Abstract

Philosophers have mistakenly used the phrase “motivation by rightness *de re*” to refer to motivations whose objects are right-making features. If the “*de re*” qualifier were used correctly, this phrase would instead refer to a motivation whose object is that the agent acts rightly, but that represents moral rightness under another description. Moreover, many instances of ordinary moral deliberation cannot be fully explained by the hypothesis that deliberators just care about the particular right-making features that they see at stake in their circumstances. They can only be explained by taking deliberators to have a further, overarching concern: a concern to strike the right balancebetween everything important at stake. And these overarching concerns are often instances of motivation by rightness *de re*, correctly construed, given the conceptual role of ordinary deliberators’ judgments that a certain action strikes the right balance. These observations challenge various criticisms of motivation by rightness *de dicto*.

Keywords

moral deliberation, moral motivation, moral metasemantics, moral fetishism, de re, de dicto, conceptual role semantics

10

Deliberation and Moral Motivation

*Zoë Johnson King*

1. Introduction

This chapter makes three contributions to the literature on moral motivation.

First, I identify a kind of moral motivation whose significance has been overlooked because its possibility has been obscured: a motivation whose object is that the agent acts rightly, but that represents acting rightly under another description. This possibility has been obscured by a mistake in the way philosophers working on moral motivation have employed the *de re*/*de dicto* distinction. We use the phrase “motivation by rightness *de re*” to refer to motivations whose objects are right-making features. But the “*de re*” qualifier is supposed to indicate that, although our description of an agent’s attitude refers to the very same object as her attitude, nonetheless she may represent this object under another description (as in Frege puzzles). And the relationship between right-*making* features and rightness is “making”, not identity. So motivations whose objects are these features are not motivation by rightness *de re*, properly so-called; rather than a single entity with two names, we have two distinct entities that are metaphysically related. Correctly construed, the phrase “motivation by rightness *de re*” refers to motivations whose object is that the agent acts rightly, but that may represent moral rightness—not the right-making features, but rightness itself—under another description.

Second, I argue that a familiar variety of moral deliberation cannot be fully explained by the hypothesis that deliberators just care about the particular morally significant things that they see at stake in their circumstances. I focus on cases in which the agent can tell that multiple important things are at stake and that they call for different kinds of responses from her, such that anything she might do would promote each of them to some extent but nothing she can do would promote them all fully. In such a case, we could just choose randomly. But we do not usually do that. Instead, we try to assess the relative importance of the things at stake and to identify the normative relationships that they bear to one another, pausing periodically to ask ourselves whether there may be more things that matter at stake than those we have so far taken into account. I arguethat we can only explain these episodes of deliberation if we take deliberators to have an overarching concern—one that goes beyond their concerns for the particular morally significant things at stake, and that is about striking the right balance between everything at stake that matters.

Third, I pose a trilemma for critics of motivation by rightness *de dicto*, assuming that many quotidian and innocuous motivations count as motivation by rightness *de re* (correctly construed). Philosophers who criticize motivation by rightness *de dicto* should clarify what it is to which they object: their target could be the object of this motivation, its content, or something else about it. The latter kind of criticism is not specific to motivation by rightness *de dicto* and is thus beside the point. Meanwhile, criticisms of any motivation whose object is that the agent acts rightly face death by counterexample if it turns out that myriad innocuous motivations count as motivation by rightness *de re*. And criticisms that target the content of motivation by rightness *de dicto* face an explanatory challenge: if there is nothing wrong with being motivated to act rightly under another description, then why would there be something wrong with being motivated to act rightly under this very description?

These three contributions are logically independent. But lurking in the background is a fourth, potentially more significant, contribution to the literature. For there is reason to think that the familiar kind of motivation described in Section 3 is an instance of the overlooked kind identified in Section 2, triggering the trilemma set out in Section 4. That is to say: there is reason to think that the motivations that drive the sort of deliberation I am interested in are instances of motivation by rightness *de re* (correctly construed). This is because people engaged in these episodes of deliberation often conceive of their aim in non-moral or not-obviously-moral terms, and yet, from our theoretician’s point of view, we may still identify the object of their motivation as <that they act rightly>. We may do this if normative terms’ conceptual role is at least a significant part of what determines their reference, since, in deliberators’ minds, the judgment that a certain act strikes the right balance often plays the rightness-role. Thus the chapter ends with a metasemantic conjecture: the true theory of moral metasemantics counts the motivations of these deliberators as among those whose object is that the agent acts rightly. If this is correct, then the chapter’s other contributions jointly identify a trilemma for critics of motivation by rightness *de dicto*: they are challenged by the fact that there is a way of being motivated to act rightly (*de re*, correctly construed) that is ubiquitous and apparently unproblematic.

2. Moral Motivation

Philosophers of language and mind use the *de re*/*de dicto* distinction to distinguish two ways we can describe people’s mental states. Here’s an illustration of the distinction as it is standardly drawn: suppose that Lois Lane wants to interview Superman, and Superman is in fact mild-mannered Clark Kent, but Lois does not know this. Does Lois want to interview Clark Kent? The answer depends on how the question is construed. Lois does not have a motivation with the content “I interview Clark Kent”. And, if asked whether she wants to interview Clark Kent, she would say that she does not—her mild-mannered colleague wouldn’t be interesting to her readers. So Lois is not motivated to interview Clark Kent *de dicto*; that is, not under the description given by the name “Clark Kent”. But we know that Lois is motivated to interview Superman. And we know that Superman is, in fact, Clark Kent—those are two names for the same person. So we know that Lois is motivated to interview someone who is, in fact, Clark Kent. We can therefore say that Lois is motivated to interview Clark Kent *de re*; that is, the man himself.

Specialists disagree about how best to understand what is going on here. But the basic idea is that we use the “*de re*” and “*de dicto*” qualifiers in attitude-reports to indicate whether we mean to convey something about the way someone thinks of the object of her attitude (*de dicto*), or just to say what that object is (*de re*). The “*de re*” qualifier signals that the object of the agent’s attitude is the entity to which our description refers, but that she herself may think of it in a different way than our description suggests—as in the Superman example. By contrast, the “*de dicto*” qualifier signals that we want to convey something about the way the agent’s attitude represents its object. The traditional way to put this is to say that her attitude represents its object “under a description” or under a “mode of presentation” that matches the terms in our attitude-ascription. But some contentious assumptions about mental content, on which I do not wish to take a stand in this chapter, are built in to this way of thinking. So, for present purposes, I use the phrase “under a description” to talk about whatever is going on in *de dicto* attitude-reports, while remaining neutral as to precisely what this is.

Ethicists and metaethicists sometimes apply the *de re*/*de dicto* distinction to moral motivation, following this widely cited passage in Michael Smith’s *The Moral Problem* (1994, p. 75):

Good people care non-derivatively about honesty, the weal and woe of their children and friends, the well-being of their fellows, people getting what they deserve, justice, equality, and the like, not just one thing: doing what they believe to be right, where this is read *de dicto* and not *de re*. Indeed, commonsense tells us that being so motivated is a fetish or moral vice, not the one and only moral virtue.

There are two ways to read Smith here. First, we might take the phrase “what they believe to be right” to refer to token acts—those that the agent believes to be right. On this reading, Smith is distinguishing among agents who are motivated to perform particular acts that they believe to be right. He distinguishes those whose motivation represents the act under the description “doing what I believe to be right” from those whose motivation represents it under another description, such as “being honest” or “spreading rumors” or “upsetting Karen”. Alternatively, we might take the phrase “what they believe to be right” to refer to types of act. On this reading, Smith is distinguishing among agents who have general motivations to perform actions of a certain kind. He distinguishes those with a general motivation whose content is that they perform acts that they believe to be right from those for whom there is at least one type of act such that the agent (a) has a general motivation to perform acts of this type and (b) believes that acts of this type are right. For example, the latter sort of agent may think that honest acts are right and also have a general motivation to be honest.

Both readings are somewhat unsatisfying. On the token-act reading, it is difficult to square the distinction drawn at the end of Smith’s sentence with his immediately preceding claim that good people care about such things as honesty, equality, and justice. None of those things are token acts. So, if the motivations that constitute motivation to do what one believes to be right *de re* are motivations to perform token acts, then they have the wrong sort of object to appear on Smith’s list of the motivations that are the mark of a good person. Moreover, Smith says that good people’s concerns are “non-derivative”, but the idea of a non-derivative motivation to perform a token act seems odd—like an arbitrary compulsion. Our motivations to perform particular acts usually derive from general motivations to perform acts of certain types and beliefs that the acts are of one or more of these types.[[1]](#footnote-2) This pushes us toward the type-of-act reading. On that reading, the general concerns that Smith lists as the concerns of good people might constitute motivation to do what one believes to be right *de re*, so long as the agents also have the corresponding beliefs (i.e., beliefs that honest acts are right, that acts that prevent “weal and woe” are right, and so on). However, on this reading, all manner of nefarious motivations also count as motivation to do what one believes to be right *de re*. For example, some people care non-derivatively about promoting white supremacy and believe that acts promoting white supremacy are right. These motivations count. In general, if someone believes that acts with a feature that she cares about are right, on this reading she counts as motivated to do what she believes to be right *de re*, even if the feature is in fact horrific. Thus, motivation to do what one believes to be right *de re* could be good or it could be horrific.

In search of a genre of motivation that seems incontrovertibly good, contemporary ethicists sometimes part ways with Smith and take the agent’s beliefs out of the picture. It is now less common to talk of agents who are motivated to do what *they believe* to be right *de re* and more common to talk of those who are motivated to do what *is* right *de re*. And the scholars who use this locution assume that we can make sense of it by taking the type-of-act reading of Smith and simply omitting the part about the agent’s beliefs; just as we say that someone is motivated to do what she *believes* to be right *de re* if there is a feature such that she is motivated to perform acts with this feature and believes that acts with this feature are right, similarly, the idea is that we should say that someone is motivated to do what *is* right *de re* if there is a feature such that she is motivated to perform acts with this feature and it is in fact among the right-making features. Many philosophers now think that this kind of motivation is what makes someone a good person; inspired by recent discussions of the case of Huckleberry Finn, we find it intuitive that what makes someone good is that she cares about what are in fact the right-making features, whether or not she sees them as such. So, this view does a better job of identifying motivation by rightness *de re* with something that seems incontrovertibly good than either reading of Smith. My guess is that this explains the view’s contemporary popularity.

But this view misuses the “*de re*” qualifier. To repeat: the qualifier is supposed to be used to signal that the object of an agent’s attitude is the very same entity to which our description refers, though she herself may consider it under a different description—as in the Superman example. And contemporary usage says that someone is motivated by rightness *de re* if she is motivated to perform acts with features that are in fact right-making. But right-making features are not the property of moral rightness considered under another description. If rightness were identical to the right-making features, then acts would be made right by their rightness, which is incoherent; it cannot simply be <that A is right> that makes <that A is right> the case.[[2]](#footnote-3) Rather, the term “making” in the phrase “right-making feature” picks out an asymmetric metaphysical relation—the “makes it the case” relation.[[3]](#footnote-4) So, when we use the phrase “motivation by rightness *de re*” to refer to a motivation whose object is a right-making feature, we use the “*de re*” qualifier as if it meant something like “Look down the Great Chain of Being, and look below the thing I just mentioned!”, rather than meaning what it usually means. This is not how the qualifier is supposed to be used. Superman doesn’t make Clark Kent the case; he *is* Clark Kent.

The upshot is that, in the attempt to identify a variety of motivation that is clearly good, ethicists and metaethicists have used the phrase “motivation by rightness *de re*” incorrectly. This is not a deadly sin. But, once we recognize it, we can ask what the phrase would refer to if we used it correctly. Following the Superman example, the answer is clear: if the “*de re*” qualifier were used correctly, then “motivation by rightness *de re*” would refer to motivations whose object is that the agent acts rightly, but that might represent the property of moral rightness under another description.

The possibility of this sort of motivation has been obscured by our terminological confusion. Contemporary discussions of moral motivation often proceed as if the only phenomena to be discussed are an explicit concern with acting rightly, which we call “motivation by rightness *de dicto*”, and the various motivations whose objects are right-making features, which we call “motivation by rightness *de re*”. But this is too narrow a range of options. Just as Lois Lane can be motivated to interview Clark Kent under another description, so too can ordinary folk be motivated to act rightly under another description. Thus, the contemporary literature contrasts two species of moral motivation that are not the only games in town.

It is also fairly common for us to use the phrase “*de re*” as shorthand for “*de re*, but not *de dicto*”. But, technically, the attitude-ascriptions that have true *de dicto* readings are a proper subset of those that have true *de re* readings. For example, Lois wants to interview Superman *de re*, since she wants to interview the person who is in fact Superman—indeed, she wants to interview him under this very description! Similarly, if we use the qualifiers correctly, then motivation by rightness *de dicto* is a species of motivation by rightness *de re*. This point can be missed when ethicists and metaethicists talk about people “who are motivated by rightness *de dicto* but not *de re*”, meaning people who are motivated to act rightly but not motivated by what are in fact the right-making features. In fact, it is impossible for someone to be motivated by something *de dicto* without being motivated by it *de re*.

3. Deliberation

Enough about moral motivation for now. Let’s talk about deliberation for a section.

Picture the scene: in a café, you aren’t sure whether it would be better to get the vegan sandwich or the one made from local ingredients. You arranged to meet up with your cousin later, and you’re exhausted from work and stressed about a meeting that went terribly, but you already canceled your last plan with your cousin and can’t tell whether this is a good enough reason to cancel again. One of your friends is texting you about a tumultuous breakup while another is texting about a sudden bereavement, and you don’t know who you should reply to first. That reminds you: another friend seems to be mistreating their new partner, but you aren’t close to the partner and aren’t sure whether it’s your place to intervene. You’re in a rush on your way to teach when the person in front of you in line at the café suddenly bursts into tears.

These examples are intended to be five realistic, quotidian illustrations of a familiar kind of case: someone can tell that multiple important things are at stake in her circumstances and that at least one of them calls for one sort of response, while at least one other calls for a wholly different sort of response. In such cases, we usually deliberate about what to do. We do this initially by coming up with actions we might perform and thinking about how well they fare when evaluated with an eye to each consideration at stake. Canceling on your cousin saves energy for the post-mortem conversations about the meeting that will happen tomorrow, but it’s disrespectful of your cousin’s time and might convey the impression that she isn’t important to you. Going out honors your commitment to your cousin and brightens her day, but leaves you in a worse position to navigate those difficult conversations. Changing the plan to something less time-consuming, or calling your cousin for a chat, might fare moderately well with respect to everything at stake and thus be a good compromise. And so on.

Call cases like this *conflict cases*. In conflict cases, the deliberator could in principle survey every act that she can perform (specified at an appropriate level of grain) and assess the precise extent to which each of them would promote each of the various things at stake.[[4]](#footnote-5) She could draw a little table in her head or on a napkin to jot down the results. But, having completed these tasks, she would still face a further question: given *everything* at stake, and given the extent to which each available act would promote each thing, what should she do?

In many conflict cases, it is clear to us that anything we might do would promote each thing at stake to some extent, and equally clear that nothing we can do would promote them all fully. Having realized this, we might give up, shrug our shoulders, and flip a coin—saying something like, “Oh well, you can’t win ʼem all!”, or, “Oh good, no matter what I do I’ll promote everything important to some extent!”. But we do not usually give this flippant response. And there seems something unsavory about it. The flippant response seems callous, and perhaps even reckless; at the very least, it seems inappropriately insensitive to the fact that multiple important things are all at stake and are pulling you in different directions. If you decided to cancel on your cousin or to deprioritize your bereaved friend’s texts on the basis of a coin-flip, it would seem as though your cousin or friend had legitimate grounds for complaint.

Instead of giving the flippant response, we usually engage in what I will call *continued deliberation*. Having determined the extent to which each available act promotes each thing at stake, we do not stop there. Instead, we draw detailed comparisons between these things, assess their relative importance, and try to figure out which normative relationships—such as relations of defeat among reasons, lexical priority among duties, and realization among values—they bear to each other. We think about whether it’s more important to reduce animal suffering or lower our “food miles”. We think about the conditions under which people are released from obligations to follow through on prior commitments. We think about how to compare one friend’s greater need to the other’s shortage of alternative confidantes. We think about whether we have the standing to tell someone how to conduct their romantic relationships if doing so protects an apparent victim. We try to come up with a way to respond to this crying stranger that helps them without making us inexcusably late for class. And so on.

Some conflict cases involve ample time for continued deliberation, while others place us under severe time constraints. Sometimes deliberation is cut short and we are forced to act before reaching a verdict about what we should do in light of everything at stake and the various degrees to which and ways in which they all matter. On these occasions, we typically feel rushed and frustrated, as though we did not get to finish what we started. But, when we have time, we sometimes *go looking for further morally significant considerations* before reaching our verdict. This means that we reflect carefully on our circumstances, aiming to identify the aspects that matter as exhaustively as we can, and then to identify the sort of response that each thing that matters calls for from us. Indeed, we do something like this all the time, though usually implicitly; as we go about our days we are receptive to evidence indicating that something that we know matters might be at stake or that something that we know to be at stake might matter. But during continued deliberation we periodically engage in this sort of monitoring more actively. We ask ourselves something like, “OK, and is there anything else that I should take into account?”, mentally scanning the normative landscape for further factors to consider before reaching our overall verdict about what we should do.

I take it that these phenomena will seem quotidian and familiar to anyone who sometimes finds themselves in conflict cases of the sort just described. This is because the phenomena *are* quotidian and familiar. Continued deliberation is a ubiquitous part of everyday moral life.

The question I now want to ask is: What motivates these deliberative episodes? Why do we engage in them?

One answer is that people engage in continued deliberation because they want to do the right thing. Some evidence for this natural answer is that deliberators often say things like, “Oh, I really want to do the right thing—I just wish I knew what it was!”. And, once they settle on a course of action, when asked why they are doing what they are doing they often reply “Because it’s the right thing to do”. Thus it often looks as though they engage in continued deliberation because they want to act rightly.[[5]](#footnote-6)

But motivations to act rightly *de dicto* get a bad rap. As we have seen, Smith says that “common sense tells us” that this motivation is “a fetish or moral vice”. Similarly, Jamie Dreier says that this motivation can seem “disgusting” (2000, p. 624), Jonas Olson that it seems “suspect” and “perverted” (2002, pp. 90–1), and Julia Markovits that it seems “plainly cold” (2010, p. 204). And that’s just some people whose names start with J. The full list of philosophers who have expressed aversion to motivation by rightness *de dicto* is long. Relatedly, some authors have suggested that this sort of motivation could only manifest in bizarre and unappealing ways. For example, Teemu Toppinen says that someone motivated by rightness *de dicto* will respond to any temptations to act contrary to what she thinks is right by “thinking furiously about the rightness of the act in and of itself”, rather than thinking about its right-making features (2004, p. 312). And David Shoemaker ([2007](#B13), pp. 88–90) gives us the following characterization:

They would be just like psychopaths, stricken by a similar sort of tunnel vision, albeit with respect to a different set of goals. The psychopath is focused on his own immediate needs or desires, so the general reasons of morality do not matter to him and thus will always lose out when they conflict with his needs or desires (as they often do). The fetishist, in contrast, does care about morality—he is obsessively devoted to it, after all—but here it is the second-personal reasons of his fellows that do not matter to him, and thus it is those reasons that will always lose out when they conflict with what he takes to be the demands of morality generally. Insofar as he would be a kind of bloodlessly calculating “morality machine,” then, his incapacity for the kind of interpersonal relationships constitutive of membership in the moral community would leave him, as with the psychopath, external to it.

In short, plenty of philosophers stress how averse they are to people motivated to act rightly and how weird such people must be. These authors make being motivated to act rightly sound awful. And that looks like bad news for those of us who regularly engage in continued deliberation. For if we do so because we are motivated to act rightly, then it looks as though we are subject to all these criticisms—we are cold, furious, disgusting morality machines.

But can this be? Most people who engage in continued deliberation in conflict cases don’t seem as freakish and disturbing as these caricatures suggest. This may be because there is something wrong with the caricatures (as I will later argue). Or it may be because, contrary to appearances, ordinary deliberators are not motivated to act rightly. And that is exactly what some philosophers have recently argued; they have argued that we can fully explain deliberation by taking deliberators to be solely concerned with the particular morally significant things that they see are at stake in their circumstances, and that a further motivation to act rightly is not needed.

For example, here is Nomy Arpaly ([2015](#B2), pp. 148–9):

People who ask themselves “What is the right thing to do?” often start asking and deliberating because something bothers them that is genuinely morally relevant. They wonder, say, whether to tell someone his spouse is cheating—displaying concern for both *truthfulness* and *another person’s well-being*. To use another example, suppose our deliberator wants to have an abortion but wonders whether or not it would be right. Suppose what causes her to stop and deliberate is the risk that by having an abortion she would be *killing a person*. One way to put it would be that the deliberator wonders whether *abortion is murder*—she already knows that murder is wrong, so her view of morality is not wide open. Thus the commonsense deliberator’s decision to stop and deliberate speaks well of her. So does the deliberation itself: whether she wonders if a fetus is a person or ends up weighing the possible personhood of the fetus against her possible *right to control her body*, she is considering morally relevant things . . . In short, there is no value in moral fetishism—but a real, live moral fetishist is hard to find!

And here is Brian Weatherson ([2014](#B17), pp. 160–1):

Sometimes we stop and think, what would be the best thing to do in a certain kind of case? . . . I agree this is a good practice . . . But [consider] an observation also by Smith ([1994](#B15), pp. 40–1), that moral inquiry has “a certain characteristic coherentist form”. I think (not originally) that this is because we’re not trying to figure out something about this magical thing, the good, but rather because we’re trying to systematise and where necessary reconcile our values. When we’re doing moral philosophy, we’re often doing work that is more at the systematising end, trying to figure out whether seemingly disparate values have a common core. When we’re trying to figure out what is right in the context of deciding what to do, we’re often trying to reconcile, where possible, conflicting values. But as long as we accept that there are genuinely plural values, both in moral and prudential reasoning, we shouldn’t think that a desire to determine what is right is driven by a motivation to do the right thing, or to live a good life, as such.

These authors hold that we engage in continued deliberation in conflict cases because we care about the things that conflict in the cases. What motivates us is an assortment of concerns with being truthful, promoting well-being, not murdering, respecting rights, and so on.

But that cannot be the whole story. For these specific concerns, by themselves, explain almost none of the characteristic phenomena of continued deliberation. Indeed, it is unclear why someone would engage in continued deliberation at all if she were concerned *solely* with the particular things that she sees at stake in her circumstances. To repeat: in these deliberative episodes we draw comparisons between the things at stake, assess their relative importance, and try to figure out what normative relationships they bear to one another. If someone *only* cared about truthfulness and well-being (say), in and of themselves, then why would these things’ relative importance or the normative relationships between them matter to her? For such an agent, learning of the conflict amounts to learning that she cannot get everything she wants. So one might expect that, rather than engaging in continued deliberation, she would give up and flip a coin.

It is also difficult for Arpaly and Weatherson to explain why we feel dissatisfied if deliberation is cut short and we are forced to act before reaching an overall judgment about what we should do given everything at stake. We feel this dissatisfaction even if we choose an act to perform and are sure of the precise extent to which it promotes each thing at stake. Indeed, we knew (or were in a position to know) all this about our chosen act long before we began continued deliberation: we could tell this much about the act just by looking at its row in the table on our mental napkin. Yet, when continued deliberation is cut short, we react as if there is something else that we were aiming at and are not sure we have achieved. This is hard to understand if we were all along aware of the precise extent to which our now-chosen act promotes each of the things we care about.

Lastly, Arpaly and Weatherson cannot explain the parts of continued deliberation in which we go looking for further morally significant considerations. The hypothesis that deliberators are motivated by concern for the particular morally significant things that they see at stake in their circumstances does not explain this at all. For, when we go looking for further morally significant considerations, we do so because we worry that there may be things that matter at stake *besides* those we currently see, and we want to discover any such things and then take them into account in our overall verdict. This clearly cannot be explained by our concerns for the particular things at stake that have so far occurred to us, since our concern extends beyond them.

Here is my suggestion: we engage in continued deliberation because we have an overarching concern, over and above our concerns for each particular thing that matters at stake in our circumstances. For the sake of a label, we can say that we are motivated to *strike the right balance* between everything at stake that matters, in light of the various degrees to which and ways in which they all matter and the different responses that they accordingly call for from us. However, this label does not mean that ordinary deliberators conceive of their overarching concern in precisely these terms. On the contrary, ordinary deliberators often do not consider the nature of their overarching concern at all. And, when they do, they often articulate their aim in non-moral or not-conspicuously-moral terms. People without philosophy backgrounds rarely use the term “moral” and do not use it as moral philosophers do; the folk term connotes a high-minded, holier-than-thou, judgmental sort of attitude. Thus many ordinary folk who engage in continued deliberation cast their overarching concern in other terms. We describe ourselves as thinking about *what’s most important* or *what really matters*, or we say that we are *juggling things*, or that we are *just trying not to be a shitty person* or not to be a *dickhead* or an *asshole* or a *jerk* (depending on where we live), or similar. Similarly, I have referred to the considerations that deliberators try to balance as “morally significant considerations”, but they do not think of them in such terms. Nonetheless, it remains the case that their deliberative episode is motivated by an overarching concern that goes beyond their concerns for the various particular things at stake, and is about striking the right balance between them.

This hypothesis explains why we consider the relative importance of and normative relationships between the things at stake; doing so is part of trying to strike the right balance between them. It also explains why we feel frustrated when deliberation is cut short; we feel frustrated because we *are* frustrated, which is because we aimed to figure out how to strike the right balance and have been prevented from doing so. And it explains why we go looking for further morally significant considerations; we are motivated to strike the right balance between *every*thing that matters at stake in our circumstances, and not merely those things that have so far occurred to us.

Thus, the hypothesis that continued deliberation is driven by a concern to strike the right balance explains the phenomena much better than the hypothesis that it is driven solely by concerns for the particular things that deliberators see at stake in their circumstances. Of course, these are not the only possible explanations. I cannot survey every other possible explanation. So, I content myself with having challenged the main rival explanation that has some defenders in the contemporary literature and made the case for my hypothesis over theirs.

4. Deliberation and Moral Motivation

Motivation by rightness *de dicto* gets a bad rap, as we have seen. The literature is replete with criticisms and caricatures of this sort of motivation. But critics of motivation by rightness *de dicto* are sometimes unclear about what exactly it is that they are criticizing. They have three main options: the *object* of this motivation, its *content*, or something else about it—perhaps its strength in certain people, or the single-mindedness with which certain people may pursue it. Since the third option is just “something else”, this list is exhaustive.

Let’s start with the last option: criticizing features of motivation by rightness *de dicto* that are independent of its content and object. Critics who take this route may succeed in identifying something that seems weird or bad about certain motivations. But they fail to identify a specific problem for motivation by rightness *de dicto*. For example, consider the worry that motivation by rightness *de dicto* might make someone unattractively preoccupied with moral rightness. It is true that this motivation could manifest in this way; someone may tear her hair out over a minor moral problem, deliberating far longer than the gravity of the situation warrants or continuing past the point where further thought can be expected to yield new insight. But any motivation, with any content and object, could manifest in this excessive way. Notably, this holds of concerns for any right-making feature: for instance, someone may tear her hair out over whether a minor slight is disrespectful, deliberating far longer than the gravity of the situation warrants, or she may think about whether truthfulness requires telling the whole truth long past the point where further thought can be expected to yield new insight. And so on. There is a general risk of motivations’ becoming all-consuming. But this is not a specific problem with caring about acting rightly; it could happen to any motivation, so to speak. And the fact that a certain kind of motivation *could* manifest in an excessive way does nothing to impugn the status of the motivation’s ordinary, moderate versions. For example, the fact that OCD is an unwelcome condition clearly does not show that there is always something wrong with wanting to wash your hands.

Some caricatures of motivation by rightness *de dicto* face this objection. Consider Toppinen’s description of the person who responds to temptations to act contrary to what she thinks right by thinking “furiously” of the rightness of the act. This is no weirder than thinking “furiously” of an act’s fairness, of its truthfulness, of the fact that it promotes well-being, and so on, if in all cases the agent just thinks about the fact that the act possesses a property and gnashes her teeth. That would be a weird thing to do. But it is a weird thing to do whichever property the agent has in mind. Similarly, recall Shoemaker’s warning of the “tunnel vision” of the “bloodlessly calculating morality machine”. It does sound scary for someone to be indifferent to other people’s needs and rights whenever they conflict with what she takes to be the general demands of morality. But it sounds equally scary for someone to be indifferent to people’s needs and rights whenever they conflict with what she takes to be the more general demands of honesty, justice, equality, or whatever. So, the “bloodlessly calculating machine” caricature also describes a bad way that motivations can manifest, but one that is not specific to motivations to act rightly. But, once again: the fact that there could be creepy, unappealing versions of a motivation does nothing to impugn the status of its ordinary, moderate versions. We can have too much of a good thing.

In the literature following Smith, some philosophers discuss cases in which motivation by rightness *de dicto* is an agent’s sole motivation, while others discuss cases in which it is one of many motivations. Those in the former camp are typically also vulnerable to this objection: it seems weird and bad for someone to have just one motivation, no matter what its object and content are. Someone who only has a single motivation may be unattractively preoccupied with its object, but, as we just saw, these preoccupations are no less attractive when their object is rightness than when it is a right-making feature (or any other feature). And a similar point applies to those who worry that someone motivated solely by rightness *de dicto* might be mistaken about what is right, thus performing what are in fact morally atrocious acts in the name of her ostensibly virtuous goal. That is a reasonable worry. But it does not go away when we imagine an agent who is motivated by a right-making feature rather than by rightness. No feature is such that being motivated by it automatically renders the agent omniscient about its precise nature and extension; someone can care about well-being while being mistaken about the nature of well-being, can care about justice while being mistaken about what justice consists in, and so on. Any feature is such that someone motivated solely by it might be mistaken about which acts possess it, thus performing what are in fact morally atrocious acts in the name of her ostensibly virtuous goal. It is important to think about how to evaluate these well-meaning but morally mistaken agents.[[6]](#footnote-7) But, once again, what we have here is a potential problem with motivations that is independent of their content and object, rather than a problem that arises specifically for motivations to act rightly.

For their criticisms to target motivation by rightness *de dicto* particular, then, scholars must take one of the first two routes: they must criticize the object of a motivation—the fact that the agent is motivated to act rightly—or its content—the fact that she is motivated to act rightly under that very description.

This means that critics of motivation by rightness *de dicto* had better hope that motivation by rightness *de re*, correctly construed, is rare. For recall that the phrase “motivation by rightness *de re*” refers to motivations whose object is that the agent acts rightly but that may represent acting rightly under another description. Suppose that motivation by rightness *de re*, so construed, is ubiquitous and unproblematic—that is, suppose that many motivations that strike us as utterly quotidian and familiar turn out to count as motivation by rightness *de re*. If that is so, then critics of motivation by rightness *de dicto* who take either the first (object) or the second (content) route are in trouble. For those who claim that there is something objectionable about all motivations whose object is that the agent acts rightly face death by counterexample, given the quotidian and familiar instances of motivation by rightness *de re*. And those who say that there is something uniquely bad about motivation by rightness *de dicto* face an explanatory burden: they must explain why, if other varieties of motivation to act rightly are unproblematic, nonetheless there is something wrong with being motivated to act rightly *de dicto*. And it is difficult to see what could discharge this burden. That is, it is difficult to see why, if there is nothing wrong with being motivated to act rightly under another description, there would be something wrong with being motivated to act rightly under this very description. On the contrary, one would think that the mere fact that someone conceives of her motivation’s object in a particularly perspicuous way cannot turn her from a good person into a bad one.

So, if motivation by rightness *de re* (correctly construed) turns out to be ubiquitous and unproblematic, then critics of motivation by rightness *de dicto* face a trilemma: their view is straightforwardly false, they have an explanatory burden to discharge, or their criticisms miss the mark entirely.

Now, consider the following argument:

1. Someone is motivated by rightness *de re* if she has a motivation whose object is that she acts rightly, but that represents acting rightly under another description.
2. People engage in continued deliberation in conflict cases because they are motivated to strike the right balance between everything that matters at stake.
3. Engaging in continued deliberation in conflict cases is ubiquitous and unproblematic.
4. Being motivated to strike the right balance between everything that matters at stake is a way of having a motivation whose object is that one acts rightly, but that represents acting rightly under another description.
5. People who engage in continued deliberation in conflict cases have a motivation whose object is that they act rightly, but that represents acting rightly under another description. *(from 2, 4)*
6. People who engage in continued deliberation in conflict cases are motivated by rightness *de re*. *(from 1, 5)*
7. Motivation by rightness *de re* is ubiquitous and unproblematic. *(from 3, 6)*

I defended premise 1 in Section 2; I defended premise 2 in Section 3; and premise 3 is partially an easily confirmed empirical claim (continued deliberation is ubiquitous) and partially a moral claim whose denial would have drastic consequences for our evaluation of very many ordinary people very much of the time (continued deliberation is unproblematic). So, if premise 4 is also true, then it looks as though motivation by rightness *de re* is indeed ubiquitous and unproblematic. As well as being interesting in and of itself, this result would trigger the trilemma for critics of motivation by rightness *de dicto*.

Is premise 4 true? I think it might well be. The motivations that drive continued deliberation—motivations to strike the right balance between everything that matters at stake—are good candidates for being motivations whose object is that the agent acts rightly but that represent acting rightly under another description. After all, many ordinary deliberators conceive of their aim in non-moral or not-obviously-moral terms. The description “under which” they are motivated is given by these terms: *juggling everything* or *acting in a way that reflects what really matters*, or *not being an asshole*, or whatever. All that remains to be seen, then, is whether these thoughts in fact refer to acting rightly.

This is a question in moral metasemantics; the question is about what it takes for the property of moral rightness to be referred to by someone’s mental state. And moral metasemantics is a relatively young field. So, my conclusions in this section are tentative. But I am hopeful. That is because of the focus, within this field, on the contribution to reference made by moral judgments’ *conceptual role*. Judgments about what is morally right have a kind of priority in deliberation about what we should do all-things-considered and in evaluation of our own and others’ actions, and they have a close tie to the reactive sentiments. And many moral metasemanticists think that this conceptual role plays a central role in securing referents for the terms in which our moral judgments are cast. This, we think, must be the case in order to explain why agents in different communities manage to co-refer, and thus to substantively disagree rather than talking past one another, despite enormous variation in their external environments and their first-order moral beliefs. What explains why they disagree, the story goes, is that their judgments cast in moral terms play the same role within their cognitive architecture. In short: our terms count as referring to moral stuff in virtue of what we do with them.[[7]](#footnote-8) So, if your judgments cast in a term *t* have a certain characteristic kind of priority in deliberation and evaluation and a certain characteristic tie to the reactive sentiments, then your term *t* plays the conceptual role of my term “right”, and may thereby get to refer to rightness.

It is empirically verifiable that this is true of many deliberators’ thoughts about striking the right balance—or about *juggling everything*, or *acting in a way that reflects what really matters*, or *not being a shitty person*, and so on. Whatever description it is “under”, for deliberators, the judgment that a certain act strikes the right balance often plays the rightness-role within her cognitive architecture. This judgment is associated with the reactive attitudes (e.g., we feel guilty if we think that a certain act strikes the right balance but choose to do something else for selfish reasons), and it carries a lot of weight in deliberation about what to do all-things-considered and in evaluation of our own and others’ actions. Simply put: we engage in continued deliberation because striking the right balance matters to us in exactly the way that acting rightly matters to people motivated by rightness *de dicto*.

Here is my conjecture: the true theory of moral metasemantics will count the motivations that drive continued deliberation in conflict cases as among those whose object is that the agent acts rightly. This is a conjecture, since a full defense of a particular theory of moral metasemantics would be out of place in this chapter. But, as I’ve said, I’m hopeful. The importance of securing univocity across speakers has already led several metaethicists to embrace conceptual role semantics for moral terms (see e.g., Wedgwood [2001](#B18); Enoch [2011](#B6); Eklund [2017](#B5)), while others have found ways to accommodate the contribution of conceptual role to reference for moral terms within a classically Lewisian metasemantics (e.g., Dunaway and McPherson [2016](#B4); Williams [2018](#B19)). And there are views in moral metasemantics that do not place conceptual role front and center, but still might secure the result that ordinary deliberators’ thoughts are about moral rightness (partly) in virtue of their role in deliberators’ cognitive architecture. For instance, views that place the judgments about co-reference of informed speakers front and center might secure this result (see e.g., Schroeter and Schroeter [2014](#B12)), since informed speakers might recognize deliberators’ thoughts as co-referential with their rightness-thoughts in light of these thoughts’ conceptual role. Teleosemanticists (e.g., Sinclair [2012](#B14)) might surmise that, given this role, the concepts in which deliberators’ thoughts about striking the right balance are couched have developed to serve the same function as our concept of moral rightness. And so on. Since there are many promising ways for my conjecture to turn out to be true, I am content to tentatively rest my hat on it, punting the details to the metasemanticists.

Two points of clarification are in order. First: I do not expect that *every* time a deliberator thinks something like “I’m trying to juggle everything”, her thoughts will turn out to have acting rightly as their object. That is because there are many other balances between competing considerations that we can aim to strike, besides the morally right balance; we may aim to strike the *prudentially* right balance, the *aesthetically* right balance, the *all-things-considered* right balance, the right balance *between justice and honesty* (bracketing all other morally significant considerations), and so on. Determining which balance someone aims to strike, and thus what the object of her motivation is, will often be a tricky business, since the content of her thoughts—if it is limited to a minimal description like “juggle everything”—will not pin it down. Instead, we must examine the form that the agent’s deliberation takes and the conceptual role of her strikes-the-right-balance judgments. For instance, someone who simply aims to strike the right balance between justice and honesty will not go looking for any further morally significant considerations. And someone’s strikes-the-right-balance judgments will not be tied to the reactive sentiments if it is the aesthetically right balance, rather than the morally right balance, at which she aims. My contention is that conflict cases in which someone engages in continued deliberation because she is motivated to act rightly *de re* (correctly construed) are quotidian and ubiquitous. But I do not think that all cases of deliberation whatsoever fall into this category. We can determine precisely how many and which cases fall into this category by examining deliberators’ judgments’ conceptual roles.

Second: on some versions of conceptual role semantics, sameness of conceptual role secures sameness of *referent*—that is, two terms with the same conceptual role refer to the same object—but on other versions, sameness of conceptual role secures sameness of *concept*—that is, two terms with the same conceptual role have the same conceptual content. For example, the terms “tipping birds” and “single-leg deadlift” may have the same conceptual content in virtue of our judgments couched in these terms’ being tokened in response to the same exercise, leading us to perform this exercise, and so on. If the terms in which ordinary deliberators conceive of the object of their deliberation turn out to have the same conceptual *content* as the term “morally right”, this will be an interesting and surprising result. For it will mean that ordinary deliberators are motivated by rightness *de dicto* after all; the concept of moral rightness is tokened in their deliberation, though they token it by means of a less clear and more convoluted phrase than “moral rightness”. This is an even worse result for those, like Arpaly and Weatherson, who want to resist the thought that ordinary deliberators are motivated by rightness *de dicto*. For it means that ordinary deliberators are indeed motivated by rightness *de dicto*, and so they are wrong. So, I am not worried about this possibility. But I take myself to have shown that, even on a weaker version of conceptual role semantics on which sameness of conceptual role only guarantees sameness of referent, we should still expect that many ordinary deliberators will qualify as motivated to act rightly—not *de dicto*, but *de re*, where the “*de re*” qualifier is used correctly.

5. Conclusion

Let’s review.

I have issued a novel challenge to the familiar view that there are two types of moral motivation, motivation by rightness *de dicto* and motivation by rightness *de re*, and that there is something amiss with motivation by rightness *de dicto* that is not similarly amiss with motivation by rightness *de re*. I first argued that this is the wrong way to set up the options, noting that philosophers have applied the *de re*/*de dicto* distinction to moral motivation in a confused way that interprets the phrase “motivation by rightness *de re*” as referring to motivations whose objects are right-making features. I observed that if the “*de re*” qualifier were used correctly then this phrase would refer to a motivation whose object is that the agent acts rightly, but that may represent the property of moral rightness under another description. I then refuted the view that ordinary deliberation can be fully explained by supposing that deliberators just care about the particular right-making features they see at stake in their circumstances, arguing that many quotidian and familiar phenomena cannot be so explained. I argued that continued deliberation in conflict cases can only be explained by taking deliberators to have a further, overarching concern: a concern to strike the right balance between everything important at stake. Finally, I suggested that there is good reason to think that the misunderstood variety of motivation from Section 3 will often turn out to be an instance of the overlooked variety of motivation from Section 2, given the conceptual role of ordinary deliberators’ judgments to the effect that a certain act strikes the right balance. And I argued that this presents critics of motivation by rightness *de dicto* with a trilemma: their criticisms either face lots of counterexamples, bear an explanatory burden, or miss the mark entirely.

The upshot is that ethicists and metaethicists should be less cavalier with the *de re*/*de dicto* distinction, as applied to moral motivation. Careful work in the philosophy of mind and language, along with careful attention to the phenomena of everyday moral life, reveals that the sweeping statements we have couched in these terms cannot withstand critical scrutiny.

Another upshot is that many more people are motivated to act rightly than critics of this sort of motivation typically think. Those who identify motivation by rightness *de dicto* with a bizarre caricature have no trouble observing that few real agents approximate the caricature. But, if my argument is sound, then there are ways of being motivated to act rightly that look nothing like these caricatures. Indeed, assuming that you are a normal person who deliberates in conflict cases in the way that I have described, you too may be moral fetishist. Welcome to the club.[[8]](#footnote-9)

References

Aboodi, Ron (2017). “One Thought Too Few: Where *De Dicto* Moral Motivation is Necessary”. *Ethical Theory and Moral Practice* 20: 223–37.

Arpaly, Nomy (2015). “Huckleberry Finn Revisited: Inverse Akrasia and Moral Ignorance”. In Randolph Clarke, Michael McKenna, and Angela Smith (eds.), *The Nature of Moral Responsibility*, pp. 141–56. New York: Oxford University Press.

Dreier, James (2000). “Dispositions and Fetishes: Externalist Models of Moral Motivation”. *Philosophy and Phenomenological Research* 61(3): 619–38.

Dunaway, Billy and McPherson, Tristram (2016). “Reference Magnetism as a Solution to the Moral Twin Earth Problem”. *Ergo* 3(25). <https://doi.org/10.3998/ergo.12405314.0003.025>.

Eklund, Matti (2017). *Choosing Normative Concepts.* Oxford: Oxford University Press.

Enoch, David (2011). *Taking Morality Seriously: A Defense of Robust Realism.* Oxford: Oxford University Press.

Hare, Richard M. (1991). *The Language of Morals*. Oxford: Oxford University Press.

Horgan, Terence and Timmons, Mark (1991). “New Wave Moral Realism Meets Moral Twin Earth”. *Journal of Philosophical Research* 16: 446–65.

Johnson King, Zoë (2020). “Praiseworthy Motivations”. *Noûs* 54(2): 408–30.

Markovits, Julia (2010). “Acting for the Right Reasons”. *Philosophical Review* 119(2): 201–42.

Olson, Jonas (2002). “Are Desires De Dicto Fetishistic?”. *Inquiry* 45(1): 89–96.

Schroeter, Laura and Schroeter, François (2014). “Normative Concepts: A Connectedness Model”. *Philosophers’ Imprint* 14(25): 1–26.

Shoemaker, David (2007). “Moral Address, Moral Responsibility, and the Boundaries of the Moral Community”. *Ethics* 118(1): 70–108.

Sinclair, Neil (2012). “Metaethics, Teleosemantics and the Function of Moral Judgments”. *Biology and Philosophy* 27(5): 639–62.

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

Toppinen, Teemu (2004). “Moral Fetishism Revisited”. *Proceedings of the Aristotelian Society* 104(3): 305–13.

Weatherson, Brian (2014). “Running Risks Morally”. *Philosophical Studies* 167(1): 141–63.

Wedgwood, Ralph (2001). “Conceptual Role Semantics for Moral Terms”. *Philosophical Review* 110: 1–30.

Williams, Robert (2018). “Normative Reference Magnets”. *Philosophical Review* 127(1): 41–71.

1. Let’s not over-intellectualize this, though. For example, a motivation to draw one’s hand away from a fire can derive from an aversion to pain and a representation of the fire as painful. This can all happen sub-personally. The point is that motivations to perform particular acts are not intrinsic. [↑](#footnote-ref-2)
2. Thanks to Ralph Wedgwood for this way of putting the point. [↑](#footnote-ref-3)
3. Here I take no stand on exactly how moral grounding works. Any view according to which the fact that an act has a certain right-making feature (e.g., the fact that it is honest) partially or wholly makes it the case that the act is right, but is not identical to the fact that the act is right, serves my purposes. And these metanormative assumptions are shared by my primary interlocutors in this chapter. They think that ordinary deliberation is often motivated by a concern for what are in fact right-making features, but, to make this plausible in light of empirical facts about the diverse things by which people are in fact motivated, they must say (and do in fact say) that there are multiple right-making features and that these features overlap substantially with the things that ordinary people care about. And the multiple right-making features are plainly not identical to each other. So, given that identity is transitive, rightness cannot be identical to these features. Of course, it is possible to think that there is just one right-“making” feature and that it is in fact identical to rightness (as, for example, Moore seems to have done). Philosophers who say this are not my primary target in this chapter, although they are impaled on the content horn of the trilemma that I present in Section 4. Thanks to Francesca Bunkenborg for helpful discussion of this point. [↑](#footnote-ref-4)
4. I intend for the term “promote” to be read without consequentialist connotations, such that “promoting” something of moral significance need not amount to increasing the total quantity of it in the world and can encompass whatever is the appropriate way to respond to the thing in question. [↑](#footnote-ref-5)
5. Aboodi ([2017](#B1)) has argued that some kinds of moral deliberation can only be explained by taking the agent to have motivations with *de dicto* moral content. I am sympathetic to this argument. But, as will soon be evident, I think that we can go further—I think that some episodes of deliberation are driven by motivations with no obvious *de dicto* moral content, but that these motivations’ object may nonetheless be that the agent acts rightly. [↑](#footnote-ref-6)
6. I defend a proposal in Johnson King (2020). [↑](#footnote-ref-7)
7. This is also widely—though, of course, not universally—thought to be the lesson of the “Moral Twin Earth” thought-experiment (Horgan and Timmons [1991](#B8); cf. the similar but more racist “missionary and cannibals” case in Hare [1991](#B7)). [↑](#footnote-ref-8)
8. Previous versions of this chapter were presented at the Chapel Hill Metaethics Workshop 2018, a colloquium talk at Rice University, a virtual colloquium talk at Humboldt University, the Virtual Metaethics Workshop 2020, and the Madison Metaethics Wokshop 2020. I am grateful to participants in all these sessions for formative feedback. I am also grateful to Tom Dougherty, Daniel Fogal, Alex King, Nick Laskowski, Stephanie Leary, Ralph Wedgwood, Daniel Wodak, and Alex Worsnip, who commented on earlier drafts of the chapter, and to my dissertation committee members, who discussed the tiny seeds of these ideas with me long before they germinated. [↑](#footnote-ref-9)