivation. *Philos Stud*. doi:10.1007/s11098-016-0645-z
- Smith M (1994) *The moral problem*. Wiley, Oxford
- Smith M (1996) The argument for Internalism: reply to miller. *Analysis* 56(3):175–184
- Smith M (2002) Evaluation, uncertainty and motivation. *Ethical Theory Moral Pract* 5(3):305–320
- Smith M (2004) *Ethics and the a priori: selected essays on moral psychology and meta-ethics*. Cambridge University Press, Cambridge
- Smith M (2013) On the nature and significance of the distinction between thick and thin ethical concepts. In: Kirchin S (ed) *Thick Concepts*, Oxford.
- Svavarsdóttir S (1999) Moral cognitivism and motivation. *Philos Rev* 108(2):161–219
- Weatherston B (2014) Running risks morally. *Philos Stud* 167(1):141–163
- Williams B (1981) *Moral Luck: Philosophical Papers 1973–1980*. Cambridge, Cambridge University Press
- Zimmerman M J (2008) *Living with uncertainty: the moral significance of ignorance*. Cambridge University Press, Cambridge

<sup>35</sup> I discuss the relevant significance of such non-deliberative dispositions in Aboodi 2015: 310, 312.

<sup>36</sup> I make some progress with this project in Aboodi MS.