

Deliberation and Moral Motivation

ZOË JOHNSON KING
New York University

Note to readers: This is a draft. Nonetheless, feel free to share it with wild abandon – just bear in mind, if you want to cite it, that it might change.

1. Introduction

In this paper I make three related-but-independent 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 the property of moral rightness under another description. To explain why this possibility has been obscured, I discuss a smorgasbord of confusions and mistakes in the ways philosophers working on moral motivation have applied the *de re/de dicto* distinction. Most importantly, I note that it is an error to use the phrase “motivation by rightness *de re*” to refer to motivations whose objects are right-making features.¹ This is because the “*de re*” qualifier is supposed to indicate that, although our description of an agent’s attitude refers to the same object as her attitude, nonetheless she may represent this object under another description (as in standard Frege puzzles). But 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 are dealing with distinct entities that bear a metaphysical relationship to one another. Correctly construed, the phrase “motivation by rightness *de re*” would refer to motivations that represent the property of moral rightness – not right-making features, but rightness itself – under another description.

Second, I challenge the popular view that ordinary deliberation can be fully explained by the hypothesis that deliberators just care about the particular morally significant things that they see are at stake in their circumstances. I argue that there is a quotidian, ubiquitous kind of deliberation that cannot be so explained. 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 these things to some extent but nothing she could do would promote them all fully. She might then shrug her shoulders and flip a coin. But we do not usually do that. Rather, we engage in what I call *continued deliberation*: we try to assess the relative importance of the things at stake and to identify the normative relationships (such as defeat or lexical priority) 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 thus far taken into account. I argue that we can only explain continued deliberation if we take deliberators to have an overarching concern – one that goes

¹ The other errors on the smorgasbord just serve to explain how scholars were pushed toward this mistaken usage.

beyond their concerns for the particular things that they see are at stake in their circumstances, and 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*, on the assumption that very many quotidian and innocuous-seeming motivations count as motivation by rightness *de re* (correctly construed). Philosophers who object to motivation by rightness *de dicto* should be clearer about whether their criticisms target the object of such a motivation, its content, or something else about it such as its potential strength or single-mindedness. The latter kind of criticism is not specific to motivation by rightness *de dicto* – it could happen to any motivation, so to speak – and is thus beside the point. Meanwhile, criticisms of motivations whose object is that the agent acts rightly face death by counterexample if there are innocuous-seeming motivations that count as motivation by rightness *de re*. And criticisms that target the content of motivation by rightness *de dicto* then face an explanatory challenge: if there is nothing wrong with being motivated to act rightly under another description, why should there be something wrong with being motivated to act rightly under a particularly perspicuous description?

These three contributions are independent of one another. The point about what the phrase “motivation by rightness *de re*” would refer to if the “*de re*” qualifier were used correctly holds even if the rest of the paper does not describe a motivation of this kind. The point about how to understand the motivations that drive continued deliberation holds even if these motivations’ existence lacks any further metaethical significance. And the points that critics of motivation by rightness *de dicto* should be clearer about what their objections target, but that this lands them in a trilemma if myriad ordinary motivations turn out to count as motivation by rightness *de re*, hold even the rest of this paper does not show that this assumption is satisfied.

Nonetheless, lurking in the background is a fourth, more significant, potential contribution to the literature. For there is reason to think that the ubiquitous kind of motivation described in the paper’s second section is an instance of the overlooked kind identified in the first section, triggering the trilemma set forth in the third section. That is to say: there is reason to suspect that the motivations that drive continued deliberation are instances of motivation by rightness *de re* (correctly construed). This is because ordinary folk engaged in continued deliberation frequently conceive of their aim in non-moral or not-obviously-moral terms, but, from our theoretician’s point of view, we may still identify the object of their motivation as <that they act rightly>. For there is good reason to think that normative terms’ conceptual role is at least a significant part of what determines their reference. And, in deliberators’ minds, the judgment that an act strikes the right balance often plays the rightness-role. Thus the paper ends with a metasemantic conjecture: the true theory of moral metasemantics, whatever it is, counts motivations to strike the right balance as among those whose object is that the agent acts rightly. If this is correct, then the paper’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

Bear with me through a technical, somewhat historical, first section. This section surveys and then criticizes the way the *de re/de dicto* distinction has been applied in the literature on 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 and bespectacled 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 may then say that Lois is motivated to interview Clark Kent *de re*; that is, the man himself.

Specialists disagree about how best to understand this kind of case. But the basic idea is that we use “*de re*” and “*de dicto*” qualifiers in attitude-reports to indicate whether we mean to say something about the way someone thinks of the object of her attitude (*de dicto*), or just to say what that object is (*de re*). Using the “*de re*” qualifier signals that the object of the agent’s attitude is the entity to which our description refers, though she herself may think of it in a different way than our description suggests – as in the Clark/Superman example. By contrast, using 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 a “mode of presentation” that matches the terms we use in our attitude-ascription. But some contentious assumptions about mental content are built in to this way of thinking, and I need not and do not wish to take a stand on those issues in this paper.² 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 it is.

Ethicists and metaethicists have begun to 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.

This passage admits different readings of the distinction that Smith is trying to draw, because the phrase “what they believe to be right” in the penultimate sentence can be read in (at least) two ways. First, it might refer to particular acts – those the agent believes to be right. Second, it might refer to kinds of act – the kinds such that the agent believes that acts of this kind are right.

On the particular-act reading, Smith is distinguishing among agents who have motivations to perform acts that they believe to be right. The agent’s motivation may represent such an act under the description “doing what I believe to be right”, or under some other description such as “being honest” or “spreading rumors” or “upsetting Karen”. The first would be a case of motivation to do what they believe to be right *de dicto*. The rest would be cases of motivation to do what they believe to be right *de re*; in these cases, there is an act such that the agent both believes it to be right and is motivated to perform it, but what she finds appealing about it is something other than its rightness.

The phrase “act that I believe to be right” picks out a kind of act: acts that the thinker believes to be right. So, when Smith writes of people who are care about “doing what they believe to be right” *de dicto*, we can

² The main issue is that it we can draw the *de re/de dicto* distinction when attributing mental states to thinkers who are not language-users, for whom talk of “descriptions” seems odd. Thanks to Tom Dougherty for discussion of this issue.

instead read him as talking about people with a more general motivation to perform acts of a certain kind: acts that they believe to be right. This general motivation will generate particular motivations to perform particular acts when the agent comes to believe that she believes them to be right. For the contrast in Smith's passage to make sense, we must then find a *de re* reading of the phrase "what they believe to be right" that still refers to kinds of act. The most natural way to do this is to say that someone is motivated to do what she believes to be right *de re* if there is at least one feature such that she believes that acts with this feature are right and she also has a general motivation to perform acts with this feature. For example, she may think that honest acts are right and also have a general motivation to be honest.

Smith's passage has in fact been read in a variety of ways. For example:

[The sentence 'Kalista desires to do what is right'] could mean that for each thing that is in fact right, Kalista desires to do that thing. Or it could mean that Kalista has a desire whose content is: to do whatever is right. Smith calls the first reading the *de re* reading and the second the *de dicto* reading.

...says Jamie Dreier (2000, p.621). And:

If I have a desire *de re* to do what is right, then I care about the thing or the act itself, in virtue of the characteristics that make it the right thing (to do) or the right act. If I have a *de re* desire to help the needy, then it is this thing I care about, I have an immediate desire, for example, to give money to the Red Cross. This is all there is to the intentional object of my desire.

... says Jonas Olson (2002, p.90). And finally:

The features that the person in question believes make actions right consist in features she thinks are had by actions that are right. Her desire to perform actions that have these features is consequently a desire to perform actions she judges to be right. However, it is a desire *de re*, not a desire *de dicto*. It is a desire to perform actions that have certain features, such as helping people in need.

...says Caj Strandberg (2007, p.253). Here Dreier takes the particular-act reading, and Strandberg the kind-of-act reading. Olson's reading is somewhere in between: his idea of an agent motivated to do what is right *de re* is of someone motivated to perform a particular act that she believes to be right — for example, giving money to the Red Cross — and for whom this desire is "immediate", yet somehow also held "in virtue of the characteristics that make it the right thing to do".

But these interpretations of Smith's passage are all riddled with problems. Start with Olson's reading. This sort of reading is popular; one often hears summaries of the lesson of Smith's passage to the effect that good people are motivated to perform particular acts that they believe to be right, where it is the acts' right-making features that appeal to the agent and explain her motivation to perform them. However, this cannot be what Smith had in mind, as this kind of motivation cannot appear on Smith's list of what "good people" care about. For he says that good people care *non-derivatively* about things like honesty, justice, and equality. And if someone is motivated to perform a particular act in virtue of its right-making features, then their motivation to perform the act is *derivative*: it derives from general motivations to perform acts with the relevant features (e.g. honest acts, just acts) and beliefs that the act in question has those features. Derivative

motivations are explicitly ruled off of Smith's list of what good people care about. So, if a motivation to do what one believes to be right *de re* is derivative, then it does not make the list.

What about non-derivative motivations to perform particular acts? Maybe motivation by rightness *de re* is "immediate", as Olson puts it, in the sense of being non-derived. In other words, maybe "all there is to the intentional object of my desire" really is <that I give money to the Red Cross>, rather than something like <that I give money to the Red Cross, so as to help the needy>. But this idea is even less promising. For it is hard to see how immediate desires to perform particular acts, divorced from all beliefs about the acts' moral significance, could be good at all. Non-derivative desires to perform particular acts sound like compulsions, which are not normally thought to be the mark of a good person. Moreover, it is still hard to see how these motivations could appear on Smith's list of what good people care about. For he says that good people care non-derivatively about things like honesty, justice, and equality. None of these things are particular acts. So if a motivation to do what one believes to be right *de re* involves non-derivative motivations to perform particular acts, then it is again hard to see how it could make the list.

We avoid these problems by taking the kind-of-act reading, as Strandberg does. Perhaps the non-derivative motivations of a good person are general motivations to perform acts with "features that she thinks are had by actions that are right", such as a motivation to perform acts that help people in need. I suspect that this is, in fact, what Smith had in mind; on this reading, if someone believes that (e.g.) honest acts, just acts, and equal acts are right, then she can have the cares that Smith says make her a good person and thus count as motivated to do what she believes to be right *de re*. However, on this reading, all manner of more nefarious motivations can 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. On this reading, as long as the agent believes that acts with the feature she cares about are right, she is motivated to do what she believes to be right *de re*, even if in fact the features are 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 hear talk of agents who are motivated to do what *they believe* to be right *de re* and more common to hear 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 kind-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 believes that acts with this feature are right and is motivated to perform acts with this feature, 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 it is in fact among the features that make acts right — whether the agent knows this or not — and she is motivated to perform acts with this feature. Many ethicists now think that this kind of motivation is what makes someone a good person; inspired by recent discussions of Huckleberry Finn, we find it intuitive that what makes someone good is that she cares about what are in fact right-making features,³ whether or not she recognizes them as such. So, this view does a better job of identifying motivation by rightness *de re* with something incontrovertibly good than all the readings of Smith just surveyed. My guess is that this explains the view's contemporary popularity.

But this popular contemporary view misuses the "*de re*" qualifier. To repeat: the qualifier is supposed to be used to signal that the object of the agent's attitude is the very same entity to which our description refers, though she herself may think of it in a different way than our description suggests — as in the Superman

³ Especially in Arpaly (2003) and Markovits (2010). For dissent, see Sliwa (2016) and Johnson King (2020).

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 among the right-making features. 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 itself, which is incoherent; it cannot simply be <that A is right> that makes <that A is right> the case. Rather, the term “making” in the phrase “right-making feature” picks out an asymmetric metaphysical relation — the “makes it the case” relation.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 are using 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. And that is not the standard use of the qualifier. Clark Kent doesn’t make Superman the case; he *is* Superman.

The upshot of all this discussion is that, in the attempt to identify a variety of motivation that clearly makes someone a good person, ethicists and metaethicists have confused ourselves and started to use the phrase “motivation by rightness *de re*” in an incorrect way. Once we see this, we can ask what this phrase would refer to if we used it in the correct way. Following the Superman example, the answer is clear: if the “*de re*” qualifier were used in the usual way familiar from philosophy of language, then “motivation by rightness *de re*” would refer to a motivation whose object is that the agent acts rightly, but that represents the property of rightness — not right-making features, but rightness itself — under another description.

The possibility of this sort of motivation has been obscured by our terminological confusion. Contemporary discussions of moral motivation now 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 we can now see that this is too narrow a set 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.

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 together 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 aren’t sure who you should reply to first. Another friend seems to be mistreating their new partner, but you aren’t close to the partner and can’t tell 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 coffee shop suddenly bursts into tears.

These five examples are intended to be 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,

⁴ 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 itself identical to the fact that the act is right, will serve my purposes here.

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 of the considerations at stake. Canceling on your cousin saves energy for the post-mortem conversations about the meeting that will no doubt happen tomorrow, and might thus help you to push for the reforms you've been arguing for, 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 up her day, but leaves you in a worse position to push for those reforms. Changing the plan to something less time-consuming, or calling your cousin to explain what has happened and have a shorter chat, might fare moderately well with respect to everything at stake and thus might be a good compromise.

Call cases like this *conflict cases*. In conflict cases, the deliberator could survey every act that she can perform (specified at an appropriate level of grain) and assess the precise extent to which each act would promote each of the various things at stake.⁵ She could even draw a little table in her head or on a napkin to record the results of her reasoning. 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 of these things, what should she do?

In many conflict cases, it is clear to us that anything we might do would promote each of the things 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 to settle what to do — saying something like, “Oh well, you can't win 'em all!”, or, “Oh good, no matter what I do I'll promote everything that matters to some extent!”. But people in conflict cases do not usually give this flippant response. And there seems something unsavory about it. The flippant response seems callous, and perhaps reckless; at the very least, it seems inappropriately insensitive to the fact that multiple important things are at stake and are pulling you in different directions.⁶ If you decided to cancel on your cousin or deprioritize your bereaved friend's texts on the basis of a coin-flip, then it would seem as though your cousin or friend had legitimate grounds for complaint.

Instead of giving the flippant response, people in conflict cases usually engage in what I will call *continued deliberation*. Having figured out the extent to which each available act promotes each thing at stake, we do not stop there. Instead, we draw detailed comparisons between the things, assess their relative importance, and try to figure out which normative relationships — such as relations of defeat among reasons, of lexical priority among duties, and of realization among values — they bear to each other. We think about whether it's more important to reduce animal suffering or to 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 against the other friend's closer relationship to us and shortage of other 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 suddenly-crying stranger that might help them, if indeed there is anything we can do to help them, without making us inexcusably late for class.

As these examples suggest, some conflict cases involve ample time for continued deliberation, while others place us under time constraints. Sometimes deliberation is cut short and we are forced to act before reaching

⁵ I intend for the term “promote” here to be read without its consequentialist connotations, such that “promoting” something of moral significance does not necessarily amount to increasing the total quantity of it in the world, and can encompass whatever is the appropriate way to respond to the morally significant thing in question.

⁶ For more on the idea of moral recklessness see Field (forthcoming).

an overall 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.

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 identify the sort of response that each thing that matters calls for from us. There is something like this that we do all the time, though usually implicitly; as we go about our days we are receptive to evidence 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?”, and we mentally scan the normative landscape for further factors to consider before reaching an overall verdict about what we should do.⁷

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. And I take it that this is because they *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 natural answer is that people engage in continued deliberation because they want to do the right thing in their circumstances. Some evidence for this natural answer is that deliberators often say things like, “Oh, I really want to do the right thing here — 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 sometimes reply “Because it’s the right thing to do”. Thus it looks as though they engage in continued deliberation because they want to act rightly.

But motivation by rightness *de dicto* gets a bad rap. As we have seen, Michael Smith says that “common sense tells us” that this sort of motivation amounts to “a fetish or moral vice”. Jamie Dreier says that this sort of motivation can seem “disgusting” (2000, p.624), Jonas Olson that it seems “suspect” and “perverted” (2002, pp.90-91), Julia Markovits that it is “plainly cold” (2010, p.**). And that’s just the people whose names start with J. The full list of authors who have expressed aversion to motivation by rightness *de dicto* is long. Relatedly, some philosophers 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 by thinking about the features of the act that make it right (2004, p.312). And David Shoemaker (2007, 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

⁷ To anticipate: I do not mean to suggest that we think of the further factors that we seek under the description “morally significant considerations”. We may think of them in some other way. I will discuss this point shortly.

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 are keen to 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; if we do so because we are motivated to act rightly, then 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 anywhere near as freakish and disturbing as these caricatures would suggest. This may be because there is something wrong with the caricatures (as I will argue later). Or it may be because, contrary to appearances, ordinary deliberators are not in fact motivated to act rightly. That is exactly what some philosophers have recently argued: they have argued that we can fully explain deliberation by understanding deliberators to be solely concerned with the particular important things that they see are at stake in their circumstances.⁸ According to these philosophers, particular concerns for the particular things at stake suffice to explain the phenomena associated with ordinary deliberation — a further motivation to act rightly is not needed.

For example, here is Nomy Arpaly (2015, 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, 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, pp. 40–41), 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

⁸ I use the factive “see” because, for these philosophers, there is an important difference between deliberators who are correct about which things that matter are at stake in their circumstances and those who are in correct — and, among the latter, an important difference between those who are incorrect about what is at stake (a non-moral mistake) and those who are incorrect about what matters (a moral mistake). Clearly, there are also plenty of deliberators who are incorrect about what is at stake or about what matters. But, for the sake of argument, in this paper I am happy to restrict our focus to cases in which the deliberator is correct.

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 not a concern with acting rightly, but an assortment of concerns with being truthful, promoting well-being, not murdering, respecting rights, and so on.

But this 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, if any, 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 have everything she wants. So one might expect that, rather than engaging in continued deliberation, she would at this point abandon deliberation 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 a 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 just by looking at the act's row in the table we drew 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 that we have achieved. This uncertainty 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. When we go looking for further morally significant considerations, we do so because there may be things that matter at stake besides those that we currently see, and we want to discover any such things and then take them into account in our overall verdict about what we should do. 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.

Someone could argue that, in addition to being concerned with the various particular things we see at stake, we have an overarching desire that we act in a manner that reflects the relative strengths of our degrees of concern for these things, which prompts continued deliberation.⁹ But this hypothesis fits poorly with the

⁹ Thanks to Nick Laskowski for this suggestion.

phenomenology. These deliberative episodes don't feel like introspection. When you're trying to figure out whether it's more important to reduce animal suffering or lower your food miles, your aim is to figure out which cause is *more important*, not which one you personally prefer. Likewise, trying to figure out whether you have the standing to confront a friend about their romantic relationship doesn't feel like trying to figure out how much you want to confront them – someone who thinks that she lacks the standing to blame often knows that she very much wants to do so, and is asking herself a different kind of question. Similarly, trying to figure out whether you have sufficient reason to cancel on your cousin feels different if there are moral reasons to conserve your energy than if you just don't want to go out; the latter involves thinking (in part) about what you want and how strongly you want it, whereas the former just involves thinking about what's most important. Moreover, as I've indicated, these deliberative episodes all center around trying to *figure something out*. They consist of careful moral reasoning, not self-discovery. And they leave open the possibility that, having figured out the relative importance of and normative relationships between all the things that matter at stake, we will then see that our degrees of motivation for these things are out of whack, and that we accordingly need to do some work on ourselves. This strongly suggests that we are not trying to figure out how much we personally happen to care about these things, but how important they are.

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. But it is not a concern with striking our preferred balance between these things. For the sake of a label, we might say that deliberators 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 starkly different responses that they accordingly call for from us. However, this terminological label does not mean that deliberators think about their overarching concern in precisely these terms. On the contrary, ordinary deliberators typically 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 in the way that moral theorists do; the folk term connotes a high-minded, judgmental attitude with an air of grandstanding.¹⁰ 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 similar. Similarly, I have referred to the considerations that deliberators try to balance as "morally significant considerations", but they surely 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, not their preferred balance, but 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, because we were trying to figure out how to strike the right balance and were prevented from doing so. And it explains why we periodically go looking for further morally significant considerations; we are motivated to strike the right balance between *everything* that matters at stake in our circumstances, not merely those things that have so far occurred to us.

Thus, the hypothesis that continued deliberation is driven by motivation to strike the right balance explains the phenomena much better than the hypothesis that it is driven solely by concerns for the particular things at deliberators see at stake in their circumstances. Of course, these are not the only possible explanations. I

¹⁰ For the notion of "moral grandstanding" see Tosi and Warmke (2016).

cannot survey every other possible explanation here. So I content myself with having challenged the only rival explanation that has some defenders in the contemporary literature, and having made the case for my hypothesis over theirs.

4. Deliberation and Moral Motivation

Motivation by rightness *de dicto* gets a bad rap, as we just saw. The literature is replete with caricatures of this kind of motivation and derogatory dismissals of people who are so motivated. But critics of motivation by rightness *de dicto* are sometimes unclear about what exactly it is that they are objecting to.¹¹ They have three main options: they may criticize the *object* of such a motivation, its *content*, or something else about it – perhaps its strength in certain people, or the single-mindedness with which certain people might pursue it. Since the third option here 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 well succeed in identifying something that seems weird or bad about some possible motivations. But they fail to identify a problem that pertains specifically to 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 so motivated 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 of the right-making feature(s): for example, 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 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 certain things – it could happen to any motivation, so to speak. And the fact that a certain kind of motivation *could* manifest in a weird or bad way does nothing to impugn the status of more ordinary versions of the motivation. The fact that OCD is an unpleasant 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* are vulnerable to this objection. For instance, consider Toppinen’s description of the person who responds to temptation to act contrary to what she thinks is right by thinking “furiously” of the rightness of the act. This is no weirder than thinking “furiously” of an act’s fairness, its truthfulness, 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 whatever 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

¹¹ Not only does this hold of philosophers working on moral motivation, but it also holds of philosophers working on many other topics, as Smith’s “fetishism” intuition is now well-known and so it is common for authors to invoke it as a premise in their arguments about other things. For example, Weatherson (2014) invokes the fetishism intuition in an argument about moral uncertainty, Markovits (2010) invokes the intuition in an argument about moral worth, and Hills (2009) invokes the intuition in an argument about moral testimony. As one might expect, philosophers who deploy the fetishism intuition in passing are typically even less clear about what they take themselves to be objecting to than those in the literature responding directly to Smith.

more 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, again, the "bloodlessly calculating machine" caricature describes a bad way that a motivation can manifest, but one that is not at all specific to motivations to act rightly. But, to repeat, the fact that there could be creepy, unappealing, excessive versions of a motivation does nothing to impugn the status of its ordinary, moderate versions. We can have too much of a good thing.

For criticisms that target motivation by rightness *de dicto* in particular, then, scholars must take one of the first two routes. They must either take themselves to be criticizing the object of a motivation – that is, the fact that the agent is motivated to act rightly – or its content – that is, 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 understood, is rare. To see why, recall my argument from section 1: correctly understood, the phrase "motivation by rightness *de re*" refers to motivations whose object is that the agent acts rightly but that represent acting rightly under another description. Now suppose that motivation by rightness *de re*, so construed, is ubiquitous and unproblematic – that is, assume that very 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. Those who claim that there is something objectionable about all motivations whose object is that the agent acts rightly would 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* that does not apply to motivation by rightness *de re* would face an explanatory burden: they would have to 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 might discharge this burden. It is difficult to see why, if there's nothing wrong with being motivated to act rightly under another description, there would be something wrong with being motivated to act rightly under that very description. On the contrary, one would think that the mere fact that someone conceives of the object of her motivation in a particularly perspicuous way should not 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. Either their view is straightforwardly false, they have an explanatory burden to discharge, or their criticisms miss the mark.

Why does this matter?

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 in their circumstances.
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 in one's circumstances 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. (2, 4)

6. People who engage in continued deliberation in conflict cases are motivated by rightness *de re*. (1, 5)
7. Motivation by rightness *de re* is ubiquitous and unproblematic. (3, 6)

I argued for premise 1 in section 1. I argued for premise 2 in section 2. 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* that we just discussed.

So, is premise 4 true? I think that it might well be. The sorts of 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. As I emphasized earlier, many ordinary deliberators conceive of their aim in non-moral or not-obviously-moral terms. So the description “under which” they are motivated will be given by these terms: *juggling everything* or *acting in a way that reflects what really matters* or *not being a shitty person*. All that remains to be seen, then, is whether these thoughts in fact refer to acting rightly.

This is a question in moral metasemantics; a question about what it takes for the property of moral rightness to be referred to by someone’s mental state. And moral metasemantics is a young field. So my conclusions in this section must be tentative. But I am hopeful. That is because of the extensive focus, within this field, on the contribution to reference made by moral judgments’ *conceptual role*. Moral judgments have a kind of priority and importance in deliberation about what we ought to do all-things-considered and in evaluation of our own and others’ actions, and they have a close tie to the reactive sentiments. And a growing number of moral metasemanticists think that the conceptual role of moral judgments plays a central role in securing referents for our moral terms. This, we think, must be the case in order to explain why agents in different communities manage to co-refer, thus substantively disagreeing rather than talking past each other, despite enormous variation in their external environments and in their first-order moral beliefs. What explains why they disagree, despite applying their moral terms to different things and thus tracking different lower-level properties, is that judgements cast in these 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.¹² So, if your judgments cast in a term *t* function to settle questions about what people ought to do, how we ought to evaluate actions and agents, and which reactive attitudes are fitting, then your term *t* plays the role of my term “right” in your cognitive architecture, and may thereby get to refer to rightness.

As an empirical matter, this seems to be true of many deliberators’ thoughts about striking the right balance – or, more specifically, 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 typically plays the conceptual role of a judgment that the act is right within her cognitive architecture. This judgment is associated with the reactive attitudes (e.g. deliberators feel guilty if they think that a certain act strikes the right balance but choose to do something else for selfish reasons), carries a lot of weight in deliberation about what to do all-things-considered and in evaluation of our own and others’ actions, and so on. Simply put: we engage in continued deliberation because striking the right balance matters to us in exactly the way that acting rightly matters to those motivated by rightness *de dicto*.

¹² This is widely thought to be the chief lesson of the “Moral Twin Earth” thought-experiment (Horgan and Timmons 1991, cf. the structurally similar but more racist “missionary and cannibals” example in Hare 1991).

Here is my conjecture: whatever the true theory of moral metasemantics turns out to be, it will count the motivations that drive continued deliberation in conflict cases as among those whose object is that the agent acts rightly. I offer this as a conjecture, since a full defense of a particular theory of moral metasemantics would be out of place in this paper. But I take it that there is good reason to be hopeful here. The importance of securing univocity across speakers with widely divergent beliefs about the extension of moral terms and similarly different external environments has already led a number of prominent philosophers to embrace versions of conceptual role semantics for such terms (see e.g. Wedgwood 2001, Enoch 2011, Eklund 2017). Others have found ways to accommodate the important contribution of conceptual role to reference within a more classically Lewisian metasemantics for moral terms (e.g. Dunaway and MacPherson 2016, Williams 2018). And there are views in moral metasemantics that do not place conceptual role front and center, but that still might end up securing the result that ordinary deliberators' thoughts are about moral rightness in virtue of the role such thoughts play in deliberators' cognitive architecture. For instance, views that place the judgments about co-reference of informed speakers front and center might end up securing this result (see e.g. Schroeter and Schroeter 2014), as informed speakers might recognize deliberators' thoughts as co-referential with their rightness-thoughts in light of these judgments' conceptual role. Teleosemanticists (see Sinclair 2012) might surmise that, given this role, the concepts in which deliberators cast their thoughts about striking the right balance are developed to serve the same function as our concept of moral rightness. And so on. Since there are so many ways for this conjecture to turn out to be true, I am content to tentatively rest my hat on it, punting the details to the metasemanticists.¹³

We need an important point of clarification at this juncture. 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. It must be possible for multiple terms to have the same conceptual content, as this holds of synonyms from across languages (the classic example is “dog” and “chien”). On some versions of conceptual role semantics, synonyms from across languages are like pairs of terms within a language with the same conceptual role.

If it turns out that the terms in which ordinary deliberators conceive of the object of their deliberation 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. That is to say, it will mean that it is the very concept of moral rightness itself that figures in the content of their deliberation, though they token it by means of a less clear and more convoluted phrase than “moral rightness”.¹⁴ This is an even worse result for those, like Arpaly and Weatherston, who want to resist the thought that ordinary deliberators are motivated by rightness *de dicto*; it means that ordinary deliberators are indeed motivated by rightness *de dicto*, and so they are straightforwardly 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 guarantees sameness of referent rather than sameness of concept, we should still expect that very many ordinary deliberators very much of the time will qualify as motivated to act rightly — not *de dicto*, but *de re*, where the “*de re*” qualifier is used correctly.

¹³ Shout-out to Preston Werner for teaching me most of what I know about metasemantics.

¹⁴ Part of why this is so surprising is that ordinary deliberators may deny that they are motivated to act rightly — saying something like “I don’t care about this morality stuff; I’m just trying not to be a jerk!”. But we usually do not offer *de dicto* attitude-ascriptions in cases in which the agent does not believe that the relevant term refers to the object of her attitude; we do not say that Lois is motivated to interview Clark Kent *de dicto*. So it is odd that someone can be motivated to act rightly *de dicto* while expressly denying that the phrase “acting rightly” refers to the object of her motivation.

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 represents the property of moral rightness – not right-making features, but rightness itself – under another description. I then refuted the popular 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 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 the second section will often turn out to be an instance of the overlooked variety of motivation from the first section, 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 a lot 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 – may reveal that the sweeping statements that we have couched in these terms cannot withstand critical scrutiny.

Another upshot, if everything that I have said so far is correct, 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 correct, then there are ways of being motivated to act rightly that look nothing like these caricatures. Thus, assuming that you are a normal person who deliberates in conflict cases in the way that I have described, you too are a moral fetishist. Welcome to the club.

REFERENCES

- Arpaly, Nomy (2003). "Moral Worth". *The Journal of Philosophy* 99(5): 223-245.
- 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.141-156. New York, NY: Oxford University Press.
- Dreier, James (2000). "Dispositions and Fetishes: Externalist Models of Moral Motivation". *Philosophy and Phenomenological Research* 61(3), 619-638.
- Dunaway, Billy and McPherson, Tristram (2016). "Reference Magnetism as a Solution to the Moral Twin Earth Problem". *Ergo* 3.
- Eklund, Matti (2017). *Choosing Normative Concepts*. Oxford, UK: Oxford University Press.
- Enoch, David (2011). *Taking Morality Seriously: A Defense of Robust Realism*. Oxford, UK: Oxford University Press.
- Field, Claire (forthcoming). "Recklessness and Uncertainty: Jackson Cases and Merely Apparent Asymmetry". Forthcoming in *Journal of Moral Philosophy*. URL: <https://philpapers.org/rec/FIERAU>
- Hare, Richard M. (1991). *The Language of Morals*. Oxford: Oxford University Press.
- Hills, Alison (2009). "Moral Testimony and Moral Epistemology". *Ethics* 120: 94-127.
- Horgan, Terence and Timmons, Mark (1991). "New Wave Moral Realism Meets Moral Twin Earth". *Journal of Philosophical Research* 16:446-465.
- Johnson King, Zoë (2020). "Accidentally Doing the Right Thing". *Philosophy and Phenomenological Research* 100(1): 186-206.
- Markovits, Julia (2010). "Acting for the Right Reasons". *Philosophical Review* 119(2): 201-242.
- Olson, Jonas (2002). "Are Desires De Dicto Fetishistic?". *Inquiry* 45(1): 89-96.
- Schroeter, Laura and Schroeter, Francois (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):639-662.
- Sliwa, Paulina (2016). "Moral Worth and Moral Knowledge". *Philosophy and Phenomenological Research* 93(2): 393-218.

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

Toppinen, Teemu (2004). "Moral Fetishism Revisited". *Proceedings of the Aristotelian Society* 104(3): 305-313.

Tosi, Justin and Brandon Warmke (2016). "Moral Grandstanding." *Philosophy and Public Affairs* 4(3): 197-217.

Weatherson, Brian (2014). "Running Risks Morally". *Philosophical Studies* 167(1): 141-163.

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.