

Rational Moral Ignorance

Abstract: What should a person do when, through no fault of her own, she ends up believing a false moral theory? Some suggest that she should act against what the false theory recommends; others argue that she should follow her rationally held moral beliefs. While the former view better accords with intuitions about cases, the latter one seems to enjoy a critical advantage: It seems better able to render moral requirements ‘followable’ or ‘action-guiding.’ But this tempting thought proves difficult to justify. Indeed, whether it can be justified turns out to depend importantly on the rational status of epistemic akrasia. Furthermore, it can be argued, from premises all parties to the moral ignorance debate should accept, that rational epistemic akrasia is possible. If the argument proves successful, it follows that a person should sometimes act against her rationally held moral convictions.

1 Introduction

What should a person do when, through no fault of her own, she ends up believing some false moral theory?¹ Should she do as the theory would recommend? Or should she act against the moral principles she believes in? The following example illustrates the general question.

Frances’s Dilemma: Frances knows several diligent, insightful, and well-informed Kantian ethicists. She regards each one as trustworthy with respect to ethical issues and is rational in so doing. On the basis of their testimony, she comes to have a rational, but false, moral belief. Specifically, she comes to believe that lying is always wrong, no matter the consequences. The next day, an emergency arises: Frances must decide whether to lie to a would-be murderer to save the life of her friend. Should Frances lie or tell the truth? (adapted from Weatherson ms., p. 12)

On the one hand, there is pressure to say that Frances should tell the lie: In general, one should protect one’s friends from murderers. On the other hand, if Frances does tell the truth, a compelling defense of her choice is available: She did what she believed was right, and she had good reason to believe as she did.² There is some tension here.

¹Moral ignorance has been discussed by several authors, including Arpaly and Schroeder (1999, 2014), Arpaly (2002, 2003, 2015), Rosen (2004), Sepielli (2009, 2014, 2016, 2018), Zimmerman (2010), Harman (2011, 2015), Weatherson (2015, ms.), Hedden (2016), Pittard and Worsnip (2017), Johnson King (2020, forthcoming), and Podgorski (2020). This paper’s focus is narrower—*epistemically rational* moral ignorance. See Harman (2011, pp. 460–463; 2015, pp. 75–78), Podgorski (2020), Sepielli (2018), and Weatherson (ms., pp. 12–16) for relevant discussion.

²One way to react is to deny the setup: Frances simply couldn’t have been rational in believing as she did. This reaction might be motivated by a certain highly demanding epistemological picture, according to which people are always rationally required to believe the moral truth. See Guerrero (2007), McGrath (2009), Harman (2011), Sliwa (2012), and Wieland (2015) for discussion of closely related issues. For the purposes of this paper, we’ll assume that it is possible to have rational, false beliefs about what’s right.

Here is the plan. After clarifying the question at hand (§2), we will distinguish three candidate norms that might be thought to apply to someone like Frances (§3): the *transparent* norm (which requires an agent to act in accord with her moral beliefs), the *translucent* norm (which requires an agent to act in accord with the moral beliefs it would be rational to have), and the *opaque* norm (which is not sensitive to an agent's moral beliefs at all, even when those beliefs are rationally held). The transparent norm is widely thought to be something of a non-starter by those on both sides of the debate. So the chief task of the paper is to compare the translucent and opaque norms, with respect to some relevant desiderata. On the one hand, we'll see that the opaque norm more faithfully tracks what a good person would do (§4). At the same time, it must be acknowledged that the opaque norm seems problematically *unfollowable*, in some important sense. This suggestion will receive close scrutiny (§5). In what sense is the opaque norm unfollowable? Is there some salient standard of followability that it fails to meet, which the translucent norm does meet? Ultimately, the answer to this question will be shown to depend upon the rational status of epistemic *akrasia*: If rational epistemic akrasia is possible, then the translucent norm and the opaque norm will—contrary to appearances—turn out to be on a par, with respect to followability. And furthermore it can be argued, from premises all parties to our debate must accept, that rational epistemic akrasia is indeed possible. If the argument is to be trusted, it indicates that a moral norm can require a person to act against her own rationally held moral beliefs without thereby being unfollowable in any useful or important sense.

2 Clarifying the Guiding Question

2.1 Objective and Subjective Requirements

Put simply, our guiding question is: “What should someone like Frances do?” But this question requires clarification. To see this, it will be helpful to consider one very simple reaction to our puzzle, which we'll call the ‘easy answer.’

What should Frances do? That's an easy one! Just look at the setup of the case. From the setup, we can see that Frances's belief is false. Presumably, her belief is false because one is required to lie when the would-be murderer arrives at one's door. So

Frances should tell the lie.³

Though tempting, the easy answer may be a bit too quick. To see why some have found this line unsatisfying,⁴ it is helpful to appeal to a popular distinction between objective and subjective moral requirements.⁵ Alan Gibbard (2005, pp. 343–344) characterizes the distinction as follows:

We can ask what one ought to do in light of all the facts. Alternatively, we can ask what one ought to do in light of available information... Standardly in moral theory, we distinguish what a person ought to do in the objective sense and what she ought to do in the subjective sense.

To illustrate, we can apply the distinction to a simple example.

Doctor's Dilemma: A doctor must decide whether to give a certain medicine to her ailing patient. In fact, the patient possesses a rare, undetectable allergy to the medicine and would be better off without it. But the doctor has no way of knowing about this; she has every reason to think that the medicine will help her patient recover. Should she give him the medicine or not?

If we take the objective/subjective distinction seriously, the question asked above is ambiguous. Objectively, the doctor should not give her patient the medicine, for he is allergic to it. But subjectively, the doctor should give her patient the medicine, for her evidence suggests that it will help him recover.

With this distinction in place, the easy answer may seem less appealing. All sides can agree that, objectively, Frances should tell the lie. *That's* what was stipulated in the setup. But an important question remains: In light of her information, what should Frances do? This question isn't settled by the easy answer, and it is the question that occupies us in this paper.⁶

³ Weatherson expresses sympathy with this line (ms., p. 12).

⁴ See Harman (2015, p. 58) and Sepielli (2016, p. 2952).

⁵ For a sampling of authors who appeal to such a distinction, see Ewing (1953), Brandt (1976), Pollock (1979), Unger (1975), Wedgwood (2003), Setiya (2004), Gibbard (2005), Schroeder (2009), and Harman (2011, 2015). More recently, some, such as Pittard and Worsnip (2017), have expressed skepticism about 'should' or 'ought' is ambiguous between (exactly) these two senses. This issue will be discussed in the next subsection.

⁶ Given the distinction just drawn, one might wonder whether the moral beliefs which generate our puzzle (e.g. Frances's belief that she's required to tell the truth) themselves refer to objective or subjective requirements. In our central example, it is natural to think of Frances as having both of these beliefs. But we can remain neutral on this issue more broadly. In discussing the Frances case and others, we can simply refer to an agent's "belief about what's right" without

2.2 When Should We Disambiguate?

At this point, a distinct worry arises. Though disambiguating objective and subjective requirements may help us disarm the easy answer, it opens us up to a more sophisticated deflationary attack. The thought runs as follows. Once we permit more than one type of requirement into our normative ontology (e.g. objective and subjective), we arguably open the door for further subdivision. Why stop at exactly two?⁷ More pointedly, in our central example, why can't there be, in effect, two subjective 'shoulds'—one which requires Frances to protect her friend and another which requires her to do what she believes is right? If disambiguation is appropriate in the case of the doctor, why not in the case of Frances too? Responding to this line of reasoning will require us to think more carefully about when disambiguation is called for and when it isn't.

Consider, for the moment, a sparse normative ontology, one which holds up the objective 'should' as the only morally important 'should.' What sorts of difficulties does such a picture face? Which roles is the objective 'should' unable to play, if any? Two ideas suggest themselves.

One potential complaint against the objective 'should' is its failure to track what a person of morally good character would do. When we say that a person should do *X*, we may be saying, in part, that a good person would do *X*. And the objective 'should' seems not to respect this. In the doctor example, a good person would surely give her patient the medicine, but the objective 'should' recommends the opposite.

A second complaint against the objective 'should' is that it seems to issue *unfollowable* instructions, in some important sense. If our doctor were to see a long series of patients (some of whom were allergic to the medicine and some of whom were not), the objective 'should' would enjoin her to *give the medicine to all and only those whom it will help*. It would be unfair to expect the doctor to conform to this rule.

So it seems that there are at least two important roles that the objective

specifying whether the belief is about objective rightness, subjective rightness, or both. The arguments found in this paper will apply to all varieties of case.

⁷ See Sepielli (2014), Pittard and Worsnip (2017), and Kagan (2018). See also Jackson (1991), who warns of "an annoying profusion of 'oughts'" (1991, p. 471).

'should' manifestly fails to play, which are restated below.⁸

Role 1 (Character): Track what a good person would do.⁹

Role 2 (Followability): Be followable in some important sense.

Because these roles cannot be played by the objective 'should,' we have reason to seek an alternative, a subjective 'should,' to do this missing work. Must we think that there is one unambiguous subjective 'should' that plays these roles perfectly? Certainly not. Perhaps some degree of sacrifice with respect to at least one of these desiderata is necessary. If so, we might end up with a choice about where and how much sacrifice to permit. We could privilege the character role and end up with one 'should,' we could privilege the followability role and end up with another 'should,' or we could do something else. On this picture, we could end up with two, three, or even infinitely many different subjective 'shoulds' corresponding to different ways of weighing these desiderata. But while we will not take a stand on many subjective 'shoulds' there are, we will at least assume this: If a given norm fares *strictly worse*, with respect to these desiderata, than some rival norm, then, absent some other role for the norm in question to play, the norm isn't one that deserves a place in our (subjective) moral theorizing.

In what follows, we'll consider various norms that might be thought to apply to someone like Frances in order to investigate how effectively these different norms can play the roles we have identified. We will concentrate on one popular norm, which would require Frances to tell the truth. But as we will see, the norm is outperformed by a different norm, which would require Frances to tell the lie.

⁸ Perhaps there's a third role we've neglected, which is to *delineate those acts for which a person cannot aptly be blamed or criticized*. This is certainly an important role, but I think it's a mistake to construe this role as being entirely independent from those already outlined. We will confront this issue at the end of the paper. For discussion of blame in the face of moral ignorance, see Arpaly and Schroeder (1999, 2014), Arpaly (2002, 2003, 2015), Driver (2001), Markovits (2010), Zimmerman (2010), and Shoemaker (2011).

⁹ Presumably, there are some situations where good people would act in different ways. In such cases, multiple options might be permissible. What we're looking for is a norm that tracks what *any* good person would do.

3 Three Norms: Opaque, Transparent, Translucent

What type of norm might apply to an agent like Frances? We'll consider a few possibilities. To get a feel for what these different norms look like, we'll think about how their respective contents vary (or don't vary) depending upon which moral theory is correct. For illustration, we'll focus on two moral theories: Utilitarianism (which requires us to maximize utility) and a rights-based view (which requires us not to violate people's rights). Before we examine the three types of norms which are, in effect, candidates for a subjective 'should,' it is worth first considering the objective norm, for contrast.

True Moral Theory	Objective Norm
Utilitarianism	"Do whatever, in fact, maximizes utility."
Rights-based View	"Do not, in fact, violate anyone's rights."

The objective norm does not depend on an agent's evidence at all. It would require our doctor to withhold medicine from her ailing patient. As we've seen, this norm does not play either of our two roles effectively. It does not track what a good person would do, nor is it remotely followable. We can do better.

The objective norm can be contrasted with various subjective norms, all of which do depend on an agent's evidence in some way. We'll start with what can be called the *opaque norm*. Though the opaque norm does depend on an agent's evidence, it does not depend on her evidence *about which moral theory is correct*. In this way, it is subjective but also 'anchored' to the correct moral theory.

True Moral Theory	Opaque Norm
Utilitarianism	"Maximize expected utility."
Rights-based view	"Minimize the likelihood that you violate someone's rights."

Unlike the objective norm, the opaque norm requires our doctor to give the medicine to her patient. At the same time, the opaque norm would require Frances to tell the lie (at least, given any plausible moral theory).

Though the opaque norm may capture something important, it arguably leaves something to be desired on the followability front. One might complain: "Gee, the opaque norm is great for those who have solved ethics—but not so helpful to the rest of us!" This is why we've labeled it the 'opaque' norm—someone who isn't sure which moral theory is correct (i.e. virtually everyone) won't be sure what the opaque norm requires and therefore won't be

in a position to use it. In response, we might look for a different norm, one which can be useful even in the absence of moral omniscience.

The simplest norm of this type can be called the *transparent* norm, which requires a person to *do whatever she believes is right*—regardless of which moral theory is true.¹⁰

True Moral Theory	Transparent Norm
Utilitarianism	“Do whatever you believe is right.”
Rights-based view	“Do whatever you believe is right.”

Now here’s a very followable norm. Notably, you don’t need to solve ethics to determine what the norm recommends; you only need to be aware of your own moral beliefs. But while conforming to this norm may not be especially difficult, it isn’t clear that this norm represents an ideal worth aiming at. As Elizabeth Harman points out, people who do reprehensible things might well believe those things to be right, but this doesn’t seem to excuse their behavior (2011, pp. 457–458, 2015, p. 65). Andrew Sepielli, writing from the other side of the debate, agrees with Harman on this point and suggests a possible path forward.

It would be absurd to let a Nazi off the hook for heinous acts just because he was very confident in the moral view upon which he based those acts (Harman 2011, 2015). What seems more relevant is how reasonable or well-grounded that confidence is. (2018, p. 6)

Accordingly, some, including Sepielli, have proposed norms which are similar to the transparent norm—but modified to include a ‘reasonableness’ or ‘rationality’ condition, which serves to exclude actors like the imagined Nazi.

In the spirit of Sepielli’s suggestion, here is what we’ll call the *translucent* norm, which requires us to *do whatever it is rational for one to believe is right*.¹¹

¹⁰ Ewing (1953) who may have introduced the notion of the subjective ‘should,’ seemed to have this norm in mind. Gracely (1996) and Gustaffson & Torpman (2014) propose a related norm: Do whatever is recommended by the moral theory of which you are most confident. Schroeder (2009) defends an account of the subjective ‘should’ with the following consequence: If a person believes that, objectively, she should ϕ , then subjectively, she should ϕ . See also Schroeder (2017).

¹¹ For ease of presentation, we have construed the translucent norm as simply as possible, but it is critical to note that a variety of more sophisticated norms have been proposed, which are similar to the translucent norm in spirit. Way and Whiting (2016a) come the closest to defending the translucent norm as stated. They argue that a person should ϕ if she justifiably believes she should ϕ . Kiesewetter (2016) defends a related view. A different approach, favored by Lockhart (2000), Ross (2006), Sepielli (2009, 2018), and Enoch (2014), requires a person to *maximize expected*

True Moral Theory	Translucent Norm
Utilitarianism	“Do whatever it is rational for one to believe is right.”
Rights-based view	“Do whatever it is rational for one to believe is right.”

We’ve labeled this norm “translucent” because it seems to occupy a kind of middle ground, with respect to followability, between the transparent and opaque norms. On the one hand, the translucent norm seems less followable than the transparent norm: While the transparent norm’s prescriptions are simply a function of the agent’s own moral beliefs, the translucent norm’s prescriptions are a function of the moral beliefs it would be rational for the agent to have. On the other hand, the translucent norm seems more followable than the opaque norm: Determining what the opaque norm requires will involve solving ethics—a tall order, to say the least. But determining what the translucent norm requires will only involve rationally assessing one’s own evidence about ethics.

So it may seem that the translucent norm strikes an ideal balance between unpalatable extremes. The transparent norm makes compliance easy but worthless, since even the Nazi can conform to it. The opaque norm makes compliance valuable but exceedingly difficult to attain, since it is anchored to the correct moral theory. The translucent norm appears to be the Goldilocks approach. But this appearance is misleading. In what follows, we’ll assess how effectively the translucent and opaque norms play the roles outlined earlier. First, we’ll examine some grounds for thinking that the opaque norm enjoys an advantage with respect to the character role. Next, we’ll turn to the issue of followability. Intuitively, you might expect that the translucent norm will enjoy an advantage over the opaque norm here. But as it turns out, this tempting thought proves difficult to justify.

objective moral value. As Hedden (2016) observes, this approach has the virtue of being sensitive to the epistemic probabilities of the various moral theories as well as the stakes each theory would assign to the agent’s decision. In our central example, the view would only recommend that Frances tell the truth if her rational confidence in the no-lying principle were high enough to offset the potential moral disaster, in the event that lying is permissible. All of the arguments in this paper that concern the translucent norm will apply equally to this view, too. See Hedden (2016), Kieswetter (2016), Podgorski (2020), or Johnson King (forthcoming) for further discussion.

4 Character and the Virtues

4.1 *Judgments about Cases*

In cases when the two norms conflict, would a good person conform to the opaque norm or the translucent one? One reason for thinking that the opaque norm better tracks what a good person would do derives from intuitive judgments about cases, such as Frances's Dilemma. If Frances really is a good person, the thought goes, surely she'll care more about protecting her friend's life than she'll care about "doing the right thing," and so she'll tell the lie (as the opaque norm requires). The translucent norm, in contrast, requires Frances to tell the truth, which seems like the wrong result. Other examples, such as the case of Huckleberry Finn, can be used to similar effect.¹² If the judgments these cases elicit are to be trusted, it follows that a good person will at least sometimes act against her moral convictions (thereby violating the translucent norm and its ilk). But some have expressed skepticism about whether these cases can do the relevant sort of work, despite the intuitions they elicit.¹³ For this reason, it is worth exploring approaches to this question that do not rely so heavily on intuitions about the very sorts of cases at issue. As it happens, an alternative approach is available—one that involves thinking about how specific character virtues are applied in the face of moral ignorance. Even without relying centrally on contested intuitions, it can be argued the translucent norm does not track what a virtuous person would do as faithfully as the opaque norm does.

4.2 *Argument from the Virtues*

We can start with an innocuous observation: Good people tend to have certain traits (kindness, fairness, generosity, courage, etc.). Though this seems a truism, it turns out to provide some useful leverage. For we can explore how these more specific traits interact with one's (rationally held) beliefs about what's right. Consider the following case.

¹² Arpaly and Schroeder (1999, 2014), Arpaly (2002, 2003, 2015), Harman (2011, 2015), Weatherston (2015). Relatedly, Smith (1994) takes it as a premise that being motivated to do as one deems 'right' constitutes a vicious 'moral fetishism.' But some have disputed this proposition (see Sepielli 2016). Below, we'll explore what follows for the present debate even if we allow, contra Smith, that a concern for what's right may constitute a virtue rather than a vice.

¹³ See Zimmerman (2010, p. 201), Kieseewetter (2016, p. 767), or Johnson King (2020, p. 188).

Public Speaking: Allie has been invited to speak in front of a large audience about a painful and deeply personal subject. The prospect terrifies her; she has never spoken in front of such a large group before. In addition, Allie happens to have a false moral belief: She believes—rationally, we can suppose—that sharing personal stories such as hers with large audiences is morally wrong because it is immodest to do so. As a result, she believes that it would be wrong to share her story publicly. At the same time, she knows that sharing her story could be of great service to others who are struggling with issues similar to those she faced. For this reason, she decides to accept the invitation, despite her reservations.

Some might react to this case by asserting, flat out, that Allie’s decision reflects well on her character. I find this verdict intuitive also. But in the present context, this appeal to intuition does not add much to the examples cited previously. A proponent of the translucent norm could say, in response, that if Allie were really a good person, she would care deeply about doing the right thing and would, hence, not share her story if she truly believed it were wrong to do so.

Even if it is controversial whether Allie’s decision reflects well on her character overall, there is another judgment that is, I think, more secure: Allie’s decision reflects well on her in at least one respect—it was *courageous*. If this is so, it follows that at least one virtue, courage, does not require one to do what one rationally believes is right. After all, Allie clearly acts courageously despite doing something she rationally believes to be wrong.

Is this feature unique to courage? Or do other character virtues share this insensitivity to one’s rationally held beliefs about what’s right? It seems to me that other virtues do share this feature. Take generosity, for instance. Suppose Scrooge rationally believes that giving to homeless people is wrong because such people “are there by their own fault and deserve to suffer the consequences.” Despite his mistaken belief, Scrooge acts generously when he invites a person who is homeless to sleep in his apartment on a cold night. As is true with courage, one can be generous even as one acts against one’s rationally held moral beliefs. For many other virtues, the same pattern seems to hold.

These observations give rise to a question. If individual virtues like courage, generosity, etc. do not require that one does what one deems right, why should the property of *being virtuous*—which presumably involves having these more specific—work any differently? Why not think it, too, attaches directly to a certain sort of conduct—quite apart from whatever moral principles an agent

happens to believe? And, of course, if one can be virtuous without acting in accordance with one's moral beliefs (rational or otherwise), then the translucent norm is not going to track what a good person would do—at least, it won't do so perfectly, and it presumably won't do so as faithfully as the opaque norm would.

In response to this challenge, a proponent of the translucent norm might argue that *being disposed to do what one deems right* is a virtue of its own. They might say:

True, most virtues do seem to allow or even require a person to act against what she believes to be right. But there's a virtue you didn't emphasize, *conscientiousness*, which is all about doing what one believes to be right. Part of being a good person is having this virtue.¹⁴

The suggestion being made above is a controversial one: What some would see as a virtue of conscientiousness, others would see as a vicious (or at least neutral) "moral fetishism."¹⁵ So traditionally, there is something of an impasse here. But for present purposes, I think we can make progress even if we grant that conscientiousness is a virtue. That is, I think we can reply to the objection posed above on its own terms.

For even if we grant that conscientiousness is a virtue, it doesn't follow that the perfectly virtuous person will always conform to the translucent norm. There are still other virtues (courage, generosity, etc.), with which conscientiousness can clash. When such a clash happens, a good person will be conflicted: Her conscientiousness will incline her to do what she deems 'right,' but other virtues will incline her to do the opposite. It's natural to think that these conflicts will be resolved in different ways, depending upon the specifics of each case. But the defender of the translucent norm cannot allow this. To obtain the result that the translucent norm *perfectly, in all cases* tracks what a good person would do, it has to be the case that conscientiousness always wins these conflicts. Is this plausible? In defense of the idea, a proponent of the translucent norm might take either of two approaches—one which is very bold, and one which is more modest.

¹⁴ Hurka (2014, pp. 498–499) argues that it is a virtue to be (defeasibly) disposed to follow one's conscience, where 'following one's conscience' involves acting as one believes right. Johnson King (forthcoming) pursues a similar line. See Arpaly (2015, pp. 147–151) for a reply to Hurka.

¹⁵ Smith (1994, Ch. 3, §5).

First, she might take a bold stance, asserting simply that *conscientiousness is the only virtue*—contrary to what we have been assuming. On this sort of view, it just isn't true that a good person would be kind, generous, courageous, and so on. Fundamentally, a good person would be motivated to do what's right, but not, say, to treat people well. This hard line view is coherent, but I suspect that few would be prepared to go this far.¹⁶

Alternatively, the translucent norm advocate might opt for what would seem a more modest view. She might say that while kindness, generosity, courage and the like are certainly virtues, conscientiousness enjoys a special priority over them enabling it to win all inter-virtue conflicts that may arise. On this picture, a good person would be super-conscientious (always doing what she believed was right when applicable), while also being kind, generous, and so on, insofar as was possible.

In response, one might ask what role these secondary virtues would be playing, if any, on the foregoing picture: If the good person always does what she believes is right, in what sense can she be said to possess other virtues? Is she ever motivated by kindness, for example? Wouldn't conscientiousness be doing all the real work?

One possible answer is that these other 'secondary' virtues, such as kindness, shape the good person's behavior when conscientiousness is silent—that is, when she believes that there are multiple permissible options.¹⁷ The idea can be illustrated by way of the following example.

Libertarian Samaritan: Xena rationally believes in a libertarian moral view, according to which there are no positive duties to provide aid at all. One day, she comes upon someone in severe need and must decide whether to lend him

¹⁶Rosen (2004) comes the closest to defending this view. Rosen is concerned with blameworthiness rather than character. He argues that the only way for a person to be *blameworthy* is by acting against her own moral beliefs. Rosen's view allows that a sincere Nazi could be blameless, which I think few would be willing to accept. Even fewer would allow that sincere Nazis are good people.

¹⁷Interestingly, this answer actually isn't available to those who defend the translucent norm's expectational cousin, which tells an agent to maximize expected objective moral value (see fn. 11). As Hedden (2016) observes, this norm is going to be quite demanding, much more demanding than, say, the translucent norm. Like other norms that require the maximization of some value, this norm will virtually never be silent. If conscientiousness is defined correspondingly, then it too will virtually never be silent, in which case, the 'secondary' virtues would be almost entirely inert in shaping the good person's behavior.

assistance. Given her moral commitments, Xena believes that she is permitted to help him, but she also believes she is permitted not to. What should she do?

If Xena is a good person, she will always do what she believes is right. But here, her moral beliefs are silent. According to her libertarian ideology (to which she rationally subscribes), both options are permissible. Correspondingly, conscientiousness is silent. So even if Xena is maximally conscientious, there is still space for other aspects of her character to show themselves. And plausibly, if she's kind, she'll provide the help.

But there is an observation worth making here. Recall that our aim in this section is to assess how well the translucent and opaque norms respectively track what a good person would do. In the Libertarian Samaritan case, all sides can agree that if Xena is a good person, she will offer assistance. The opaque norm presumably requires her to do precisely this.¹⁸ The translucent norm does not. So in this one case, the opaque norm more faithfully tracks what a good person would do than its rival.

Unless the defender of the translucent norm is to argue boldly that conscientiousness is the only virtue (which hardly involves remaining neutral with respect to first-order ethical questions), it should be acknowledged that the translucent norm does not perfectly track what a person of good character would do, and that the opaque norm seems to fare better on this score. Sometimes, a good person will be motivated by something other than her moral convictions.

5 Followability

Even if the conclusion of the previous section is correct, it doesn't follow that the translucent norm deserves no place in our moral theorizing. We outlined two important roles earlier—character and followability—and it may be that no norm can play both roles perfectly. Though the translucent norm may not track what a person of good character would do as faithfully as the opaque norm does, the translucent norm seems followable in a way that the opaque norm isn't.

¹⁸ Assume, for the sake of argument, that Xena's beliefs are not correct—or in other words, that this is a case of rational moral ignorance. Given this assumption, only the opaque norm will require Xena to provide the assistance.

5.1 Setting the Stage

I want to start by drawing out the *intuition* that the translucent norm is more followable than the opaque one. The content of the opaque norm depends heavily upon which moral theory happens to be correct. If utilitarianism is true, the norm will require us to maximize expected utility; if a rights-based view is true, the norm will instead require us to minimize the chance of committing a rights violation. And of course, ethics is hard. It's hard to know which moral theory is correct. For this reason, it's hard to know what the opaque norm requires. In contrast, we seem to know precisely what the translucent norm requires. It requires us to *do whatever it is rational to believe is right*. Admittedly, it may not be easy to conform to this norm. Evaluating the evidence in a rational manner can be difficult. But still, the translucent norm still seems to be helpful in a way that the opaque norm isn't. When someone tells me to conform to the opaque norm, it's as if they're telling me "Solve ethics!" But when someone tells me to conform to the translucent norm, it's as if they're telling me "Follow your evidence." The latter seems a much more reasonable request.

There are a few simple ways one might try to capture this intuition. First, it might be suggested that we can explicitly state the content of the translucent norm but not the opaque norm. But this assessment oversimplifies the situation. In what sense can we specify exactly what the translucent norm requires? If we're unsure about what it's rational to believe in a particular case, then at best, we'll be able to express the translucent norm's prescriptions as a conjunction of conditional norms.

If it's rational to believe that *X* is right, do *X*; if it's rational to believe that *Y* is right, do *Y*; etc.

Viewed this way, it is evident that we cannot specify explicitly what the translucent norm requires. And moreover, it must be acknowledged that the opaque norm can be characterized in a similar fashion.

If utilitarianism is true, maximize expected utility; if a rights-based view is true, minimize the likelihood of committing a rights violation; etc.¹⁹

¹⁹ A proponent of the opaque norm might resist this construal. After all, if utilitarianism is true, then the opaque norm requires us simply to *maximize expected utility*. If anything, this

There is a structural parallel here. We will have to work a bit harder to identify a relevant asymmetry.

A second approach can be motivated as follows.

To assess how followable a norm is, just ask: *What's the chance I'd succeed in conforming to the norm if I tried?* Successfully conforming to the opaque norm involves either solving ethics (unlikely) or simply guessing which moral theory is correct (also unlikely). But successfully conforming to the translucent norm just involves evaluating my evidence rationally (relatively likely) and then doing what I believe is right.

This is a natural observation. If an agent were to try to conform to both norms in turn, her success rate for the translucent norm might well be greater. But what follows from this kind of followability advantage?

Consider the following norm: "When you roll two dice, whatever you do, please make sure they don't land double sixes!" The probability of conformity with this norm is high, but it is clearly unfollowable in an important sense: It wouldn't be fair to expect someone to conform to this norm. Earlier, when we invoked followability, the motivating thought was that, if a norm is followable, it had better be fair to expect someone to conform to its prescriptions. (Recall that this is what the objective norm was lacking.) So it's not enough to show that, say, the probability of conformity with the translucent norm is higher. What is needed is not a difference in degree, but a difference in kind: To show that the translucent norm enjoys a relevant advantage over its rival, we have to show that it meets some important *standard of fairness* that the opaque norm fails to meet.

5.2 Trackability

What sort of standard would achieve the above goal? Returning to the objective norm is a good first step. For we already know that the objective norm seems egregiously unfollowable in just about every relevant sense. Intuitively, it isn't at all fair or reasonable to expect someone to conform to it. But why not?

Here's an idea. Why is the objective norm unfollowable? Because no one on Earth (or anywhere else) could consistently and reliably conform to it!²⁰ This

observation makes the opaque norm seem somewhat more followable. To make things as difficult on ourselves possible, we will set this point aside.

²⁰ Here, it is advisable to distinguish *conforming to a norm*, on the one hand, from *using a norm to guide one's decision-making*, on the other. Whether one conforms is a purely extensional matter;

thought can be captured by appeal to a standard which might be called *trackability*.

Trackability: A norm is *trackable* if and only if there could be a person who, given her psychological makeup, is certain to succeed in conforming to it.

To illustrate, it will help to think about a toy norm that fails to meet this standard. Suppose a schoolteacher has the following unfair classroom rule: To pass her class, students must correctly predict the result of a fair coin flip. Of course, in any individual case, or in any series of cases, a given student can get lucky. But there's no strategy that is certain to succeed. For this reason, the teacher's rule is not trackable.

For similar reasons, the objective norm is not trackable either. After all, it requires us to choose whichever option will, in fact, produce the morally best outcome. In any given case, this may involve predicting which way a coin will land or guessing on the basis of insufficient evidence whether a patient is allergic to some medicine. Since there is no strategy certain to succeed at such tasks, the objective norm, like the classroom rule, is untrackable.

What about the translucent and opaque norms? Are they trackable? It seems clear that both of them are. Take the translucent norm first. To show this norm to be trackable, we must engineer a being that would always conform to it. Imagine that Lucy is epistemically rational—she always believes what her evidence supports. Furthermore, let's add that Lucy's sole goal in life is, in her own words, "to do the right thing, whatever that happens to be." For this reason, Lucy always does what she believes is right. Given Lucy's rationality and her approach to decision-making, it is clear that she will conform to the translucent norm in all possible situations. So the translucent norm is trackable.

Let's try the opaque norm next. Here, things are a bit more complicated, since the content of the opaque norm depends upon which moral theory is correct. In principle, these contents might differ with respect to trackability. But in practice, I think, the opaque norm will turn out to be trackable for most, if not

whether one is suitably guided depends upon the details of her deliberative activity. See Smith (2012, 2018), Way and Whiting (2016b), and Hughes (2018) for helpful discussion of this latter notion. The first standard we'll examine is conformity-centered; the ones we'll examine thereafter are guidance-centered.

all, moral theories. For illustration, let's fix, by stipulation, a correct moral theory. Specifically, let's stipulate that morality consists in equal parts promotion of welfare and respect for autonomy. With this provision in place, we can formulate the opaque norm explicitly: *Do what you expect will promote welfare and respect autonomy*. Can we engineer someone who is psychologically certain to conform to this norm? It seems clear that we can. Imagine that Opal (like Lucy) is epistemically rational—she always believes what her evidence supports.²¹ Furthermore, let's add that Opal's sole goal in life is, in her own words, "to do whatever promotes welfare and respects autonomy—whether it's right or not!" Opal may not know it or even care, but she will conform to the opaque norm in all possible situations. So the opaque norm is trackable too.

Thus, trackability cannot ground an asymmetry between the translucent and opaque norms. Both norms meet this standard. But trackability, as a standard of followability, arguably leaves something to be desired. It is quite weak, as the following exchange illustrates.

You: I'm in a bit of a pickle. I'm not sure which moral theory is correct. What should I do?

Oracle: Hmm. Here's some advice: *Conform to the opaque norm*.

You: That's entirely unhelpful. I don't know what the opaque norm says! How can you expect me to conform without knowing this?

Oracle: You doubt the norm I've given you? Unlike the objective norm, the opaque norm is trackable. Some beings conform perfectly, in all possible situations.

You: ...

It's easy to sympathize with a desire for more guidance. And, admittedly, the translucent norm does seem to provide it. The translucent norm, it seems, just has to be followable in a more robust sense. Can we find it?

5.3 Rational Executability

Before we examine what seems to me the most promising approach, we must first look at a very strong standard, which we can call *perfect executability*.

Perfect Executability: A norm is *perfectly executable* if and only if a person is certain

²¹ It's not obvious whether this provision is needed. Interestingly, Parfit (2011, ch. 5), in a discussion of practical rationality, suggests that an epistemic condition should be applied to an agent's normative beliefs, but not to her non-normative ones. See Mueller (2017) for relevant discussion.

to succeed in conforming to it by ‘trying’ (that is, by doing what she believes is required of her).²²

This standard is too strong. Both norms fall far short. There is no guarantee that a person will succeed in conforming to the translucent norm if she tries, for she might make an epistemic mistake. And there’s certainly no surefire way to conform to the opaque norm either. So the standard of perfect executability cannot be our dividing wedge.²³ But perhaps it will lead us to one.

There seems to be something special about the translucent norm. Suppose that I am an unabashed “moral fetishist”: I am resolutely disposed to do what I believe is right. Then, so long as I reason well (that is, so long as my beliefs about what’s right are rational), I’m guaranteed to succeed in conforming to the translucent norm. And this is significant. After all, reasoning well is something that it seems appropriate to expect from people, whether they succeed or not.

This insight allows us to describe the standard of followability that will finally drive a wedge between our two norms.

Rational Executability: A norm is *rationaly executable* if and only if the norm is perfectly executable for anyone whose beliefs about what’s required of her are rationally held.

For reasons given above, the translucent norm is rationally executable. And the opaque norm certainly isn’t. After all, someone who reasons well and who ‘tries’ in the relevant sense will tell the truth in the case of Frances, while the opaque norm recommends telling the lie. So, we have found a standard of followability that only the translucent norm meets, as the table below indicates.

	Trackability	Perfect Executability	Rational Executability
Objective Norm	no	no	no
Opaque Norm	yes	no	no
Translucent Norm	yes	no	yes

What follows from this? The finding might be somewhat less momentous than it seems. Yes, it is true that only the translucent norm is rationally

²² A related approach is to say that a norm is followable just in case one is always in a position to know what it requires. But Williamson’s (2000) anti-luminosity argument indicates that a knowledge-based standard such as this would disqualify virtually any norm.

²³ It’s worth noting, here, that the transparent norm actually meets (or at least comes close to meeting) this standard. This reinforces the fact that, if we reject the transparent norm (as authors on both sides advocate), we are all sacrificing at least *some* kind of followability for other desiderata.

executable. But is rational executability a kind of followability *worth wanting*? It is not clear. Suppose we learned that instructions for adhering to a certain special norm were located at the center of a large maze. And suppose we knew that these instructions were so clearly presented that anyone who found them would have no trouble following them exactly. We might say that the special norm is *maze-executable*—that is, the norm is perfectly executable *to anyone who can traverse the maze*. Is that, in itself, valuable? Does it follow that the special norm is robustly followable? Naturally, it depends: How difficult is the maze?

The situation with rational executability is similar. The translucent norm is followable by anyone who believes rationally about certain subjects. But how difficult is it to be a rational believer? If it's quite difficult, then the translucent norm, in effect, inherits that difficulty. Here's the bottom line: Rational executability is a kind of followability worth wanting only to the extent that rational norms are themselves followable in a deep and interesting sense. Are they?

5.4 How Followable Are Rational Norms?

The preceding question is, of course, not easily settled. Its answer will depend upon controversial issues within epistemology—including, for one, the internalism/externalism debate. But as we have just developed some standards for assessing the followability of any given norm, it seems appropriate to proceed by assessing the followability of rational norms using these very standards— while staying as neutral as we can with respect to nearby epistemological debates.

To vindicate the translucent norm, rational norms had better turn out to be *more followable* than the opaque norm. (After all, if there were parity here, the translucent norm would be, in effect, no more followable than its rival.) Do rational norms fare better than the opaque norm?

Consider perfect executability first. Clearly, rational norms are not perfectly executable: Those who try to meet them do not always succeed.

What about rational executability? Might rational norms meet this weaker standard? Things get interesting here. The very question seems viciously circular, but it turns out to be coherent. Somewhat surprisingly, though, there is

reason to doubt that rational norms meet this standard.²⁴

Before we can explain why rational norms may not be rationally executable, we should clarify the question even being asked. Suppose that all of Emily's *epistemological* beliefs (e.g. her beliefs about which theory of epistemic rationality is correct) are rationally held. Further, suppose that Emily is what we might call an *epistemic fetishist*: Her sole goal in life is to be a rational believer. For this reason, whenever she discovers that her preferred theory of rationality recommends that she adopt some attitude toward a certain proposition, she instantly does so. Here's the question we're after: By proceeding in this way, could Emily ever run afoul of epistemic rationality? If rational norms are rationally executable, then she couldn't.

Now, it is very tempting to think that Emily, by her very nature, would necessarily be immune from ever violating rational norms. After all, she always reasons *about* rationality in a perfectly rational manner, and from there, she always proceeds *precisely* as her favored account of rationality recommends. How could this process be anything less than rationally perfect? What else is Emily to do? Well, another thing Emily *could* do is to knowingly flout her own favored account of rationality, thereby doing something she herself thought was irrational. But this is a combination that makes little sense from the inside. It seems irrational, if anything is.

Several authors have argued along these very lines.²⁵ In different words, the issue at hand is whether an agent can ever be rationally required to have an *epistemically akratic* combination of attitudes: that is, a combination of attitudes such that one of the attitudes rationally recommends against the other (for

²⁴ The precise formulation of rational executability is highly relevant here. Note that a norm is rationally executable when it is perfectly executable for an agent who is specified to have rational beliefs *about what's required of her*. This qualification is important. If the agent were specified to be a fully rational believer, then rational norms would trivially meet this standard. But the goal, for the defender of the translucent norm, is to show that rational norms are followable in a robust sense. To show that a norm is robustly followable, one must do more than identify a standard which it trivially meets.

²⁵ See Feldman (2005), Kolodny (2005), Smithies (2012), and Titelbaum (2015). See Greco (2014) for a different sort of case against epistemic akrasia. For defenses of the possibility of rational epistemic akrasia, see Wedgwood (2011), Coates (2012), Elga (2013), Horowitz (2014), Christensen (2016), Dorst (2019), and Weatherson (ms.)—though these authors acknowledge that epistemic akrasia may typically be irrational.

example, believing p while also believing that belief in p is irrational). There certainly seems to be something uncomfortable about an agent's holding such a pair of attitudes. I feel the intuitive discomfort.

And I do think that epistemic akrasia is *usually* and *typically* irrational. But in certain cases, there is reason to think that akratic combinations of attitudes actually can be rational. In a paper discussing these issues, Sophie Horowitz (2014) has argued that while epistemic akrasia is very often irrational, it can be rational in peculiar cases that cause expected rationality and expected accuracy to diverge, in a certain sense. Specifically, she notices that, sometimes, an agent can be forced to choose between a belief that is more likely to be rational (by her own lights) and a belief that is more likely to be accurate (by her own lights). In such cases, Horowitz suggests that an agent should aim at accuracy. If this is right, then in certain circumstances, an agent like Emily, who only cares about rationality, can end up on the wrong side.

Horowitz describes an instructive example illustrating this phenomenon (p. 736), which is inspired by Williamson's case of the unmarked clock (2014, p. 979). Christensen (2016, pp. 413–416) proposes a different sort of case with a similar accuracy-rationality divergence, but which appeals to higher-order evidence rather than vagueness. Both types of examples are well worth examining in detail; if the reader is convinced by either, it suffices for our purposes here. But I want to propose a third type of case that can exemplify the rationality-accuracy divergence, one which is structurally parallel to the example we considered at the outset of this paper.²⁶ In our central example, Frances, rationally comes to believe a dubious ethical principle. In the following example, Emily will rationally come to believe a dubious epistemological principle.

Consider the following view concerning the epistemology of mathematical claims.

Proof Chauvinism: Rationality isn't just about accuracy; it's also about responding to the right reasons. For mathematical propositions, the right reasons are proofs. Though one can obtain very accurate mathematical beliefs by deferring to reliable sources, rationally speaking, one should never be highly confident in a mathematical proposition until one can derive an adequate proof.

²⁶ See Feldman (2005, p. 96) for a similar example, though the case he discusses does not explicitly impose the critical accuracy/rationality divergence.

Though Proof Chauvinism is not an especially plausible epistemological theory, in certain circumstances, someone like Emily might well be rational to believe in its truth. For instance, we can imagine that all of Emily's epistemology professors—who are generally quite reliable—unanimously and independently endorse Proof Chauvinism. Upon learning of this epistemological consensus, Emily could well be rational in coming to believe that Proof Chauvinism is true.²⁷

Suppose Emily finds herself in precisely this predicament. She rationally believes that Proof Chauvinism is true. Later, she reflects on the Pythagorean Theorem, a statement which she'd always been highly confident of, but never learned how to prove. She notices that Proof Chauvinism recommends that she become far less confident of this alleged theorem until she can prove it. Should she remain confident in it or not?

Given Emily's epistemic fetishism, it is clear what she'll do: She'll reduce confidence in the theorem. But epistemically speaking, should she proceed this way? I think not. Emily should flout Proof Chauvinism's advice, despite rationally regarding it as a constraint on rationality. Why?

The Proof Chauvinist allows that one's opinions may become quite *accurate* if one defers to reliable sources but maintains that one's opinions will be *less rational* for such accuracy-enhancing deference. In light of this strange feature of Proof Chauvinism, it seems to me that Emily should instead reason as follows:

If I remain confident in the Pythagorean Theorem, my opinion will probably be less rational, but more accurate. If I reduce confidence, my opinion will be more rational, but probably less accurate. I don't care about being rational; I care about being right. So I will remain confident.

Though it was suggested that epistemic akrasia "makes little sense from the inside," in the imagined situation, the foregoing reasoning seems quite sensible. Given a conflict between what one considers 'more rational' and what one considers 'more accurate,' my inclination is to say that, if one is rational, one will

²⁷Titelbaum (2015) argues against the possibility of this sort of case. But his argument assumes that epistemic akrasia is irrational, which is precisely what I'm angling to deny. So there's something of a dialectical stalemate here: He takes the irrationality of epistemic akrasia as a premise and uses it to establish what he calls the "fixed point thesis," whereas I am—in describing the case as I have—taking the falsity of the fixed point thesis as a premise and using it to establish the possibility of rational epistemic akrasia. His $P \rightarrow Q$ is my $\neg Q \rightarrow \neg P$. Each of us rejects the other's starting point, but, strictly speaking, the inferences we draw are compatible with one another.

favor accuracy. So, in the end, I think that if Emily were rational, she would believe the Pythagorean Theorem and regard herself as irrational for so believing.²⁸ If this is right, then an agent like Emily can run afoul of epistemic rationality. And so rational norms turn out not to be rationally executable.

In response, though, one might appeal to a version of the thought we considered much earlier—that ‘should’ sometimes admits of multiple readings. Sure, perhaps there is a sense in which Emily should flout Proof Chauvinism. But even so, it doesn’t follow that there’s no sense in which Emily would be irrational to proceed this way—simultaneously believing Proof Chauvinism to be true while disregarding its advice. On this picture, the question “What should Emily believe?” has no single correct answer. There may be two varieties of rationality at play here, one which can require akrasia and one which forbids it.

I am sympathetic to the general suggestion that there might not be just one species of rationality. But I dispute the specific contention that it would be irrational, in some important sense, for Emily to flout Proof Chauvinism’s recommendations. To see why, it’s useful to consider why one might think that epistemic akrasia is problematic in the first place. What’s wrong with epistemic akrasia? Horowitz provides what seems to me a compelling story. Why is it bad to believe something of the form “ p , but my belief in p is irrational”? It’s bad because, generally speaking, rational beliefs are *likely to be true*. By endorsing an akratic package, you commit yourself to something of the form, “ p , but my belief in p is unlikely to be true” (pp. 738–740). This suggestion is corroborated by Alex Worsnip (2018), who endorses something very close to an anti-akrasia norm. Notice how similarly he motivates his principle: “As such, to believe while judging oneself to lack sufficient evidence amounts to holding that p is true, but also isn’t especially likely—in light of all the available information—to be true” (p. 17). If this is the motivation for positing an anti-akrasia principle, then the Proof Chauvinism case is the one type of situation where we’d expect the principle not to apply. So I think it’s difficult to resist the intuitive verdict elicited by the case, even if one is open to the thought that there might be multiple rational norms, which can come into conflict with one another.

All of this indicates that rational norms turn out not to be rationally

²⁸ Compare Horowitz (2014, pp. 735–740), Christensen (2016, pp. 413–416).

executable after all. An agent like Emily, who cares only about being a rational believer, can still end up believing irrationally. If this is right, then—barring the discovery of an alternative followability standard not considered here—rational norms are not themselves followable in any distinctive sense. At best, they are merely trackable—a standard which the opaque norm meets also. Since conforming to the translucent norm necessarily involves conforming to rational norms, it seems that the translucent and opaque norms are, with respect to followability, in the same boat.

5.5 *Applying What We've Learned*

So... what lessons should be drawn from all of this? To conclude our discussion of followability, it will be useful to return to the case of Frances to apply what we've learned.

According to the opaque norm, Frances is required to tell the lie despite the fact that her evidence suggests that she should do the opposite. The complaint against this picture was that it makes morality unfollowable. Frances certainly could succeed in conforming with the opaque norm (and most in her shoes probably would), but if she does so, it would only be because *she happens to care about the right things* (in this case, protecting her friend). So unless you've solved ethics, you'd better hope that you, too, care about the right things, if you wish to conform to the opaque norm consistently. This doesn't seem to provide the right sort of guidance.

In response, we considered an alternative, the translucent norm, which promised to be more followable than its opaque rival. But given the preceding discussion, it's not clear that it can be. Why not? To follow the translucent norm, one had better be able to conform to the norms of (epistemic) rationality. But how difficult is that? If the translucent norm is to enjoy a relevant advantage, it had better not be the case that the norms of rationality turn out to be on a followability par with the opaque norm.

But that's precisely what I've argued. Cases of rational epistemic akrasia show that the norms of epistemic rationality are, in a way, quite like the opaque norm. In the case of Emily, for example, the norms of epistemic rationality require her to maintain belief in the Pythagorean Theorem despite the fact that

her evidence suggests that she should do otherwise. (Note the parallel with the opaque norm in the case of Frances.) Emily certainly could succeed in conforming with this epistemic norm (and most in her shoes probably would), but if she does so, it would only be because *she happens to care about the right things* (in this case, truth). Her evidence suggested that she should prioritize believing-for-the-right-reasons over accuracy. But rationality requires her to aim at accuracy anyway. The very same complaints about followability that seemed to plague the opaque norm arise equally here.

The foregoing discussion can be seen as posing a challenge to proponents of the translucent norm. If the translucent norm is to be relevantly more followable than the opaque norm, then it had better be the case that epistemic norms themselves are relevantly more followable than the opaque norm. The challenge is to identify a standard of followability that differentiates them. If this challenge cannot be met, then we can extract the following moral. If one really wants to put forward a norm which proves to be more followable than the opaque norm, the proposed norm had better not appeal directly to epistemic rationality in the way the translucent norm does. Once epistemic norms have been invoked, one has given up the game.

Here is one last takeaway, framed in perhaps a more positive light. Throughout this discussion, we have assumed that, sometimes, what is in fact epistemically (or morally) important can come apart from what it's rational to believe is epistemically (or morally) important. In the epistemic case, we've seen that an epistemically rational agent will be resolutely aimed at truth: that is, she'll endeavor to believe what's true even when she believes that she is rationally required to do something else. If the moral case functions similarly, then a morally good agent will be resolutely aimed at whatever is in fact morally valuable (kindness, justice, etc.): she'll endeavor to promote and pursue these things even when she believes she is morally required to do something else. In the epistemic case, it's not obvious that there is any followability issue. If this is so, and if the analogy is to be trusted, then, contrary to appearances, there may not be any followability issue in the moral case either, and the sort of followability demanded by translucence should not move us.

6 Conclusion

Initially, it might have seemed that the translucent norm strikes the perfect balance between unappealing extremes. On one end, the transparent norm makes compliance easy but worthless. On the other end, the opaque norm makes compliance valuable but quite difficult to attain. The translucent norm occupies a seeming middle ground between these norms—a potential Goldilocks approach. But as we've seen, it is hard to make the case that the translucent norm improves upon the opaque norm in any appreciable way. The translucent norm doesn't track what a good person would do as faithfully as the opaque norm does, and despite appearances, it is no more followable than the opaque norm is. If these are the only two relevant dimensions, the opaque norm seems to outperform its translucent rival.

To conclude, it is worth connecting these observations to something that has been largely absent from the discussion—blame. Earlier, we set forth two important roles we wanted subjective moral norms to play. In focusing on just these two roles, one might have thought that we neglected to consider an important third role for subjective moral norms: *to delineate those acts for which a person cannot aptly be blamed or criticized*. This is certainly an important role. But it is a mistake, I think, to construe this blame role as existing independently of the two roles we concentrated on throughout our discussion. To play the blame role effectively, it is plausible that a norm should both (i) track what a good person would do (since bad people do blameworthy things, and good people don't) and (ii) be followable in such a way that makes it fair to expect someone to conform to it (since this perhaps makes it fair to blame someone for not doing so). If all of this is right, then a norm will be capable of playing the blame role to the extent that it can play the character and followability roles. Since the opaque norm has outperformed the translucent norm with respect to these, the opaque norm seems better positioned to track blameless conduct. Initially, one might have thought that, if anyone would be free from blame, it would be a perfectly rational, maximally conscientious agent. But as things currently stand, this seems not to be the case.²⁹

²⁹ For comments that improved the paper, I would like to acknowledge Nomy Arpaly, Bob Beddor, Ben Blumson, Anna Brinkerhoff, David Builes, Harry Chalmers, Keith DeRose, Jamie

References:

- Arpaly, N. (2002). Moral worth. *Journal of Philosophy*, 99, 223–245.
- (2003). *Unprincipled Virtue*. Oxford University Press.
- (2015). Huckleberry Finn revisited: inverse akrasia and moral ignorance. In R. Clarke, M. McKenna, & A. Smith (Eds.), *The Nature of Moral Responsibility* (pp. 141–156). Oxford University Press.
- Arpaly, N. & Schroeder, T. (1999). Praise, blame, and the whole self. *Philosophical Studies*, 93, 161–188.
- (2014). *In Praise of Desire*. Oxford University Press.
- Brandt, R. (1976). Epistemology and ethics, parallel between. *The Encyclopedia of Philosophy*, 3, 6–7.
- Christensen, D. (2016). Disagreement, drugs, etc.: from accuracy to akrasia. *Episteme*, 13, 397–422.
- Coates, A. (2012). Rational epistemic akrasia. *American Philosophical Quarterly*, 49, 113–124.
- Dorst, K. (2019). Higher-order uncertainty. In M. Skipper Rasmussen & A. Steglich-Peterson (Eds.), *Higher-Order Evidence: New Essays* (pp. 35–61), Oxford University Press.
- Driver, J. (2001). *Uneasy Virtue*. Cambridge University Press.
- Elga, A. (2013). The puzzle of the unmarked clock and the new rational reflection principle. *Philosophical Studies*, 164, 127–139.
- Enoch, D. (2014): A defence of moral deference. *Journal of Philosophy*, 111, 229–258.
- Ewing, A. (1953). *Ethics*. London: English Universities Press.
- Feldman, R. (2005). Respecting the evidence. *Philosophical Perspectives* 95, 95–119.
- Gibbard, A. (2005). Truth and correct belief. *Philosophical Issues*, 15, 338–350.
- Greco, D. (2014). A puzzle about epistemic akrasia. *Philosophical Studies*, 167, 201–219.
- Guerrero, A. (2007)/ Don't know, don't kill: moral ignorance, culpability, and caution. *Philosophical Studies*, 136, 59–97.
- Harman, E. (2011). Does moral ignorance exculpate? *Ratio*, 24, 443–468.
- (2015). The irrelevance of moral uncertainty. *Oxford Studies in Metaethics*, 10, 53–79.

Dreier, Dave Estlund, Arianna Falbo, Tobias Fuchs, Kelly Gaus, Simon Goldstein, Eleanor Gordon-Smith, Yongming Han, Qu Hsueh-Ming, Jess Isserow, Zoë Johnson King, David Killoren, Han Li, Chad Marxen, Mike Pelczar, Abelard Podgorski, Mary Renaud, Richard Rowland, Josh Schechter, Miriam Schoenfield, Neil Sinhababu, Louanne Sungaimin, Weng Hong Tang, Brett Topey, Stephanie van Fossen, Dan Waxman, Margot Witte, Patrick Wu, Alex Worsnip, Leo Yan, Ru Ye, audiences at Yonsei University, Princeton University, and Australian Catholic University, the Faculty Reading Group at NUS, two anonymous reviewers at *Ethics*, and an anonymous reviewer at *Philosophy and Phenomenological Research*. Special thanks to David Christensen.

- Hedden, B. (2016). Does MITE make right?: decision-making under normative uncertainty. *Oxford Studies in Metaethics*, 11, 102–128.
- Horowitz, S. (2014). Epistemic akrasia. *Noûs*, 48, 718–744.
- Hughes, N. (2018). Luminosity failure, normative guidance and the principle ‘ought-implies-can’. *Utilitas*, 30, 439–457.
- Hurka, T. (2014). Many faces of virtue. *Philosophy and Phenomenological Research*, 89, 496–503.
- Jackson, F. (1991). Decision-theoretic consequentialism and the nearest and dearest objection. *Ethics*, 101, 461–482.
- Johnson King, Z. (2020). Accidentally doing the right thing. *Philosophy and Phenomenological Research*, 100, 186–206.
- (forthcoming). Praiseworthy motivations. *Noûs*.
- Kagan, S. (2018). The paradox of methods. *Politics, Philosophy and Economics*, 17, 148–168.
- Kiesewetter, B. (2016). You ought to Φ only if you may believe that you ought to Φ . *Philosophical Quarterly*, 66, 760–782.
- Kolodny, N. (2005). Why be rational? *Mind*, 114, 509–563.
- Lockhart, T. (2000). *Moral Uncertainty and its Consequences*. Oxford University Press.
- Markovits, J. (2010). Acting for the right reasons. *Philosophical Review*, 119, pp. 201–242.
- McGrath, S. (2009). The puzzle of pure moral deference. *Philosophical Perspectives*, 23, 321–344.
- Mueller, A. (2017). How does epistemic rationality constrain practical rationality? *Analytic Philosophy*, 58, 139–155.
- Parfit, D. (2011). *On What Matters*. Oxford University Press.
- Pittard, J. and Worsnip, A. (2017). Metanormative contextualism and normative uncertainty. *Mind*, 126, 155–193.
- Podgorski, A. (2020). Normative uncertainty and the dependence problem. *Mind*, 129, 43–70.
- Pollock, J. (1979). A plethora of epistemological theories. in G. Pappas (Ed.), *Knowledge and Justification* (pp. 93–113), Dordrecht: Reidel.
- Rosen, G. (2004). Skepticism about moral responsibility. *Philosophical Perspectives*, 18, 295–313.
- Ross, J. (2006). Rejecting ethical deflationism. *Ethics*, 116, 742–768.
- Schroeder, M. (2009). Means-end coherence, stringency, and subjective reasons. *Philosophical Studies*, 143, 223–248.
- (2017). Getting perspective on objective reasons. *Ethics*, 128, 289–319.
- Sepielli, A. (2009). What to do when you don’t know what to do. *Oxford Studies in Metaethics*, 4, 5–28.

- (2014). “What to do when you don’t know what to do when you don’t know what to do... *Noûs*, 47, 521–544.
 - (2016). Moral uncertainty and fetishistic motivation. *Philosophical Studies*, 173, 2951–2968.
 - (2018). Decision-making under moral uncertainty. In K. Jones, M. Timmons, & A. Zimmerman (Eds.), *The Routledge Handbook of Moral Epistemology* (pp. 508–521), Taylor and Francis.
- Setiya, K. (2004). Against internalism. *Noûs*, 38, 266–298.
- Shoemaker, D. (2011). Attributability, answerability, and accountability: toward a wider theory of moral responsibility. *Ethics*, 121, 602–632.
- Sliwa, P. (2012). In defense of moral testimony. *Philosophical Studies*, 158, 175–195.
- Smith, H. (2012): Using moral principles to guide decisions. *Philosophical Issues*, 22, 369–386.
- (2018). *Making Morality Work*. Oxford University Press.
- Smith, M. (1994). *The Moral Problem*. Blackwell.
- Smithies, D. (2012). Moore’s paradox and the accessibility of justification. *Philosophy and Phenomenological Research*, 85, 273–300.
- Titelbaum, M. (2015). Rationality’s fixed point (or: in defense of right reason). *Oxford Studies in Epistemology*, 5, 253–294.
- Unger, P. (1975): *Ignorance: A Case for Scepticism*. Oxford University Press.
- Way, J. and Whiting, D. (2016a). If you justifiably believe that you ought to Φ , you ought to Φ . *Philosophical Studies*, 173, 1873–1895.
- 2016b). Reasons and guidance (or, surprise parties and ice cream). *Analytic Philosophy*, 57, 214–235.
- Weatherson, B. (2015). Running risks morally. *Philosophical Studies*, 167, 141–163.
- (ms.). “Do Judgments Screen Evidence?” Retrieved from <<http://brian.weatherson.org/JSE.pdf>>
- Wedgwood, R. (2003). Choosing rationally and choosing correctly. In S. Stroud and C. Topple (Eds.), *Weakness of Will and Practical Irrationality* (pp. 201–230), Oxford University Press.
- (2011). Justified inference. *Synthese*, 189, 1–23.
- Wieland, J. (2015). What’s special about moral ignorance? *Ratio*, 29, 1–16.
- Williamson, T. (2000). *Knowledge and Its Limits*. Oxford University Press.
- (2014). Very improbable knowing. *Erkenntnis*, 79, 971–999.
- Worsnip, A. (2018). The conflict of evidence and coherence. *Philosophy and Phenomenological Research*, 96, 3–44.
- Zimmerman, M. (2010). *Living with Uncertainty: The Moral Significance of Ignorance*. Cambridge University Press.