

# De Dicto Concern in Cases of Ethical Vagueness

Billy Dunaway

University of Missouri–St. Louis

A person who avoids  $\phi$ -ing is *de dicto* concerned with morality just in case their reasons, or motivations, for not  $\phi$ -ing include the fact that  $\phi$ -ing is wrong, and do not include the fact that  $\phi$ -ing has features  $F_1, F_2 \dots F_n$ , where  $F_1, F_2 \dots F_n$  are the features that make  $\phi$ -ing wrong. Someone whose reasons or motivations for not  $\phi$ -ing include that fact that  $\phi$ -ing has each of the features  $F_1, F_2 \dots F_n$  is *de re* concerned with morality.

Use of the terms *de dicto* and *de re* to mark this difference in ways to be motivated morally is from Michael Smith (1994). Smith says:

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.<sup>1</sup>

---

<sup>1</sup>Smith (1994: 75). This general claim is motivated on the basis of examples, including the well-known example of the rescuer who has “one thought too many” before rescuing his wife rather than a stranger in Williams (1981: 18).

It is an appealing idea that someone who is *de dicto* concerned with morality, but not concerned with morality *de re*, is defective in some way. By not caring about the things that make actions right, agents who have only *de dicto* motivations appear not to appreciate what makes it the case that the action they perform is the one that is morally required. Even if they do the right thing, they don't do it in the right way. Someone who helps a down-on-their-luck family in need of a week's supply of groceries, and does so because this action promotes the recipients' happiness and well-being without incurring a morally comparable cost is better, in some way, than someone who helps the same family by buying the same amount of groceries, but does so only because they think helping is right.

Call this the DE DICTO DEFECT thesis. The thesis has two components: one is a *negative* component, claiming that there is a way one should not be; in particular one should not be motivated purely by *de dicto* concern for morality. The second is a *positive* component, which says how one should be instead: one's motivation should include, perhaps among other things, *de re* concern for the features that constitute moral wrongness. These claims are not unrelated: *de dicto* motivation is defective in an interesting way only if there is some other, non-defective way to be motivated in the relevant cases. Proponents of the DE DICTO DEFECT thesis do have an alternative they can point to, namely *de re* motivation. The non-defective motivation involves appreciating the facts that make it the case that one is required to act as one does. The negative and positive components of the DE DICTO DEFECT thesis thus form a natural package.

*De dicto* and *de re* concern for morality both involve a distinctive kind of

motivation. Since the motivation in question is distinctively *moral* motivation, it does not simply involve acting with moral properties (de re or de dicto) as somehow a part of one's motivation. Someone who is concerned with others' happiness and well-being, and thereby buys groceries for a needy family, does not exhibit de re *moral* motivation purely on this basis alone. If the agent treats failures in herself and others to promote others' happiness and well-being as simple failures of prudential rationality, then she is in some sense concerned de re for the properties that constitute moral rightness.<sup>2</sup>

What is missing in this case is concern for these properties in a way that not only guides action, but is also connected to morally relevant regulative thoughts and emotions of the agent. Presumably the grocery-buyer does not avoid defect simply by being de re concerned with the right-making properties, if instantiation of these properties motivates her to buy the groceries for distinctively non-moral reasons. She needs, roughly, to be disposed to feel a particular kind of moralized guilt at the thought of failing to buy the groceries, and to direct related blame-like reaction toward others who fail to buy groceries in similar circumstances.<sup>3</sup> I will call these attitudes *moralized reactive attitudes*, or *reactive attitudes* for short. In principle someone could be motivated de dicto by moral properties, yet fail to have their actions count as morally motivated, for the same reasons.

The DE DICTO DEFECT thesis has been wielded as a premise in arguments against a number of positions. Smith introduced it as a premise in an argument against externalist theories of moral motivation, according to which

---

<sup>2</sup>For example, the actor might be especially superstitious, and believe that not buying groceries for people with exactly these properties will bring bad luck.

<sup>3</sup>I won't go into detail about what exactly these moralized emotions are like. See e.g. Gibbard (1990: Ch. 7) for discussion.

making moral judgments is to be in a belief-like state—and not thereby to be motivated to act accordingly.<sup>4</sup> And more recently the DE DICTO DEFECT thesis has appeared in a number of arguments against a rational requirement to maximize expected moral value when one is uncertain as to which ethical theory is true.<sup>5</sup>

This is just a rough sketch of the content of the DE DICTO DEFECT thesis and its applications. In the following sections I will argue that it is implausible, owing to a class of examples where it is *vague* what one morally ought to do. There are several disambiguations of the thesis. On some disambiguations, being de re motivated is itself defective, owing to ethical vagueness.<sup>6</sup> Other disambiguations make the negative component implausible. More importantly, the problems for each disambiguation make the thesis inapt for one role it has played in recent literature, which is as a premise in arguments against expected value maximization in cases of moral uncertainty.

## 1 Vagueness and de re concern

Sometimes it is *vague* what one morally ought to. In these cases, it is very natural to have a desire to do the right action, whatever it is. This is a type of de dicto concern.

First, some background on the structure of ethical vagueness. Consider a series of cases  $c_1 \dots c_n$  where in each one can press a button to kill a creature to save a fully developed, adult human. In the first case  $c_1$ , the creature one

---

<sup>4</sup>See Copp (108) and Brink (1997) for examples.

<sup>5</sup>See Ross (2006) and Sepielli (2009) for more on this idea, and Harman (2014), Weatherson (2014), and Hedden (2016) for criticisms based on the DE DICTO DEFECT thesis. I discuss these criticisms in §4.

<sup>6</sup>Thus in this respect the argument is similar to the claim in Lillehammer (1997: 192) that de re moral motivation is sometimes defective.

can kill by pressing the button is an amoeba. In the last case  $c_n$ , the creature one can kill by pressing the button is another adult human. And in each case  $c_i$  between  $c_1$  and  $c_n$ , the creature one can kill is slightly more complex and developed than the creature in the case  $c_{i-1}$  immediately before it.

It is wrong to press the button in  $c_n$ : it is morally impermissible to take the life of one person to save the life of another. It is not wrong (and is in fact morally required) to press the button in  $c_1$ : one should kill an amoeba to save a human if given the chance. But for some cases  $c_i$  in between  $c_1$  and  $c_n$  it will be vague whether it is wrong to press the button. In these cases the creature one can kill will be similar enough in relevant respects, but not identical, to the human one can save. We will not want to assert that it is not wrong to press the button in  $c_i$ , and simultaneously assert that it is wrong to press the button in  $c_{i+1}$ . This is a sorites series for 'wrong', and is strong evidence that 'wrong' is vague.<sup>7</sup>

Call the borderline cases where it is vague whether pressing the button is wrong  $b_1 \dots b_n$ .<sup>8</sup> Corresponding to each borderline case is a property which, for all we know, is the property of wrongness and draws the cut-off point for wrongness at the respective borderline case. There is the property that is not instantiated by  $b_1$  but is instantiated by  $b_2$  and all of the cases after. Call this the property *Wrong*<sub>1</sub>. There is the property that is not instantiated by  $b_1$  or  $b_2$  but is instantiated by  $b_3$  and all of the cases after; this is *Wrong*<sub>2</sub>. And similarly for all of the other borderline cases. For each borderline case  $b_i$ , there is a property *Wrong*<sub>*i*</sub> that is, in most respects, identical to wrongness, at most differing by

---

<sup>7</sup>See Shafer-Landau (1995), Dougherty (2013), and [author's article] for more discussion.

<sup>8</sup>Because of higher-order vagueness, it will be vague which cases are  $b_1 \dots b_n$ . I will return to this point below.

drawing a cut-off point at  $b_i$ .<sup>9</sup>

I will begin with the assumption that one of the properties  $Wrong_1$ ,  $Wrong_2$ , etc is wrongness—that is, that one of these precisifications is the property that *constitutes* wrongness. But in cases of vagueness one cannot know which of these properties is the wrongness-constituting property. This is why one wouldn't want to say, for instance, that the button may be pressed in  $b_2$  but it would be wrong to press in  $b_3$ . Even if it is true that one may press in  $b_2$  and may not press in  $b_3$ , the fact that these are borderline cases means that one cannot know this. In short, the central phenomenon I will use to investigate the DE DICTO DEFECT thesis are cases where there is a fact about whether an action is wrong, but in principle we cannot know what this fact is.

These assumptions are most naturally associated with the Epistemicist view of vagueness.<sup>10</sup> The only assumptions we need, however, are that (a) there is a fact of the matter about which property constitutes wrongness, and (b) we cannot know what this property is. These are assumptions that will be true of some cases on a range of other views as well. Many non-Epistemicist views about vagueness will approximate the Epistemicist's way of talking about borderline cases and cut-offs.<sup>11</sup>

Even if vagueness is not as the Epistemicist describes, similar assumptions will hold on alternative approaches. That is, there will be importantly unknowable facts surrounding ethical vagueness on alternative views, even if the unknowability does not constitute vagueness, on these views. On semantic

---

<sup>9</sup>Of course wrongness has an extension in cases in addition to  $c_1 \dots c_n$  in our sorites series from above. I assume that  $Wrong_1$ ,  $Wrong_2$ , etc. do not diverge from the extension of wrongness in these cases.

<sup>10</sup>Cf. Williamson (1994)

<sup>11</sup>See the discussion of Supervaluationism's commitment to precise cut-off points in Keefe (2000: 183 ff.).

theories of vagueness, there are a number of admissible precisifications of 'wrong' which draw cut-off points at different borderline cases, where the current meaning of 'wrong' does not rule out each of the admissible precisifications.<sup>12</sup> On ontic theories, there is a vague property that 'wrong' picks out, and it is vague whether the borderline cases instantiate this property.<sup>13</sup> It will, nonetheless, be hard to know which actions are included in every precisification of 'wrong', or are not vaguely instantiated by the property it picks out. There are pairs of cases where the actions in them are nearly indistinguishable, yet only one is included in the extension of all precisifications of 'wrong', or determinately instantiates the property 'wrong' picks out. Problems for de re moral motivation, analogous to those I raise below, will arise in these cases.

Suppose that wrongness is the property *Wrong*<sub>3</sub>. In the language of the opening gloss on de dicto and de re motivation, we can suppose that the conjunction of the features which constitute wrongness  $F_1, F_2 \dots F_n$  are had by all and only the actions that instantiate *Wrong*<sub>3</sub>. One might be motivated to do the morally correct thing in all cases where there is an actions that is determinately wrong. But this does not suffice for de re moral motivation, which would involve moral concern for the property *Wrong*<sub>3</sub>.

If one is in the borderline case  $b_2$ , one can recognize that there is some pressure to press the button—after all, pressing will save the life of a human. One can also recognize that there is some pressure to not press—after all, pressing results in the death of a creature that is similar to creatures that are definitely worth saving. These are good things to care about, and someone who has these concerns is not completely in the dark with respect to the

---

<sup>12</sup>Lewis (1982)

<sup>13</sup>Barnes (2010)

content of morality. But if one is to do the morally right thing with a de re moral motivation, one needs to do more than care about the features that would give one sufficient moral reason to act in the non-vague cases  $c_1$  and  $c_n$ . One needs also to care about the property that  $b_2$  instantiates and  $b_3$  lacks, which makes it the case that it is wrong to press the button in  $b_3$  and not wrong to press the button in  $b_2$ . One needs, on the present assumption, to press the button in  $b_2$  *because* wrongness is the property *Wrong*.<sup>14</sup>

At first glance things do not look promising. Since, by hypothesis, it is in principle not possible to distinguish morally between the borderline cases, one should not be morally concerned with this property. If these appearances hold up, both the positive and negative components of the DE DICTO DEFECT thesis fail.

In the next section I will distinguish between different senses in which one might be concerned with a property. The DE DICTO DEFECT thesis can be read in multiple ways. But on each reading, a version of the foregoing is pressing.

## 2 Vagueness and the De Dicto Defect thesis

Exactly why de re moral motivation in vague cases is a bad way to be motivated depends on how the DE DICTO DEFECT thesis is spelled out. I will outline a number of ways to understand motivation in terms of different kinds of reasons that can be reasons for acting. The DE DICTO DEFECT thesis can be read in distinct ways, depending on the kind of reason involved in the thesis. One option is to read it as a thesis about what I will call the *personal*

---

<sup>14</sup>? considers a distinct worry that arises from ethical vagueness: that sorites series like the one sketched above show that there is no right thing to do. The present issue is different, since it is consistent with there being a fact about what the right thing to do is, but it is not possible to be motivated by de re moral reasons to do it.

reasons of an agent—the reasons an agent has, in some sense. In this case the thesis claims that agents whose only personal moral reasons are not *de re* are defective. Or it could be read as a thesis about what *motivates* a good moral agent, claiming that an agent in fact should be moved to decide in certain cases based on whether an available action has a particular *de re* property or not. In either case, I will argue, difficulties lurk for the thesis in cases of vagueness, and these difficulties are not properly understood when we fail to disambiguate different versions of the DE DICTO DEFECT thesis.

Each kind of reason is distinct from a *normative* reason. Normative reasons are facts that support, or count in favor of, acting in a certain way.<sup>15</sup> Normative reasons are not relevant to the DE DICTO DEFECT thesis. One might be aware of the the facts that count in favor of acting in a certain way, or one might not be aware of them. But an agent can't be defective in virtue of what her normative reasons are, in the way the DE DICTO DEFECT thesis claims. All agents have *de re* moral facts as normative reasons; some of them are just unaware of it.<sup>16</sup>

## 2.1 *Possessed reasons*

Begin with an intuitive gloss on why the DE DICTO DEFECT thesis is not a thesis about normative reasons: at a first pass, it is about the reasons non-defective agents *have* for acting. Some agents have such reasons, while others lack them, in virtue of differences in their motivational profiles. One type of reason which fits this characterization is what we can call a *personal* or a *possessed* reason. These are reasons which count in favor of acting in some

---

<sup>15</sup>See Scanlon (1998).

<sup>16</sup>Take a simple case where one can take a bet where one risks one's life savings that a fair coin-flip has come up 'heads'. Supposing the coin flip has already happened and the coin landed 'heads', one has a normative reason to take the bet: one will win the bet and keep one's life savings. But one needn't know of the result of the coin flip in order to have this normative reason.

way, but are in addition reasons that agents bear some cognitive relationship to. Suppose it is true that the fact that pressing the button in a particular borderline case instantiates *Wrong*<sub>3</sub> is a reason not to press the button in that case. It doesn't follow that agents act in borderline cases with this fact as their reason for acting.

Here are some characteristics of possessed reasons. First, possessed reasons are facts—they need to be true. If John drinks from a glass that contains liquid which is in fact gasoline, it can't be that his reason for drinking was that the glass contained gin. This is so even if he thought that there was gin in the glass before he picked it up. Since there is no gin in the glass, that there was gin in the glass can't be his (possessed) reason.<sup>17</sup>

Second, one way of attributing possession of a reason to an agent is with what I will call a *personal reason ascription*. These ascriptions take the following form:

*a*'s reason for *φ*-ing is (was) that *p*.

Here *a* names an agent, *φ*-ing an action, and *p* can be replaced with a declarative sentence. We have already noted that the sentence needs to pick out a fact.

Finally, there are further constraints on the kind of fact that can be a possessed reason and the reason ascribed in a true personal reason ascription. These need to be facts that the agent in the ascription has some epistemic access to: if John refuses the glass by claiming that it contains gasoline (which it in fact does), but has no reason to believe that it does, the corresponding personal reason ascription—that his reason for not drinking was that the glass

---

<sup>17</sup>Unger (1975). The example is from Williams (1981: 102).

contained gasoline—can't be true.<sup>18</sup> Most plausibly he also needs to *know* that the glass contains gasoline before this can be his reason.<sup>19</sup>

Return to a case of ethical vagueness, where one has a choice to press the button in one of the borderline cases  $b_1 \dots b_n$ . Wrongness is, by hypothesis, one of the properties  $Wrong_1, Wrong_2 \dots Wrong_n$ . Each of these properties draws the cut-off point for a wrong button-pressing at a different borderline case. But only one of these properties draws the cut-off point where wrongness does. If one is motivated by a *de re* concern for morality, this would mean that one is motivated by the fact that one's action has the property  $Wrong_i$ , the property that actually constitutes wrongness. Owing to vagueness, one cannot know that wrongness is  $Wrong_i$ . This complicates the relationship between facts about  $Wrong_i$  and one's personal reasons for pressing (or not pressing) the button in borderline cases.

While the unknowability of which property out of  $Wrong_1, Wrong_2, \dots$  constitutes wrongness complicates the relationship between  $Wrong_i$  (whichever property in fact constitutes wrongness) and personal reasons, it does not make it *impossible* for facts about  $Wrong_i$  to serve as personal reasons. Start with a disanalogy between the unknowability of the fact that  $Wrong_i$  constitutes wrongness, and other cases where there are facts one cannot know. Take, for example, a case where one knows that there is a coin flip which has already happened, but doesn't know whether the coin landed on the 'heads' side. Even if it is true that the coin landed on the 'heads' side, one can't have this fact as a personal reason for acting. For instance, if one is offered a bet on whether

---

<sup>18</sup>Unger (1975: 208)

<sup>19</sup>Williamson (2000), Hyman (2015), Hawthorne and Magidor (forthcoming). I will use this strong epistemic constraint on personal reasons in what follows. But many of the points I make below will apply if we use a weaker requirement.

the coin landed 'heads', one can't have as one's reason for taking the bet that the coin landed 'heads'. It is not possible to have the fact that the coin came up heads as a personal reason to bet.

But one can know something about *Wrong<sub>i</sub>*. For instance, one can know that pressing the button in *b<sub>j</sub>* instantiates *Wrong<sub>i</sub>*, when  $i < j$ . This is true by definition—when introducing *Wrong<sub>i</sub>* and its cousins, we defined each property in terms of which actions instantiate it. One can know some things about this property, including facts about what instantiates it. What one cannot know is that this property, *Wrong<sub>i</sub>*, is wrongness.

Thus unlike the result of the coin flip, the fact that pressing in *b<sub>j</sub>* instantiates *Wrong<sub>i</sub>* can in principle be one's personal reason for acting. If, for instance, one is asked to make a list about the actions that are *Wrong<sub>i</sub>*, one could have de re facts about the property among one's personal reasons for putting the actions in certain borderline cases on the list.

This is not enough to support the positive component of the DE DICTO DEFECT thesis, understood as a thesis about what one's personal reasons for acting should be. The thesis covers what we should understand to be defective (or non-defective) *moral* motivations, and so in these cases we are interested in whether the fact that pressing the button instantiates *Wrong<sub>i</sub>* can be the personal reason of someone who presses the button for moral reasons. There are two options for identifying what one's moral reason for acting is in these borderline cases. It could be a purely de re fact: that pressing the button instantiates *Wrong<sub>i</sub>* is one's reason for not pressing. Or one's reason for so acting could be a de re fact that in addition specifies the relation between *Wrong<sub>i</sub>* and wrongness: that pressing the button instantiates *Wrong<sub>i</sub>* and

*Wrong<sub>i</sub>* constitutes wrongness. But in neither case can we say that one ought to have the purported personal reason as one's reason for pressing the button.

Take the first option: that one's reason for not pressing the button in case *b<sub>j</sub>* is that pressing in *b<sub>j</sub>* instantiates *Wrong<sub>i</sub>*, full stop. One can know this claim, so in principle there is no barrier to its being one's personal reason. But it is defective to have this fact as a personal reason for acting morally. If one declines to press the button in a borderline case, and in not pressing the is disposition to feel the relevant reactive attitudes in response to agents who press the button in identical circumstances.

One way to motivate the defectiveness of one's action in this case is as follows. Ex hypothesi because of ethical vagueness, one has no idea whether pressing the button in this case falls within the cut-off point for wrongness. So, if one avoids pressing the button in a distinctively moral way purely because one believes it instantiates *Wrong<sub>i</sub>*, one is acting as if the fact that an act instantiates *Wrong<sub>i</sub>* is a moral reason when, for all one knows, this fact has no bearing on whether one morally should refrain from pressing.<sup>20</sup> An agent is making a mistake if she feels guilt at the thought of doing *Wrong<sub>i</sub>* actions, when she has no idea how *Wrong<sub>i</sub>* relates to the property of moral wrongness.

The second option is to hold that, when non-defectively motivated in a borderline case, one's personal reason for not pressing the button is the conjunctive fact: that pressing the button instantiates *Wrong<sub>i</sub>* and *Wrong<sub>i</sub>* constitutes wrongness. This fact clearly bears on how one should act morally

---

<sup>20</sup>Analogously, one would display a similar defect if one knew the outcome of the coin flip but were unaware of its relevance to the outcomes of a potential bet of one's life savings. Even if one knows that the coin landed 'heads', it is a defect to make this one's reason for betting one's life savings if one doesn't know that picking the outcome of the flip will allow one to keep one's savings and not lose them. These points are somewhat related to discussion of the 'tracking condition' in Smith (1994) and Dreier (2000).

in a borderline case. If it were one's personal reason for acting, it would be a good reason not only for refraining from pressing, but also for doing so in a way that is connected to the disposition to feel reactive attitudes of guilt and blame in response to pressing the button in identical situations. The relevance of an act's being *Wrong<sub>i</sub>* to whether one morally ought to do it would be clear, because the fact that constitutes one's reason includes the information that *Wrong<sub>i</sub>* acts are morally wrong. This solves the defectiveness of moral action with the purely de re fact about *Wrong<sub>i</sub>* as one's reason.

But one can't have the fact that pressing the button instantiates *Wrong<sub>i</sub>* and *Wrong<sub>i</sub>* constitutes wrongness as one's personal reason for acting. One can't know which property out of *Wrong<sub>1</sub>*, *Wrong<sub>2</sub>*, ... *Wrong<sub>n</sub>* constitutes wrongness. And so one can't have as a personal reason a claim that includes the fact that *Wrong<sub>i</sub>* is wrongness. It is false that one ought to have this fact as one's personal reason. It isn't false because being de re motivated in this way is a bad state to be in. It is false because this kind of de re fact isn't a personal reason one can possibly have.

Thus the positive component of the DE DICTO DEFECT thesis is false, when read as a claim about de re facts about wrongness serving as one's personal reasons.

## 2.2 *Explanatory reasons*

Not all reasons for action are personal reasons. When citing the reasons why someone acted, we can also cite the causes of their action, which needn't be personal reasons. For example the fact that my brain is releasing serotonin while I walk outside on a sunny day might explain why I smile at a passing stranger. But the release of serotonin is not my (possessed) reason for

smiling—I need not even have the belief that my brain is releasing serotonin, and even if I did I would not consider it to be a fact that counts in favor of smiling. The serotonin is part of the causal explanation for my smiling, but is not among my reasons for doing it. In this case we can say it is an *explanatory* reason for my action.

Reasons for acting which include *Wrong<sub>i</sub>* can be explanatory reasons in this sense. In particular, the de re fact about which actions instantiate *Wrong<sub>i</sub>* could be what explains our acting by being a part of the causal explanation of why the button was or was not pressed in a particular case. Even if one does not know that *Wrong<sub>i</sub>* is wrongness, it could still be that one's not pressing the button is causally explained by the fact that pressing the button instantiates *Wrong<sub>i</sub>*. These are cases where the fact that the action instantiates *Wrong<sub>i</sub>* is an explanatory reason.

Typically, the fact that an act instantiates *Wrong<sub>i</sub>* serves as an explanatory reason for an agent's action when she is disposed in some way to have her actions affected by the presence (or absence) of the property *Wrong<sub>i</sub>*. This doesn't mean that she is aware of her disposition. And it doesn't mean, even if she is aware of her disposition, that she treats her disposition as her reason for acting. The fact that one is disposed to act as if *Wrong<sub>i</sub>* is wrongness is not a normative reason for not pressing the button. Even if one knows that this is one's disposition, it is not a fact that one counts in favor a particular course of action. A non-defective agent would not take this fact to be a reason.

The DE DICTO DEFECT thesis might then be read as requiring that an agent's explanatory reasons for acting be de re facts about the property that constitutes wrongness. Since an explanatory reason need not be something one believes

at all, the fact that one can't know that *Wrong<sub>i</sub>* is wrongness does not prevent one from having this fact as an explanatory reason. This is perhaps a more promising route for the proponent of the DE DICTO DEFECT thesis to take.

With explanatory reasons of this kind, we can distinguish between two cases: those in which one *knows* that one is caused to act, in part, by the fact that an action instantiates *Wrong<sub>i</sub>*, and cases in which this fact is part of the causal explanation of one's action but one does not know that one's action is caused in this way. In short, one might know the explanatory reasons for one's action, or one might not know them.

First take someone who does have the relevant knowledge. That is, they are motivated to avoid *Wrong<sub>i</sub>* actions, and to engage in other distinctively moral behavior with respect to the property *Wrong<sub>i</sub>*. They are in a position to know that they are treating *Wrong<sub>i</sub>* as if it were wrongness. The agent who is aware that *Wrong<sub>i</sub>* plays this role in regulating their action and reactive attitudes has a *luminous* explanatory reason—it is an explanatory reason for her action that she is aware of.

Owing to ethical vagueness, one can't know that *Wrong<sub>i</sub>* is wrongness. Even if there is a luminous explanatory reason for one's action in a borderline case, one cannot know how the explanatory reason relates to what one morally ought to do. One isn't forced to treat *Wrong<sub>i</sub>* as the cut-off point for moral wrongness; one always has the option of withholding reactions of blame and praise to actions in borderline cases. Having a luminous disposition to treat *Wrong<sub>i</sub>* as wrongness leads one to take actions one knows one has no reason

to take in borderline cases.<sup>21</sup> This conflicts with the positive component of the DE DICTO DEFECT thesis.

If one's explanatory reasons are luminous, one is in a position to change one's motivations to be better. One can change so that moralized reactive attitudes are absent from each borderline case. That is, while it is true that one needs to either press or not press the button in each case, one shouldn't do it with the accompanying dispositions to have attitudes that treat the action one takes as required by morality. A defective agent retains these dispositions.

The alternative is to read the DE DICTO DEFECT thesis in terms of non-luminous explanatory reasons: the thesis might be read as requiring in part that agents have explanatory reasons for acting which involve de re facts about the property *Wrong<sub>i</sub>*, but not to know that that they are motivated to act in the relevant ways. On this reading of the thesis, agents who have facts about properties other than *Wrong<sub>i</sub>* as their only non-luminous explanatory reasons are defective. There are broadly two ways in which agents can fail to have this knowledge: such agents can have a false belief about the causes of their actions, or they can have no beliefs at all about their moral motivations.

An agent who does not know what her motivations are because she has false beliefs about how she is motivated is defective in some way—she has a false belief. So a state where one has de re explanatory reasons for acting, and moreover has false beliefs about what these reasons are, cannot be a state one is required to be in.

---

<sup>21</sup>Of course, one is forced to either press the button or not press the button in every borderline case—there is no third option. So it is not defective simply to act in a way that leads one to refrain from pressing the button only in cases that instantiate *Wrong<sub>i</sub>*, since one has to be disposed to exhibit some pattern of pressing and not pressing. But there is no reason to hold that this pattern, even if it coincides with wrongness, is required to be accompanied by the relevant reactive attitudes.

But another way for an agent to have non-luminous explanatory reasons is to have no beliefs at all about their motivations. It is not obvious that these agents are defective. They are disposed to avoid doing actions which instantiate *Wrong<sub>i</sub>*, and also to feel the appropriate reactive attitudes. They do not, however, know that they have these motivations, or these reactions. In short they are motivated in a way that treats *Wrong<sub>i</sub>* as a moral property, but do not know they are doing this. In this case it is not obvious that de re motivated agents are defective. Whether having the non-luminous de re explanatory reason is *required*, in order to avoid defect, is another question.<sup>22</sup>

To start, agents who have non-luminous explanatory reasons for acting do not necessarily have personal reasons for treating *Wrong<sub>i</sub>* as a moral property. That is, there is nothing which is their reason for avoiding actions that instantiate *Wrong<sub>i</sub>*, for blaming people who do actions that instantiate *Wrong<sub>i</sub>*, etc. (These agents act as if facts about *Wrong<sub>i</sub>* are their reasons for acting, but since they do not know that *Wrong<sub>i</sub>* is wrongness, do not have de re facts about *Wrong<sub>i</sub>* as their personal reasons.) There may be something unfortunate in acting in this way. Agents like this act in ways that have a moral character which the agent is unaware of. But it isn't a straightforward kind of irrationality or other kind of defect when one is moved to act under such ignorance: after all, not all of one's actions can be actions one has personal reasons for. Non-luminous moral action in borderline cases might just be one instance of this relatively benign phenomenon.

Here is one example. Take someone who encounters actions that are *Wrong<sub>i</sub>* early in their life and, during these encounters her psychology contingently and purely coincidentally produces a mildly unpleasant sensation at the

---

<sup>22</sup>[acknowledgment blinded for review]

prospect of performing *Wrong<sub>i</sub>* actions. This causes her to avoid performing *Wrong<sub>i</sub>* actions and, over time, she avoids them by associating the morally relevant reactive attitudes with performances of *Wrong<sub>i</sub>* actions. Even as the mildly unpleasant sensations go away, as she is accustomed to avoiding actions that are *Wrong<sub>i</sub>* in this way. A similar causal story could easily have produced a disposition to avoid actions with a distinct property that has a different cut-off point along the borderline cases. But it in this case, as it happens, the process singles out *Wrong<sub>i</sub>*, the property that in fact constitutes wrongness, even though the agent in question needn't have knowledge of it.

In this case it is not obvious that acting as if *Wrong<sub>i</sub>* is wrongness, with these explanatory reasons, is defective. The agent lacks some self-knowledge, but in general this kind of ignorance will not be avoidable in all cases). So acting with non-luminous explanatory de re reasons is consistent with the positive component of the DE DICTO DEFECT thesis.

Is one defective if one has other explanatory reasons instead? One could be motivated by de re facts about other properties—since the reasons are non-luminous, a similar causal process could make any of *Wrong<sub>1</sub>*, *Wrong<sub>2</sub>*, etc. one's explanatory reason for acting. Or, one could have non-de re explanatory reasons of this kind. But plausibly having other de re, de dicto facts about wrongness, serve as non-luminous explanatory reasons will not be defective for the same reasons as having a fact about *Wrong<sub>i</sub>* as one's non-luminous explanatory reason.<sup>23</sup> So non-luminous explanatory reasons are not good news without qualification for the DE DICTO DEFECT thesis. I will explore this idea

---

<sup>23</sup>The de re reasons, since they involve properties that have cut-off points that do not coincide with the cut-off point for wrongness, will lead one to act wrongly in some cases. But they will coincide with wrongness in other cases, and we can ask whether there is anything defective with doing the right thing with the de re explanatory reason in these cases.

in greater detail in the next section.

To sum up: the DE DICTO DEFECT thesis can be precisified in multiple ways, corresponding to different interpretations of ‘concern’ in the original formulation of the thesis. On many precisifications, the positive component looks unappealing. This is not, overall, a favorable result for proponents of the thesis. However we might expect the DE DICTO DEFECT thesis to contain an insight that makes it a legitimate tool to settle specific questions. In the next section I will look at one specific debate to which the DE DICTO DEFECT thesis has been applied, namely debates over what to do when one is morally uncertain.

### 3 The De Dicto Defect Thesis: Acting under Moral Uncertainty

It has been argued that, when one is uncertain which moral theory is true, what one rationally ought to do is maximize *expected moral value* (EMV for short). An action’s EMV is the sum of the moral goodness (or badness) of the action according to every possible moral theory, weighted by one’s credence that the theory is true. Formally, this amounts to the following:

$$\text{EMV}(\phi) = \sum_T MV_T(\phi) \times Cr(T),$$

where  $T$  is a possible moral theory,  $MV_T$  is a function that delivers the moral value of an action according to  $T$ , and  $Cr(T)$  is the likelihood that  $T$  is true.<sup>24</sup>

Call this the EMV-maximization approach. Recent literature has argued that the EMV-maximization approach on the grounds that it runs afoul of the DE DICTO DEFECT thesis. Here are some examples.

---

<sup>24</sup>See Jacob Ross (2006) and Andrew Sepielli (2009) for developments and applications of this idea. Additional examples of dilemmas involving moral uncertainty can be found in Lockhart (2000), Weatherson (2014), and Harman (2014).

Brian Weatherson (2014) suggests that EMV-maximizers have motivations that are deficient in a way that is analogous to the deficiency in someone who has *de dicto* concern for their own welfare. Weatherson's provides an analogy involving Bruce, who cares about his own welfare as such, and deliberates whether to go to a gallery on the basis of whether it would improve his welfare, whatever that turns out to be. This is, Weatherson says, a kind of *de dicto* concern for one's welfare that is deficient:

The rational person values their health, happiness and friendships (and whatever goes into the actual list of things that constitute welfare). They don't simply value their welfare, and desire to increase it. That's why it would be perverse for Bruce to go to the gallery. He would only go if he had a strange motivation. And it is why it would be perverse for Martha to turn down the steak. To do so she would have to care about morality, whatever it is, not about the list of things that Smith rightly says a good person will care about. (Weatherson 2014: 152)

And Elizabeth Harman (2014) rejects EMV-maximization for its requirement not to care about morality *de re* in the following passage:

In my view, people who act wrongfully are blameworthy not in virtue of what their moral beliefs and credences were, but in virtue of what their non-moral beliefs and credences were, and how these influenced their choices. Someone who knows she is killing an innocent person, and does so anyway, does not care adequately to avoid killing the innocent. A view of blameworthiness that can undergird the claim that false moral views do not exculpate is this:

A person is blameworthy for her wrongful behavior just in case it resulted from her failure to care de re about what is morally important: that is, from her failure to care adequately about the non-moral features of the world that in fact matter morally.

A person cares de dicto about morality if she wants to be moral. A person cares de re about morality if she wants to keep her promises, to help the needy, etc., and if keeping one's promises, helping the needy, etc. are in fact morally important. (Harman 2014: 67)

Since the EMV-maximizer is concerned with the weighted moral value of an action according to competing moral theories, she is not de re concerned with morality. She will, according to Harman, be blameworthy when she fails to do the right thing.

Both Harman and Weatherson connect the de dicto moral motivation that EMV-maximizers display with some kind of susceptibility to criticism on the agent.<sup>25</sup> There are some important clarifications to make about these claims at the outset.

First, both give a more precise reading of the kind of 'defect' alleged by the which is alleged by the DE DICTO DEFECT thesis to accompany de dicto moral motivation. For Harman the defectiveness takes the form of blameworthiness; Weatherson claims it is irrationality. These are different types of defect, and in principle an agent could exhibit one without the other.

Second, the assessments cover different cases. Weatherson's suggestion of irrationality would in principle apply even if the agent does the action which

---

<sup>25</sup>See also Hedden (2016) for a similar argument in this context.

is in fact morally required. Going to the gallery might be good for Bruce, and he might in fact do this on the grounds that it will increase his welfare. But, Weatherson says, acting in this way is irrational. Harman, on the other hand, holds that an action is blameworthy if it is in fact the wrong thing to do *and* it is motivated by mere de dicto concern for morality.

Finally, there is a difference over the positive component as well. What is required for rational motivation, according to Weatherson, is to care about “whatever goes into the actual list of things that constitute welfare”. This suggests that concern for the rightness-constituting properties is required for non-defective right action. But Harman says something weaker: “adequate” care for the morally relevant features is required to avoid blameworthiness. Adequacy comes in degrees, and can vary along multiple dimensions.

### 3.1 *Companions in guilt*

In cases of ethical vagueness, one has to press or not press the button in each case. One can't know which of these actions instantiates the property that constitutes wrongness, so one shouldn't be motivated by de re facts about this property. Cases of moral uncertainty are similar in one respect: one does not know which of one's available options instantiates the wrongness-constituting properties. (The source of this lack of knowledge could be different in the two cases.) Since there is some similarity between the two cases, proponents of the EMV-maximization view might defend themselves from charges of problematic de dicto motivation by pointing to this analogy.

We might make the analogy even stronger. It is arguable that in cases of ethical vagueness one is *rationally required* to do something analogous to EMV-maximization. When one is in a case of ethical vagueness, one is not in a

position to know whether pressing the button in the case is wrong, since one cannot know which property constitutes wrongness. But one can know in a borderline case how *likely* it is that one's action has the wrongness-constituting property, whatever it is. (For instance, it is more likely that pressing the button in  $b_{100}$  is wrong than it is that pressing the button in  $b_1$  is wrong.) Then whether one should press in a borderline case  $b_i$  is determined by whether  $EMV(\text{press in } b_i) > EMV(\text{refrain in } b_i)$ . In other words one should press the button just in case the following holds:

$$\sum_n MV_{WRONG_n}(\text{pressing in } b_i) \times Cr(WRONG_n) > \sum_n MV_{WRONG_n}(\text{not pressing in } b_i) \times Cr(WRONG_n)$$

where  $WRONG_n$  is the proposition that wrongness is constituted by the property  $Wrong_n$ .<sup>26</sup>

This appears to be a rationally permissible, and perhaps a rationally required, approach to decision-making when one finds oneself in a case of ethical vagueness. But it is an instance of de dicto concern for wrongness, since one's reasons for acting in part involve the likelihood that a some properties, which do not in fact constitute wrongness, are identical to it. So in these cases one is acting with some measure of de dicto concern for moral wrongness. If this is not a defective way to act in cases of ethical vagueness, the EMV-maximizer can argue, by analogy, that similar ways of being concerned with morality are not defective or inadequate in other cases of more prosaic moral uncertainty.

---

<sup>26</sup>This approach is discussed in greater detail in [author's article].

### 3.2 *Deeper worries*

Harman and Weatherson both claim that there is a specific defect—blameworthiness and irrationality, respectively—involved with the kind absence of de re motivation recommended by EMV-maximizers. Vagueness shows that the connection between lacking de re motivation and each kind of defectiveness is tenuous.

Take Harman's view that lack of de re concern implies blameworthiness. Having de re facts about wrongness as one's possessed reasons requires knowledge of the de re fact which is a part of one's motivation. In principle there are two different kinds of possessed reason one might have in these cases. The first is the case where one has possessed reasons involving the fact that one's act instantiates the wrongness-constituting property *Wrong<sub>i</sub>*, and is thereby wrong. But it is impossible to have personal reasons of this kind, because one cannot know that *Wrong<sub>i</sub>* constitutes wrongness. So one is not blameworthy for lacking this kind of motivation.

The more important case is where one has as one's possessed reason the de re fact that an act instantiates the precise property *Wrong<sub>i</sub>*, simpliciter. This fact makes no claim about the relationship between *Wrong<sub>i</sub>* to wrongness. It is knowable. So it can serve as one's possessed reason for acting. It is doubtful, however, that agents who fail to have this de re reason, and act wrongly in a borderline case, are deserving of blame. One should recognize that it would be wholly arbitrary to take a de re fact about *Wrong<sub>i</sub>* as one's possessed reason for acting, rather than facts about some other property that draws a different cut-off point along the borderline cases. One might then refrain from having any of these de re facts as one's reason for acting in borderline cases.

Here is an example of an alternative. One could have as one's reason for acting a fact about the conjunctive property **W**:

**W**  $Wrong_1 \wedge Wrong_2 \wedge \dots \wedge Wrong_n$ ,

where each disjunct draws a cut-off point at a different borderline case. It seems reasonable to take facts about **W** as one's reasons for acting, in view of the facts about what one can in principle know about borderline cases. Acting with the fact that an act instantiates **W** as one's possessed reason will not lead one to act wrongly in a non-borderline case. Moreover it would seem that having a more specific motivation which discriminates between borderline cases is not rational.

But someone who acts with **W** as part of their reason might act wrongly in a borderline case: by definition the borderline case will instantiate some, but not all of the conjuncts of **W**. Since **W** does not constitute wrongness, this wrong action will not be done for a de re reason concerning the property that constitutes wrongness. It does not, however, appear to be *blameworthy* to do the wrong thing with a fact involving **W** as one's reason. Additional sensitivity to the wrongness-constituting facts is not possible in a case of vagueness, and so one can hardly be blamed for lacking it.

Of course the relevant notion of concern need not be interpreted in terms of possessed reasons at all; de re facts about wrongness to serve as explanatory reasons for one's actions. These reasons can be luminous or not. Harman, then, could claim that someone who acts wrongly is blameworthy when either their luminous or their non-luminous explanatory reason for their wrongful action does not involve de re facts about  $Wrong_i$ .

If it is luminous explanatory reasons that feature in this version of the DE

DICTO DEFECT thesis, then the negative component is false. Start with Alice, who finds herself with a de re luminous explanatory reason for moral action which involves the property *Wrong<sub>i</sub>*—exactly the property that constitutes wrongness. Since Alice’s reason is luminous, she will realize that she is treating *Wrong<sub>i</sub>* as wrongness, since facts about *Wrong<sub>i</sub>* will cause her to avoid actions that instantiate it and to feel the relevant pattern or reactive attitudes. It would be understandable—and perhaps rationally required—for Alice to revise her motivations, if feasible. One way to do this would be to change her motivations to involve the property **W** instead. This would ensure morally correct actions in borderline cases, but leave open how she acts in borderline cases, where she cannot know what the morally required action is. She might then act wrongly in some borderline cases, and do so without a de re motivation.<sup>27</sup>

But, contra the negative component of the DE DICTO DEFECT thesis, Alice’s act is not blameworthy wrong action. The reasons are the same as the reasons why a wrong action in a borderline case with a personal reason involving **W** are not blameworthy.

Perhaps we should instead claim that agents whose non-luminous explanatory reasons are not de re facts about *Wrong<sub>i</sub>* are blameworthy when they act wrongly. If the relevant explanatory reasons are not luminous, then even agents who can in principle change their motivations will not be aware of them. These agents have no reasons to try to avoid de re concern for properties that draw precise cut-off points among borderline cases.

---

<sup>27</sup>There is also the option that Alice retains the disposition to not press the button in exactly all of the borderline cases where pressing the button is wrong, but does not have the moralized reactive attitudes toward these acts in the borderline cases. In such a case Alice does not act wrongly, even in borderline cases. But she doesn’t constitute an example of someone who cares adequately for the wrong-making features of actions, since she attaches no moral significance to the wrongness-constituting property. Agents who are motivated by facts about **W** instead are not blameworthy not having this motivation.

While agents won't have reason to retreat to de dicto motivations in this case, there are others who find themselves, via a similar causal-historical process, with de dicto non-luminous explanatory reasons. The negative component of the DE DICTO DEFECT thesis says that these agents will be blameworthy for their wrong actions, but this is implausible. Take a modified version of an earlier example: imagine an agent who has, instead of contingent psychological associations with the property *Wrong<sub>i</sub>*, associations that occur only for actions that instantiate **W**. As a result the explanatory reasons for her action involve the property **W**. She does not have a de re motivation. And sometimes, when it is vague what she morally ought to do, she acts wrongly, because being motivated by **W** only guarantees that she will do the right thing in cases where the ethical facts are determinate. But the only way she would end up with the a de re reason that also produces the right action in borderline cases would be to have a history of accidental unpleasantness that patterned with *Wrong<sub>i</sub>* instead. Failing to have this history is not blameworthy.

This does not, of course, show that a lack of awareness of the moral facts is *always* morally exculpating. There is room for critics of EMV-maximization to maneuver, in order to make the charge that it recommends objectionable de dicto concern for morality stick. But such a criticism has obstacles to overcome, since any general principles connecting de dicto motivation and blameworthiness will, when applied to the EMV-maximizer, need to take into account cases of ethical vagueness.

Weatherson does not claim that de re moral concern implies the absence of blameworthiness; instead he suggests that it is required by rationality.

It would be irrational is to show concern for *Wrong<sub>i</sub>* by having facts about

*Wrong<sub>i</sub>* as among one's possessed reasons for one's moral behavior; one cannot know that *Wrong<sub>i</sub>* has anything to do with morality. So having the de re possessed reason is irrational, contra the positive component of the DE DICTO DEFECT thesis. It is not, however, irrational to have as one's possessed reason the conjunctive fact that one's action instantiates *Wrong<sub>i</sub>* and that *Wrong<sub>i</sub>* is wrongness. This is not possible when we are dealing with possessed reasons, since the conjunctive fact is unknowable. The impossibility also implies that it is not rationally required to have the conjunctive de re fact as one's possessed reason; this is also inconsistent with the positive component of the DE DICTO DEFECT thesis, although in a different way.

If *Wrong<sub>i</sub>* features only in the explanatory reasons for one's action, one is still irrational if the explanatory reason is luminous. One should reflect and acknowledge that one is acting as if one knows a claim—that wrongness is *Wrong<sub>i</sub>*—which in fact one cannot know. One can retreat to a motivational profile that does not pattern moralized reactive attitudes along alongside the instantiation of *Wrong<sub>i</sub>*. This is incompatible with the positive component of the DE DICTO DEFECT thesis.

So the better option is to be non-luminous with respect to these explanatory reasons. But it can't be a requirement of rationality that one simply have some unknown cause for one's moral behavior. Sometimes—for instance when the causes of one's behavior are coincidental psychological quirks such as a mild aversion to actions that instantiate the disjunctive property **W**—one is not irrational for having one's behavior caused in this way. This is a de dicto motivation; so, having a non-luminous de re explanatory reason involving *Wrong<sub>i</sub>* is not the only way to avoid irrationality. Even if rationality

doesn't require having a de re motivation concerning wrongness-constituting properties in every circumstance, negative component of the DE DICTO DEFECT thesis should place some rational constraints on moral motivations. But when the motivations are non-luminous explanatory reasons, there are no constraints of the kind for properties that differ from wrongness only in their extension in borderline cases.

#### **4 Conclusion**

The DE DICTO DEFECT thesis can be interpreted in several ways. On some of these interpretations, it is strictly speaking false—there are cases where one should *not* have the de re motivations that positive component of the DE DICTO DEFECT thesis says one should have. Other interpretations make de re motivation permissible, but at the expense of vindicating de dicto motivation as well. The natural conclusion to draw from this is that the thesis so interpreted cannot be used, for instance, as a premise in arguments against the EMV-maximization approach to action in cases of normative uncertainty.

In reality all this shows is that the DE DICTO DEFECT thesis is not true with full generality. The exception cases could well be few and far between. Perhaps ethical vagueness is a genuinely *sui generis* phenomenon that is distinct from other types of moral uncertainty and ignorance. The original defenders of the DE DICTO DEFECT thesis may not intend to make any claims about the cases of vagueness. Thus it might seem uncharitable to object to the thesis on the grounds that it does not apply to cases of ethical vagueness—and so the unsuitability of certain kinds of de re motivation in cases of vagueness should not trouble proponents of the DE DICTO DEFECT thesis.

Unfortunately defenders of the DE DICTO DEFECT thesis cannot handle these issues simply by insisting that their thesis does not apply to cases of ethical vagueness. While cases of ethical vagueness may be apparently sui generis, the lesson we should take from them is that the DE DICTO DEFECT thesis needs to be revised, and the revisions must, without resorting to ad hoc restrictions, avoid the problems for the fully general version. It is not obvious in advance that best way to restrict the thesis will exempt requirements on de re moral motivation only to non-vague cases: as I have characterized it, ethical vagueness is similar in a number of respects to other cases of uncertainty about ethics. So we need to ask how far the requirements imposed by proponents of de re moral motivation fail to extend; the premise cannot feature uncritically in an argument against EMV-maximization.

## References

- Elizabeth Barnes. Ontic Vagueness: A Guide for the Perplexed. *Noûs*, 44(4): 601–627, 2010.
- David Brink. Moral Motivation. *Ethics*, 108(1):4–32, 1997.
- David Copp. Belief, Reason, and Motivation: Michael Smith’s *The Moral Problem*. *Ethics*, 1(33-54), 108.
- Tom Dougherty. Vague Value. *Philosophy and Phenomenological Research*, pages 1–21, 2013.
- James Dreier. Dispositions and Fetishes: Externalist Models of Moral Motivation. *Philosophy and Phenomenological Research*, 61(3):619–638, 2000.
- Allan Gibbard. *Wise Choices, Apt Feelings*. Harvard University Press, 1990.
- Elizabeth Harman. The Irrelevance of Moral Uncertainty. In Russ Shafer-Landau, editor, *Oxford Studies in Metaethics*, vol. 10. Oxford University Press, 2014.
- John Hawthorne and Ofra Magidor. Reflections on Reasons. In Daniel Star, editor, *The Oxford Handbook on Reasons and The Oxford Handbook on Reasons and Normativity*. Oxford University Press, forthcoming.

- Brian Hedden. Does MITE Make Right?: On Decision-Making Under Normative Uncertainty. In Russ Shafer-Landau, editor, *Oxford Studies in Metaethics vol. 11*, pages 102–128. 2016.
- John Hyman. *Action, Knowledge, and Will*. Oxford University Press, 2015.
- Rosanna Keefe. *Theories of Vagueness*. Cambridge University Press, 2000.
- David Lewis. Logic for Equivocators. *Noûs*, 16:431–441, 1982.
- Hallvard Lillehammer. Smith on Moral Fetishism. *Analysis*, 57(3):187–195, 1997.
- Ted Lockhart. *Moral Uncertainty and Its Consequences*. Oxford University Press, 2000.
- Jacob Ross. Rejecting Ethical Deflationism. *Ethics*, 116:742–768, 2006.
- Thomas Scanlon. *What We Owe to Each Other*. Belknap Press of Harvard University Press, 1998.
- Andrew Sepielli. What to Do When You Don't Know What to Do. In Russ Shafer-Landau, editor, *Oxford Studies in Metaethics, vol. 4*, pages 5–28. Oxford University Press, 2009.
- Russ Shafer-Landau. Vagueness, Borderline Cases, and Moral Realism. *American Philosophical Quarterly*, 32(1):83–96, 1995.
- Michael Smith. *The Moral Problem*. Wiley-Blackwell, 1994.
- Peter Unger. *Ignorance: A Case for Scepticism*. Clarendon Press, 1975.
- Brian Weatherson. Running Risks Morally. *Philosophical Studies*, 167:141–163, 2014.
- Bernard Williams. Persons, Character, and Morality. In *Moral Luck*, pages 1–19. Cambridge University Press, 1981.
- Timothy Williamson. *Vagueness*. Routledge, 1994.
- Timothy Williamson. *Knowledge and Its Limits*. Oxford University Press, 2000.