Just Do It? When to Do What You Judge You Ought to Do

Julien Dutant Clayton Littlejohn King's College London

Abstract

While it is generally believed that justification is a fallible guide to the truth, there might be interesting exceptions to this general rule. In recent work on bridge-principles, an increasing number of authors have argued that truths about what a subject ought to do are truths we stand in some privileged epistemic relation to and that our justified normative beliefs are beliefs that will not lead us astray. If these bridge-principles hold, it suggests that justification might play an interesting role in our normative theories. In turn, this might help us understand the value of justification, a value that's notoriously difficult to understand if we think of justification as but a fallible means to a desired end. We will argue that these bridge-principles will be incredibly difficult to defend. While we do not think that normative facts necessarily stand in any interesting relationship to our justified beliefs about them, there might well be a way of defending the idea that our justified beliefs about what to do won't lead us astray. In turn, this might help us understand the value of justification, but this way of thinking about justification and its value comes with costs few would be willing to pay.

1. Introduction

Few of us think that Agnes should let her conscience be her guide. Not if this means that Agnes always ought to do what she believes she ought to do. Her beliefs about most things are fallible. Her beliefs about some things are unreasonable. If her normative beliefs (i.e., beliefs about what she ought to think, feel, or do) are fallible or potentially unreasonable, nothing of normative significance should follow from the fact that she believes she ought to think, feel, or do something.

Would things be different if Agnes is remarkably reasonable? Suppose all of Agnes' beliefs about oughts were justified. If Agnes gives the matter proper consideration and justifiably judges that she ought to ϕ , could it *then* be that she should do other than ϕ ? If it were proper for Agnes to judge that she ought to ϕ , maybe it wouldn't be improper for her to respond accordingly. Wouldn't it be unreasonable for *us* to expect Agnes to do other than ϕ if, by *our* lights, she's properly drawn the conclusion that she ought to ϕ ? If it's unreasonable for *us* to expect Agnes to do other than ϕ , maybe there couldn't be any reasonable demand for Agnes to do other than ϕ . And if there couldn't be any reasonable demand for Agnes to do other than ϕ , perhaps ϕ -ing would be required or permitted.

On this way of thinking about the normative significance of justified belief, there is a principled connection between the things that justify normative beliefs and the things that justify the attitudes or actions these beliefs are about. On this way of thinking, some form of unity principle must hold:

DOWN: If A justifiably judges that she ought to ϕ , A ought to ϕ .¹

DOWN*: If A justifiably judges that it would be right for her to ϕ , it would be right for her to ϕ .

If these principles do hold, it seems to tell us something interesting about the concept of justification. While it might be an imperfect guide in some respects, it is a perfect guide when it comes to normative matters. Justified beliefs matter, at least in part, because they indicate a proper basis for thinking, feeling, and doing things. Perhaps this thought, inchoate as it is, explains some of the intuitive attraction to the justification rules for practical reason and gives expression to the idea that justification is important because of its guiding role.²

Let's contrast these principles with another kind of principle, one that brings out a very different kind of link between the oughts and the subject's normative beliefs. If we suppose that some sort of unity principle is correct, it might seem that justified normative beliefs are very different from justified non-normative beliefs. Most epistemologists think that Agnes' non-normative beliefs might be justified even if they're mistaken. If they accepted some sort of unity principle, they might propose that the reason that these unity principles hold is *not* that justification is generally a guarantee of truth, but that there's something about the normative truths that would explain why a justified judgment about what to do is, *inter alia*, an accurate judgment about what to do. This might be because the normative truths differ from non-normative truths, in part, because they are things that we stand in interesting epistemic relations to. From this perspective, it might seem that facts about what Agnes ought to do or may do carry implications for what Agnes may believe:

UP: If A ought to ϕ , A has justification to believe that she ought to ϕ .³

UP*: If it would be right for A to ϕ , A has justification to believe that it would be right to ϕ .

According to UP and UP*, Agnes can only be required to ϕ or permitted to ϕ if she has justification to believe that she is required to ϕ . Certain normative facts, on this view, are lustrous.⁴

¹ For defenses of principles in the neighborhood of DOWN, see Gibbons (2013), Greco (2014), Kiesewetter (forthcoming), Littlejohn (2012), Titelbaum (2015), and Way and Whiting (forthcoming). Some writers see these principles as trivial consequences of the factivity of justification (e.g., Littlejohn 2012), but some see them as revealing something surprising about the normative domain.

² For defenses of the justification norm of practical reason (roughly, the idea that a person has warrant for treating p as a reason for the purpose of practical reasoning if they justifiably believe p), see Fantl and McGrath (2009), Gerken (2012), Gibbons (2013), Littlejohn (2012), and Neta (2009).

³ Kiesewetter (forthcoming) offers the most careful defense of UP we have seen in the literature. Smithies (2012) offers a careful defense of a more limited principle, one that links the justificatory status of a belief with higher-order beliefs about justificatory status. For important criticisms of UP, see Srinivasan (2015).

⁴ Smithies (2012) calls a fact or condition "lustrous" iff necessarily, if that condition obtains one has justification to believe that it obtains.

There has been increased discussion of these and related bridge-principles in literature. We want to examine them carefully to see if they can be defended and to see what lessons can be drawn about the concept of justification and its role in our larger normative theories.

2. The Principles

Let's consider the full suite of principles on offer:

UP: $O\phi \rightarrow \langle BO\phi^5$ UP*: $\langle \phi \rightarrow \langle B \rangle \phi$ DOWN: $\langle BO\phi \rightarrow O\phi$ DOWN*: $\langle B \rangle \phi \rightarrow \langle \phi \rangle$

Let's first consider Kiesewetter's (forthcoming) defense of UP. He starts with a simple and intuitive idea. If Agnes, say, ought to do something, it's plausible to think that there are some reasons that make the case for ϕ -ing decisive and that Agnes is able to do what these reasons require of her. It seems plausible that Agnes should also be able to do what's required of her by responding correctly to her reasons. Requirements aren't completely esoteric. If Agnes should do something, shouldn't she be able to do this thing while being moved, in part, by the awareness that this is the thing to do? If we assume that it's generally possible for an agent who ought to do something to do it while believing or recognizing that this is what she ought to do, we can vindicate UP.

Suppose that UP is false and that there are cases where $O\phi$ is true and yet $\sim \lozenge BO\phi$. If Agnes is in such a situation, it wouldn't just be true that $O\phi$, it would also be true that $(O\phi \& O\sim BO\phi)$. This would be a situation in which $O(\phi \& \sim BO\phi)$. This is *not* something that Agnes could do while believing that she was responding correctly to reasons because that would involve, *inter alia*, believing that she ought to ϕ and failing to have that belief. Thus, if it's generally possible for agents who ought to ϕ to do so while believing that this is what they ought to do, we have our vindication of UP.

If Kiesewetter's argument establishes UP, his case for UP should give us what we need to vindicate all four principles. If we assume UP, it isn't difficult to justify DOWN*. Suppose that DOWN* is false. There would be a situation in which $\sim \lozenge \varphi$

⁵ We are reading 'Oφ' as saying that Agnes ought to ϕ and ' \Diamond BO ϕ ' and ' \Diamond ϕ ' as saying that Agnes is justified or permitted to believe she ought to ϕ and Agnes is justified or permitted to ϕ respectively. We shall assume Duality, namely that Agnes ought to ϕ iff Agnes is not permitted not to ϕ . One may want to restrict Duality in the light of the idea that nothing is either permitted or obligatory for rocks and other non-rational creatures. Restricting Duality in such a way complicates the discussion without affecting our results.

⁶ In addition to its crucial premise (if one ought to φ , then it is not logically impossible that they φ while they believe they ought to φ), Kiesewetter's argument further assumes that *oughts* aggregate (if one ought to φ_1 and one ought to φ_2 then one ought to φ_1 & φ_2) and that justified belief in conjunctive *oughts* distributes (if one has justification to believe that they ought to φ_1 & φ_2 then they have justification to believe that they ought to φ_1). The assumptions are not unassailable but fairly standard. As we don't think they are the problematic step in the argument we grant them without further ado.

but $\langle B \rangle \phi$. By Duality, O~ ϕ and by UP, $\langle BO \sim \phi$, and, granting that Duality is known a priori, $\langle B \sim \langle \phi \rangle$. Assuming that one cannot be both justified in believing one thing and justified to believe its negation, it follows that $\sim \langle B \rangle \phi$, contrary to assumption. Thus DOWN* follows from UP together with standard assumptions about Duality and the impossibility of contradictory justifications.

Let's consider UP*. The crucial assumption in the argument for UP was the idea that if Agnes ought to ϕ , it should be possible for her to ϕ while responding correctly to those reasons, something that Kiesewetter takes to involve, *inter alia*, believing that she ought to ϕ . We think that it is equally plausible to think that if Agnes may ϕ it should be possible for her to ϕ while believing that it is acceptable for her to do so. The assumption gives us an argument for UP* that parallels Kieswetter's argument for UP. For, if it were permitted to ϕ but not permitted to believe that it is so, then it would be permitted to ϕ while not believing that it is permitted to ϕ . By the analogue of Kiesewetter principle, it should then be possible for one to ϕ and not believe that it is permitted to ϕ while also believing that it permitted to ϕ and not believe that it is permitted to ϕ . But that would involve both believing and not believing that it is permitted to ϕ , which is impossible. Thus, we think that Kieswetter's argument for UP generalizes to give us UP*, too.

What about DOWN? If there is a counterexample to DOWN, it is a case in which $\langle BO\varphi \rangle$ but $\langle O\varphi \rangle$. By Duality, $\langle O\varphi \rangle$ and by UP*, $\langle O\varphi \rangle$, and, granting that Duality is known *a priori*, $\langle O\varphi \rangle$. But one cannot have justification to believe something and justification to believe its negation, so $\langle O\varphi \rangle$, contrary to assumption. Thus DOWN follows from UP* together with Duality and the impossibility of contradictory justifications.

While a generalized version of Kiesewetter's argument for UP supports the entire suite of principles, we see no good strategy for arguing from DOWN or DOWN* to UP or UP*. If you think of normative status as lustrous (i.e., something that one can justifiably believe to obtain whenever it obtains), you'll think that justified beliefs license certain responses (i.e., whenever the conditions are such that one is justified in thinking that something is right or required, one wouldn't be required to refrain from acting in light of the relevant normative belief). The converse doesn't seem to be the case. Because of this, we thought that we should divide the discussion of these principles into two sections. In §3, we'll examine views that incorporate UP and all the principles that come with it. In §4, we'll examine views that incorporate DOWN or DOWN* only.

3. Lustrous Liability

While the argument for UP has its plausibility, we think that the suite of principles just discussed will be difficult to defend. In our discussion we will assume a kind of factualism, one that says that there are facts about whether a subject ought to ϕ or may ϕ and minimal subjectivity according to which these facts are not completely independent of the perspective of the relevant agents (e.g., we'll assume that what one

⁷ The argument uses the assumption that *oughts* aggregate with *mays* (if it's permitted to ϕ_1 and obligatory to ϕ_2 then it's permitted to $\phi_1 \& \phi_2$) and the assumption that justified belief about conjunctive *mays* distribute (if one has justification to believe that they may $\phi_1 \& \phi_2$ then they have justification to believe that they may ϕ).

ought to believe is not simply the set of all truths-such an *ought* trivially entails the principles but is not the one at stake in debates over bridge principles.)

We think that cases like the following pose real problems for the package of principles just introduced:

Diving: Agnes dives into the water. After one second under the water, it's not the case that you ought to jump in to try to save her. After 600 seconds, however, you ought to jump in if you haven't already. As time passes, there will be the first second that you ought to jump in. Suppose that that second is t_n .

At t_{n-1} things were very similar to how things were at t_n but different enough that it wasn't true then that you ought to dive in to try to save your friend. When we think about these times, we learn two things. First, that there was some difference between the cases that you couldn't hope to detect was part of what makes it the case that the first second where it's true that you ought to dive was t_n . Second, that such a difference couldn't account for the resultant epistemic difference that UP tells us distinguishes the cases.

The reader is hopefully familiar with the details of Williamson's (2000) antiluminosity argument. We can deploy that argument to establish that you couldn't know at t_n that you ought to dive in and try to see if Agnes needs assistance, but that only takes us so far. It shows that the fact that you ought to dive in is not *luminous*, but doesn't show that it is not *lustrous*. To establish this further fact, we need a principle that links ignorance and irrationality. A promising principle would be this one:

NO KNOWN UNKNOWNS: If one is in a position to know that one is not in a position to know p, one doesn't have justification to believe p.

It doesn't take much reflection to see that one isn't in a position to know at t_n that one ought to jump in and that this fact is one that's known to be unknown in the relevant sense. NO KNOWN UNKNOWNS explains the apparent impermissibility of believing both that p is true and that one isn't in a position to know whether p.

Cases like this invite all kinds of interesting but ultimately distracting objections. Here is one lesson that we think we should take from this example. At no point in defending our bridge principles principles has anyone thought to challenge the epistemicist view of vagueness. That is to say, the arguments for UP haven't been offered alongside any arguments against epistemicism. If we're spotted this (admittedly controversial) assumption about vageuness, we can show that UP has some striking epistemological implications. Intuitively, the subject could not discriminate case t_n from t_{n-1} . However, if t_n represents the first moment when you're

⁸ See Smithies (2012). If the reader wants further defense of the possibility of unknowable obligation, they should read Sorensen (1995). For a helpful defense of Williamson's anti-luminosity argument against a variety of critics, see Srinivasan (2015b).

⁹ The principle is not completely uncontroversial, but it does enjoy reasonably widespread acceptance by internalists and externalists. See Adler (2002), Bird (2007), Huemer (2007), Gibbons (2013), McKinnon (2015), and Sutton (2007) for defense. For an opposing view, see McGlynn (2011).

under the obligation to jump in, you'd be justified in believing that you're now under this obligation even though you would presumably take this case and surrounding cases to be indistinguishable.

We suppose that someone who defended UP might try to brave this one ought and say that the subject *can* discriminate between the cases. They could say that it only becomes true that you ought to jump in when it is clear enough that there is good reason to do so. We think that this response misses the bullet but doesn't avoid the objection. Someone who takes this line has to decide whether they accept NO KNOWN UNKNOWNS. If they do, they'd presumably have to say that you aren't in a position to know that you don't know (even though you are) *or* that you *are* in a position to know.

We don't think that this is a good line to pursue. If someone tries to brave this out, they don't just have to say implausible things about t_n but also t_{n-1} . If t_{n-1} represents the last moment when you don't have an obligation to jump in, DOWN tells us that you wouldn't be permitted to believe that you ought to jump in. UP would then tell us that you'd be justified in believing that you wouldn't be permitted to believe that you ought to jump in.

Let's consider a second case:

Feed the Children: Agnes is faced with 1,000,000 hungry children. She knows that all but one needs to be fed today. There is one particular child that she cannot feed, but she doesn't have any way to identify the child.

We might think of this as a kind of lottery situation insofar as Agnes has the same grounds for believing each particular child is the one that cannot be fed. Because Agnes cannot know which child she cannot feed, she wouldn't know of any particular child whether she ought to feed this child. We will assume that 'ought' implies 'can' so that Agnes cannot know that she ought to feed some particular child because she cannot know whether she can feed it.) Using NO KNOWN UNKNOWNS, we can show that Agnes isn't justified in believing that she ought to feed a hungry child that has stepped forward with her empty bowl. In turn, this would be taken to show that Agnes has no obligation to feed this child. (This is what UP tells us.)

¹⁰ This claim about lotteries is not uncontroversial, but see Nelkin (2000) for an important defense for our claims about knowledge and justification.

we didn't have to construct the case using the 'ought' implies 'can' principle. We could have constructed the case using the 'ought' implies 'not merely supererogatory'. We could imagine a revised version of Feed the Children where Agnes knows that there are many hungry children and one just slightly peckish child. She might also know that it is merely supererogatory to feed a child that is known to be only slightly peckish. Under these conditions, the case would still exhibit the relevant structural features (i.e., she'd know that all but one child ought to be fed and that she cannot know with respect to any particular child whether she or he ought to be fed). We think it's clear in this case that the fact that Agnes knows she doesn't know whether she ought to feed Alice doesn't put Agnes in a position to know that it's false that she ought to feed Alice. She ought to feed the child. Alice is hungry and Agnes can help.

We don't see many good options for philosophers who accept UP and think of all *oughts* as lustrous. It would be outrageous for someone to say that there was no obligation to feed the hungry children just because, say, they knew that it would be supererogatory to feed one of them and couldn't tell which one they weren't required to feed. We think that the lesson to draw from this is simply that facts about epistemic status and practical requirement do not interact in the way that UP suggests. Once the underlying factors that determine what some subject ought to do finish their business there is always the possibility that they've done their job without providing the relevant agents with a clear path to identifying both the things that ought to be done and the things that may be avoided.

Of course, someone could contest this picture and say that the factors can only do the business of determining what ought to be done, felt, or thoughts by thereby determining that certain things may be believed or should not be believed, but then cases like the previous one and this one suggest that either our sense of when it is rational to believe something is highly unreliable or suggest that it is far too easy to subvert obligation. Consider:

Feed the Children (II): At t_1 Agnes believes on strong evidence that she ought to feed the first child, ought to feed the second child, ought to feed the third child, ... and ought to feed the millionth child. At t_2 Agnes is told by an incredibly reputable source that there is precisely one child that she should not feed. Other things equal, Agnes knows that this source is a very good source of evidence.

We think that this case raises a number of potential problems for UP.

First, it seems pretheoretically that there should be two ways of filling out the details of this case. In the first version of the case (Feed the Children (IIa)) Agnes' source *believes* on exceptionally strong evidence that there is one child that Agnes isn't required to feed but this belief is mistaken. In the second (Feed the Children (IIb)), Agnes' source *knows* that there is precisely one child that Agnes isn't required to feed. It's not difficult to see how this kind of difference could arise in cases

nature of the grounds for the preface belief prevent that belief from being justified, but we think that these kinds of testimony cases get around that problem. While

¹¹ For an interesting discussion of the different ways that these kinds of preface cases create defeating pressure, see Moeller (2015). For a general discussion about the assumptions that let us generate preface-like situations, see Worsnip (forthcoming). Some writers (e.g., Smith (2016)) try to block the preface by saying that the statistical

involving non-normative beliefs, but it's difficult to see in the present framework to keep these cases distinct. To see this, consider this argument (focusing on what happens prior to t_2 for the time being) for two surprising claims about what Agnes should do and may believe:

- 1. Agnes has the same evidence the two cases. She has sufficiently strong evidence to believe that she ought to ϕ_1 , Agnes has sufficiently strong evidence that she ought to ϕ_2 , Agnes has sufficiently strong evidence that she ought to ϕ_3 , ..., and Agnes has sufficiently strong evidence that she ought to $\phi_{1,000,000}$. [Suppose]
- 2. So, Agnes has justification to believe that she ought to ϕ_1 , Agnes has justification to believe that she ought to ϕ_2 , Agnes has justification to believe that she ought to ϕ_3 , ..., and Agnes has justification to believe that she ought to $\phi_{1,000,000}$. [Assumes that strong evidence suffices for justification.]
- 3. So, Agnes ought to ϕ_1 , Agnes ought to ϕ_2 , Agnes ought to ϕ_3 , ..., and Agnes ought to $\phi_{1,000,000}$. [(P2), DOWN]
- 4. So, Agnes ought to $(\phi_1 \& \phi_2 \& \phi_3, ..., \& \phi_{1,000,000})$. [Aggregation]
- 5. So, Agnes has justification to believe that she ought to $(\phi_1 \& \phi_2 \& \phi_3, ..., \& \phi_{1,000,000})$. [P4, UP]

It looks as if we can use an anti-skeptical assumption about the kind of evidence Agnes could have and an evidentialist assumption about justification to establish that Agnes ought to do all the things she has strong evidence to believe she ought to do. Her evidence couldn't be misleading, so her source couldn't see something that she missed that would show that she's not required to perform one of the relevant acts, so we've shown that (3) follows from (1) and that Feed the Children (IIb) is impossible. We think that this is deeply implausible. It seems rather obvious that someone with strong evidence for beliefs about what she ought to do might have misleading evidence and that some external observer might be perfectly placed to see that this

some writers (e.g., Ryan (1991)) think that the presence of the justification for believing the preface assertion indicates an absence of justification for believing some of the claims in the body of the book, some of us writing this paper side with Easwaran and Fitelson (2014) in thinking that the author could be rational in believing all the claims in the book even if they know what all the claims are and know them to be inconsistent.

¹² We'll assume that Agnes has sufficiently strong evidence to believe p if she has just the evidence of someone who is in a position to know that p. While being in a position to know p requires the truth of p, having the evidence of someone in this position arguably does not. We're supposing that Agnes has the evidence of some possible subject who could *come* to know the relevant propositions given the evidence that they had (and further factors) and this wouldn't require entailing evidence if we assume E=K (although it's a consequence of E=K that the subject who came to know p would thereby have evidence that entailed p and evidence that they could only have if p.)

subject shouldn't be required to do something even if her evidence strongly suggests that she should. We'll call this problem the *problem of infallible evidence*.

Once we've established (2), we get an argument for a surprising claim about what Agnes may believe. Because Agnes is justified in believing one million propositions about what she ought to do, we can show that she is justified in believing a massive conjunctive proposition, one that captures everything that she ought to do. Agnes isn't special in this regard, mind you. In the present framework, we're all justified in believing a giant conjunctive proposition that captures every requirement we're under. (Don't think for a second that this problem can be mitigated by trying to keep our to-do lists short and challenging the assumption that we have sufficient evidence to believe lots of propositions about what we ought to do because this merely shifts the entries on the list from things like, say, pick up the dry cleaning and bury the cat to refraining from believing that we ought to pick up the dry cleaning and refraining from believing that we ought to bury the cat.) Naturally, we'd think that the probability of these massive conjunctions would be highly improbable and not plausible candidates for justified belief or knowledge. We'll call this problem the problem of evaporating uncertainty.

Without rejecting UP, how could we tackle these problems? The only strategy we see for tackling the problem of infallible evidence would involve denying the move from (1) to (2) and rejecting the evidentialist assumption that Agnes would have justification to believe p if she had sufficient evidence to believe p. At least one of us does think that we should reject this evidentialist assumption, but this is a minority position and one that's even more implausible in the present framework. If someone wanted to defend UP by denying the evidentialist assumption, they'd have to say that in Feed the Children (IIb) Agnes ought not believe a proposition that her evidence supports even though she may believe this proposition in Feed the Children (IIa). This implies, in turn, that she'd have justification to believe that she ought not believe the relevant proposition even though her evidence supported this proposition. This, we think, is a very strange view. The view only gets worse if you think about the further evidentialist assumption that says that if Agnes is justified in believing something, she has evidence that supports it. If she ought to refrain from believing p, UP tells us that she's justified in believing this. In turn, she'd have evidence that she shouldn't believe p. That's bananas. Even if you like bananas, you then have to decide whether Agnes has this evidence in Feed the Children (IIa). If not, there's a difference in evidence between the cases just because a proposition that Agnes might believe is mistaken or Agnes has evidence that she shouldn't believe a proposition that she may believe in Feed the Children (IIa). This conflicts with the further evidentialist idea that you cannot justifiably believe something against the evidence. (If Agnes had this evidence in Feed the Children (IIa), it's still supposed to be true that she'd be justified in believing the relevant propositions about what she ought to do in spite of this evidence.) We think this line of response is hopeless.

It might seem, then, that the best thing for someone to do if they accept UP is simply embrace the conclusion. We'll have more to say about this idea momentarily, but let's just pause and think about the implications of this idea. Once we adopt this view, it starts to look like the framework isn't neutral on an important question, a question about priority. Our principles tell us that there are systematic connections between facts about what a subject may/ought to believe and may/ought to do. Should we think that the facts about what a subject ought to do, feel, or think are the

ones that have priority or should we think that facts about what a subject may/ought to believe about normative matters are the ones that determine the remainder? Once we 'solve' the problem of infallible evidence by adopting a view on which Agnes' evidence about normative matters is infallible if sufficiently strong, it looks as if we're embracing a view on which the stuff that justification is made of somehow makes it the case that she ought to do, feel, and believe the things that she's justified in thinking are to be done, felt, or thought.¹³ This is a substantial commitment to incur, but if people are willing to incur it, we hope that they'll acknowledge its potential costs and defend it.

We'll return to the problems of infallible evidence and evaporating uncertainty below. Before we move on, we'd like to discuss one further problem that preface-like cases create for UP. As we've seen, if you accept some weak assumptions about the link between justification and evidential support, it looks as if the two versions of the case should get roughly the same treatment. Thus far we've focused on what happens prior to the source's assertion that one of Agnes' normative beliefs is mistaken. What happens after Agnes is told that one of her beliefs is mistaken? Here's a wish list. We'd like to be able to accommodate three things. The first is that Agnes' evidence is the same in both cases at t_1 , the set of things she ought to do in these cases differed at t_1 and also t_2 . The second is that the testimony from Agnes' source doesn't have any effect on what she ought to do (i.e., it doesn't generate any new obligations she hadn't been under at t_1 or subvert any previous obligations she had been under at t_1). The third is that the testimony that prevents Agnes from believing with justification the massive conjunction capturing everything she ought to do but doesn't thereby prevent her from justifiably believing a massive number of the conjuncts. While we think that these are natural ways of thinking about what happens in these kinds of normative preface-style cases, UP forces us to abandon them and thus prevents us from giving a plausible description of what's happening in these kinds of cases.

3.1 Evaporating Uncertainty

We saw in the previous section that it's difficult to see in the present framework much room for normative uncertainty, uncertainty about what we may/ought to do, feel, and think. UP tells us that whenever Agnes ought to do something, she has justification to believe this. And UP* tells us that whenever she may do something, she has justification to believe this, too. The reader might wonder why we should think that we have this justification if, as it seems, normative uncertainty is not only ubiquitous but also nearly impossible to remedy. These strike us as good questions. We'll leave it to the defenders of UP to see if they can answer them.

Because we don't know what the answers would look like, we won't be able to critically assess them. What we can do, though, is point to some structural considerations that suggest that it will be difficult to square UP and UP* with some plausible points about justification, knowledge, and evidential support. Let's focus on the case of belief. If UP is correct, whenever there is some doxastic state ('D') that you ought to be in, you have justification to believe that you ought to be in this state (i.e.,

¹³ For an argument against such 'shifty' views about 'ought', see Littlejohn (2013). See Harman (2011) for a helpful discussion of the irrelevance of the difficulty of identifying obligation to our overall theory of what our obligations are.

if OD, \delta BOD). If UP* is correct, you have justification to believe that you have justification to be in D if you do indeed have this justification (i.e., if $\langle D, \langle B \rangle D$).

It's difficult to square these principles with some plausible claims about failures of iteration:

> ~KK: If you know that p, it doesn't follow that you're in a position to know that you know that p.¹⁴

> ~EE: If you have sufficient evidence that p, it doesn't follow that you have sufficient evidence that you have sufficient evidence that p.¹⁵

We can come at this from two directions. First, let's suppose that we know that there's justification to believe p iff there is sufficient evidence to believe p. Suppose Agnes has justification to believe p. If she does, UP* says that she has justification to believe that she has justification to believe p. If we combine this with our evidentialist assumption, we get the result that Agnes is justified in believing p only if she has justification to believe she has sufficient evidence to believe p. If she has justification to believe this, UP* says that she has justification to believe that she has justification to believe this. Combined with our evidentialist assumption, it should follow that Agnes has this justification only if she has justification to believe that she has sufficient evidence that she has sufficient evidence for p. As we chain these together, it's clear that whenever Agnes has justification to believe something, she has justification to believe that she has sufficient evidence that she has sufficient evidence that she has sufficient evidence (etc.) that p. Second, it seems that if Agnes knows just a bit of epistemology, she'll know that there are cases where she can have justification to believe p even though she knows that she's in a case where KK fails at some order of higher-order knowledge. Since NO KNOWN UNKNOWNS implies that she cannot have justification to believe both p and that she doesn't know whether p, we're headed for a conflict. Agnes should be able to predict that when she's in a case where knowledge or sufficient evidence will not iterate that this would preclude having justification for believing the relevant first-order proposition. This, in turn, radically and implausibly constrains the set of things that Agnes justifiably believes. In the relevant range of cases the (putative) grounds for denying justification at the firstorder would only be potential threats at higher-orders. She'd know that these threats to higher-order justification wouldn't make it easy for her to get the first-order propositions wrong.

3.2 Infallible Evidence

Let's return to the problem of infallible evidence. In discussing Feed the Children (II) we saw that there was a simple argument from an evidentialist assumption about justification and DOWN to the impossibility of having strong but misleading evidence about what you ought to do. We saw that some writers might embrace this and adopt a view on which the factors that ensure that you have sufficient evidence

¹⁴ For arguments against KK, see Williamson (2000).

¹⁵ Even the philosophers who would argue that evidence of evidence is evidence seem to agree that someone can have sufficient evidence that they have sufficient evidence for p even when they don't have this evidence and think that they can have sufficient evidence for p without having strong evidence for thinking that they have it.

for believing normative propositions thereby ensure that the relevant propositions are true. We think that this position will prove incredibly difficult to defend.

To see this, consider this *prima facie* plausible argument:

- 1. My feasible (exclusive) options are just these: ϕ_1 , ϕ_2 , and
- фз.
- 2. The best feasible option is to ϕ_1 .
- 3. If ϕ_1 is the best thing I can do, I ought to ϕ_1 .
- C. So, I ought to ϕ_1 .¹⁶

It would seem that Agnes could have sufficient evidence to believe each of these premises. However, we've seen that we can use UP and DOWN to show that if she believes on sufficient evidence that she ought to ϕ , she ought to ϕ . It seems there are two problems with this, though.

The first problem has to do with potentially misleading evidence about the value contained in options. Even if we work with some more subjective way of ranking options (e.g., in terms of expected value, not actual value) we run into difficulties when subject's have misleading *evaluative* evidence (e.g., evidence that indicates that something that lacks value has it, a lack of evidence indicating that something that is valuable has value, evidence that misleads about how much value something has, etc.) we'll run into cases where someone could have strong evidence for (2) when (2) strikes us as false. It seems fetishistic to say that someone in possession of misleading evidence about the moral permissibility of essentially non-procreative sex should act in line with this evidence if, by our lights, we see no reason to refrain from having essentially non-procreative sex.¹⁷

The second problem has to do with mistakes about ability and opportunity. It seems possible that someone could have strong but misleading evidence about their abilities or about the opportunities they'll be provided. Couldn't there be some possible world in which Agnes is in a position to know (1) and (2) and another possible world in which she has the very same evidence but (1) is mistaken? In *that* world, if she reasoned in the way just sketched, she'd believe on seemingly sufficient evidence that she ought to ϕ . However, we know that she couldn't ϕ , so we know that (C) would be false.

Some philosophers think that it's important to vindicate the intuition that if two subjects reason from just the same evidence in just the same ways they'll end up with beliefs that are justified to the same degree. If we build this in, we get some very bad results. Because it's false in the envisaged second world that Agnes ought to ϕ , it couldn't be that she had sufficient evidence to believe (1)-(3). Because she *would* have sufficient evidence to believe (1)-(3) if she had the same evidence as someone who

¹⁶ While Way and Whiting (forthcoming) like the idea that justified beliefs about oughts are infallible, they suggest that they might want to restrict the factivity of justification to normative propositions. This kind of argument, offered in Littlejohn (2012), was designed to show that such restrictions are difficult to defend. Kiesewetter (2014) anticipates this kind of worry and suggests restricting DOWN to actions that the agent ought to perform, but the difficulty with this maneuver is that we've shown that DOWN (in its unrestricted form) can be vindicated using his (forthcoming) argument for UP.

¹⁷ For further discussion of the relationship between evaluative evidence, justification, and overall obligation, see Littlejohn (2013) and Zimmerman (2010).

was in a position to know (1)-(3), it seems to follow that nobody with her evidence *could* be in a position to know (1)-(3). The upshot, then, is that Agnes couldn't have had sufficient evidence to believe (1)-(3) even in the best circumstances, not if she could have had that evidence in a bad case.

One bad implication of all of this (again, assuming that sufficient evidence suffices for justification and that UP and DOWN are correct) is that Agnes only ought to ϕ if facts about whether Agnes can ϕ supervene upon her evidence. What this means, then, is that it cannot be true to say Agnes ought to ϕ two seconds from now or that Agnes ought to wiggle her finger. At best, what Agnes ought to do will be limited to mental acts she could perform right now. We think that's a bad result. We also think that it's a bad result that Agnes couldn't have strong evidence for believing (albeit mistakenly) that her obligations aren't limited to the here and now. (If she had sufficient (but misleading?) evidence that she ought to wiggle her finger or do something two seconds from now, we'd get the result that she ought to do things like move her body or perform actions a few seconds from now and the ability to do such things doesn't supervene upon her evidence.)

The framework commits us to some skeptical implications about evidence and justification that we find implausible and some implications about an agent's obligations that we find counterintuitive. Because of this, we might reconsider the original argument for UP. The key assumption in that argument is the assumption that it must be possible to 'respond correctly' to the reasons that require you to ϕ . This involves not merely the ability to ϕ , not merely the ability to ϕ for those reasons, but also the ability to ϕ while believing that you ought to ϕ for these reasons. Why should we accept this assumption about the possibility of responding correctly? This assumption might be appealing to people who characterize normative reasons as guides, but we think that the considerations above reveal how problematic this talk of guidance might be. We see two potential ways of cashing out this metaphor to make sense of the view that reasons serve as guides. On the one hand, we might think that reasons are guides in the sense that they establish guidelines, lines we shouldn't cross. On the other, we might think that reasons are guides in the sense that they don't do their job unless they furnish us with guidance we can follow. There are some philosophers who clearly want reasons to play both roles or wear both hats, who try to argue that nothing could do the first thing unless it did the second. Forcing reasons to play both roles, we think, leads to difficulties of the sort sketched above. If we think of reasons as guides that work like clues, evidence, or indicators, it's natural to think of such things as things that have the potential to lead us astray. That's possible only if reasons don't wear both hats or play both roles. If you think, as we do, that the primary job of normative reasons is to establish guidelines (i.e., lines that we shouldn't cross) it's natural to think that a reason can do its job even if it couldn't lead us from where we are to where we ought to be. We wouldn't like the conclusion of the argument for UP and we wouldn't like the starting premise.

4. Infallibility and Fixed-Points

As noted earlier, the case against UP and UP* doesn't constitute a case against DOWN or DOWN*. If we reject UP and UP*, we remove one potential motivation for DOWN and DOWN*, but other motivations for these principles have been given in the literature.

One could try to justify DOWN and DOWN* by appeal to the thesis that justification is factive. If it's not possible to justifiably believe falsehoods, it's not possible to justifiably believe falsehoods about whether you may ϕ or ought to ϕ . Few epistemologists would find this sort of argument convincing, but a more subtle justification for a principle in the neighborhood of DOWN and DOWN* is possible. Some philosophers think that we're always required to satisfy this Enrkatic Requirement:

Enkrasia: O~(φ & BO~φ).¹⁸

Titelbaum (2015) has shown us that we can derive a special case of the Fixed-Point Thesis from this requirement:

Partial Fixed-Point Thesis: $O\phi \rightarrow \sim \langle BO \sim \phi$.

The argument is straightforward. Suppose that you ought to ϕ but are permitted to believe that you ought to $\sim \phi$. By aggregation of *oughts* with *mays*, you are permitted to both ϕ and believe that you to $\sim \phi$, which contradicts Enkrasia. Thus Enkrasia entails that if you ought to ϕ then it's not permitted to believe that you ought not to ϕ , that is, the Partial Fixed-Point Thesis. The Partial Fixed-Point Thesis allows us to rule out the possibility that certain attitudes about ϕ -ing could be rational by appeal to facts about the normative status of ϕ -ing given certain facts about the normative status of beliefs about ϕ -ing.

There is no clear argument from the Partial Fixed-Point Thesis to DOWN or DOWN*, but there is an interesting principle in the neighborhood of DOWN and DOWN* that we can justify using the Partial Fixed-Point Thesis:

$$DOWN^{**}: \langle BO\phi \rightarrow \langle \phi \rangle$$

For contraposing the Partial Fixed-Point Thesis, we get $\langle BO\phi \rightarrow \sim O\sim \phi$, and by Duality, DOWN**. Thus, we can use the Partial Fixed-Point Thesis to vindicate DOWN**.

This raises some interesting questions. How could there be a principle like DOWN**? If we assume the factivity of justification, the explanation is elementary. Anyone who is justified in believing $\Diamond \varphi$ would be justified in believing $\Diamond \varphi$, so the factivity of justification would give us the explanation we need. If we *don't* assume that justification is factive the principle might seem mysterious. The Partial Fixed-Point Thesis and DOWN** tell us that facts about what a subject ought/ought not do constrain facts about what a subject could justifiably believe *even though* there is no necessary connection between the justificatory status of beliefs in general and the accuracy of these beliefs. Why, then, would the conditions that determine the ranking of options have any bearing on justificatory status about the normative status of these options?

We saw in the above sections that certain explanations will not work. We cannot appeal to principles like UP or UP* to try to explain the links that have to hold for DOWN** or Partial-Fixed Point to be genuine principles. You can't say, for example, that potential counterexamples to DOWN** are ruled out because the only way for factors to make it the case that a subject ought not ϕ is for those factors, inter

¹⁸ For defenses of the requirement, see Gibbons (2013), Greco (2014), Littlejohn (2012, forthcoming), Smithies (2012), and Titelbaum (2015). For critical discussion, see Lasonen-Aarnio (MS).

alia, to provide a subject with the kind of evidence they'd need to form a justified belief about whether a subject ought not ϕ .

We could speculate about possible explanations as to how non-factive views of justification could be combined with principles like DOWN**, DOWN*, or DOWN, but we think that given space constraints and our limited imaginations, it might be best not to do that. Instead, we want to point out something that seems important, and that is that many of the problems that we discussed in the previous section would seem to arise for any view that combines a non-factive account of justification with DOWN**, DOWN*, or DOWN.

First, we've already seen that any view that incorporates DOWN faces a series of difficulties. If we assume that sufficient evidence is necessary and sufficient for justification, we face problems about potentially misleading evidence about a subject's abilities and opportunities. To accept DOWN and these evidentialist assumptions, we have to accept that a subject has sufficient justification to believe that they ought to ϕ only if the fact that they can ϕ supervenes upon their evidence. This, in turn, makes it impossible for subjects to have justified beliefs about what they can do moments from now and impossible for subjects to have justified beliefs about what they can do to effect changes in their environments. The skeptical costs of such a view are incredibly steep.

Second, while these worries might not arise for DOWN** or DOWN*, these views face difficult explanatory challenges. Suppose Agnes should not ϕ . According to DOWN**, she could not justifiably believe that she ought to ϕ . According to DOWN*, she could not justifiably believe that she may ϕ . If sufficient evidence is sufficient for justification, it would follow that Agnes could not have sufficient evidence to believe that she ought to ϕ or that she may ϕ . Is there any credible explanation that helps us understand why it would be impossible for Agnes to acquire this evidence? Would intuitions that otherwise justified her beliefs fail to justify and fail to provide evidence in the special case where they're misleading? Would the word of Agnes' guru carry weight only in the special cases where he's not telling her that she's permitted or required to do something that's wrong?

Someone could try to meet this challenge by saying that the presence of evidence actually plays some role in determining whether Agnes may φ. The idea is that if Agnes' evidence about what matters had been different and she had been aware of the same facts, *then* she would have one set of obligations but Agnes' obligations differ if she's aware of the same facts but has different evaluative evidence.¹⁹ This leads to some rather troubling results. On this view, we have to let agents do whatever their evidence indicates is required (DOWN**) or permitted (DOWN*). With little prior constraint on what their evidence could support, we'd have to sign off on their behavior provided that it was always guided by a well-founded belief.

To our minds, this gets things back to front. Just as we shouldn't say that Agnes is bound by her well-founded beliefs about ought to refrain from doing things others with fewer well-founded beliefs would be permitted to do, Agnes isn't freed from practical requirements that are incumbent upon others by her well-founded beliefs. If things go badly for Agnes and she gets evidence that, say, tells her that she's

¹⁹ See Zimmerman (2010) for a defense of this kind of view, a view according to which what Agnes ought to do depends upon both her evaluative evidence and her empirical evidence.

permitted to steal kidneys from the rich and distribute them to the poor or to pull sleeping drunks into the paths of trolleys, she's still required to act like a non-consequentialist.²⁰

The possibility of this shifty view highlights one of the ways that talk of 'fixedpoints' is potentially misleading. On the model that we prefer, there are fixed facts about what people should do in certain circumstances. Such facts can have little direct bearing on what evidence these subjects might have, so if they have any bearing on what these subjects may believe, it could only be because some highly unorthodox account of justification is correct (e.g., a view on which justified belief is apt belief or knowledge).²¹ Views that allow for 'shifts' in what a subject ought to do in particular circumstances that are induced by differences in evidence can try to incorporate the Partial Fixed-Point Thesis, but they cannot say that all of us, regardless of what our evaluative evidence might be, should/shouldn't act in certain ways in certain situations. Instead, they say that we're under this complex requirement to see to it that there's a mesh or match between our normative beliefs and the thoughts, feelings, or actions that these beliefs are about. We see no good explanation as to how there could be such a requirement since it seems that there's nothing particularly good about exhibiting the pretty pattern required by Enkrasia. We think that it's interesting that Enkrasia enjoys so much popularity and seems only to make sense in a framework that incorporates a very unpopular view of justification (i.e., one that says that justified beliefs are justified only if true). We leave it to the reader to find a vindication of Enkrasia that doesn't make use of this assumption and doesn't succumb to the skeptical pressures discussed above.

5. Conclusion

We've seen now two ways in which justification might function as a guide. If it infallibly identifies a guideline, this isn't because the guidelines are lustrous or luminous. If it does that, it is down to the fact that justification is an externalist notion and that the justification property is one that a belief has only if it is correct. One complaint about this view is that this kind of view conflicts with the idea that justification might be a guide in another sense, a guide that reaches us where we are and tells us where to go next. Justification cannot be valuable by virtue of playing this role because it looks like nothing could both play this role and take you where you actually need to be. If something is lustrous, it will sometimes take you where you shouldn't be. We don't see much value in a guide that we know in advance will lead you astray, not even if it's a guide that's highly visible and easy to pick out in a crowd.

We've also seen something surprising about the idea of justification as a guide. It's obvious that justification cannot play this guiding role if, say, a justified belief is justified only if it is true or constitutes knowledge, but this shouldn't encourage us to think that we could retreat to some sort of internalist view to vindicate this guidance picture. We cannot. The difficulties that arose for UP and its accompanying principles can be generated without leaving the agent's head. The problem generated

²⁰ We're assuming, of course, that we shouldn't pull sleeping drunks into the paths of trolleys to try to save others but appreciate that this is controversial and hope the reader appreciates the irrelevance of that controversy to the point we're making.

²¹ See Littlejohn (2012), Sutton (2007), and Williamson (forthcoming) for defenses of these factive accounts of justified belief.

by Diving isn't one that we can remedy by providing the subject with more knowledge or by eliminating the subject's false beliefs. The problems there arose after we gave the subject all the knowledge they could ever hope to have and without giving them any false beliefs. The problems generated by Feed the Children and Feed the Children (II) could easily be generated for internalist views that take all normative status to supervene upon a subject's mental states.

There has been an increased interest in trying to determine the value of justification. One thought might be that justification matters because of the role that justification plays in our overall normative theory. We think that this is a potentially fruitful idea, but we also think that there are some real pitfalls that we should avoid if we want to develop this idea in the right way. We see the above as a first step towards doing just that.

References

Adler, Jonathan. 2002. Beliefs Own Ethics. MIT University Press.

Bird, Alexander. 2007. Justified Judging. *Philosophy and Phenomenological Research* 74: 81-100.

Brown, Jessica. 2008. Subject-Sensitive Invariantism and the Knowledge Norm for Practical Reasoning. *Nous* 42: 167-89.

Easwaran, Kenny and Branden Fitelson. 2014. Accuracy, Coherence, and Evidence. In T. Gendler and J. Hawthorne (ed.), Oxford Studies in Epistemology.

Fantl, Jeremy and Matthew McGrath. 2009. Knowledge in an Uncertain World. Oxford University Press.

Gerken, Mikkel. 2011. Warrant and Action. Synthese 178: 529-47.

Gibbons, John. 2013. The Norm of Belief. Oxford University Press.

Greco, Dan. 2014. A Puzzle about Epistemic Akrasia. *Philosophical Studies* 167: 201-19.

Harman, Elizabeth. 2011. Does Moral Ignorance Exculpate. Ratio 24: 443-68.

Huemer, Michael. 2007. Moore's Paradox and the Norm of Belief. In S. Nuccetelli and G. Seay (ed.) *Themes from G.E. Moore*. Oxford University Press.

Kiesewetter, Benjamin. Forthcoming. You ought to ϕ only if you may believe that you ought to ϕ . *Philosophical Quarterly*.

Lasonen-Aarnio, Maria. MS. Enkrasia or Evidentialism? Learning to Love Mismatch.

Littlejohn, Clayton. 2012. *Justification and the Truth-Connection*. Cambridge University Press.

Littlejohn, Clayton. 2013. The Unity of Reason. In C. Littlejohn and J. Turri (ed.), *Epistemic Norms*. Oxford University Press.

Littlejohn, Clayton. Forthcoming. Stop Making Sense: On a Puzzle about Rationality. *Philosophy and Phenomenological Research*.

McGlynn, Aidan. 2013. Believing Things Unknown. Nous 47: 385-407.

McKinnon, Rachel. 2015. The Norms of Assertion: Truth, Lies, and Warrant. Palgrave.

Moeller, Emil. 2015. Defeasible Knowledge. Oxford University DPhil Thesis.

Nelkin, Dana. 2000. The Lottery Paradox, Knowledge, and Rationality. *Philosophical Review* 109: 373-409.

Neta, Ram. 2009. Treating Something as a Reason for Action. Nous 43: 684-99.

Ryan, Sharon. 1991. The Preface Paradox. Philosophical Studies 64: 293-307.

- Smith, Martin. 2016. Between Probability and Certainty: What Justifies Belief. Oxford University Press.
- Smithies, Declan. 2012. Moore's Paradox and the Accessibility of Justification. *Philosophy and Phenomenological Research* 85: 273-300.
- Sorensen, Roy. 1995. Unknowable Obligations. Utilitas 7: 247-71.
- Srinivasan, Amia. 2015. Normativity without Cartesian Privilege. *Philosophical Issues* 25: 273-99.
- Srinivasan, Amia. 2015b. Are We Luminous? *Philosophy and Phenomenological Research* 90:294-319.
- Sutton, Jonathan. 2007. Without Justification. MIT University Press.
- Titelbaum, Michael. 2015. Rationality's Fixed-Point (or: In Defense of Right Reason). In T. Gendler and J. Hawthorne (ed.), Oxford Studies in Epistemology 5.
- Jonathan Way and Daniel Whiting. Forthcoming. If you justifiably believe that you ought to φ, you ought to φ. *Philosophical Studies*.
- Williamson, Timothy. 2000. Knowledge and its Limits. Oxford University Press.
- Williamson, Timothy. forthcoming. Justifications, Excuses, and Sceptical Scenarios. In. J. Dutant and F. Dorsch (ed.), *The New Evil Demon Problem*. Oxford University Press.
- Worsnip, Alex. forthcoming. Belief, Credence, and the Preface Paradox. *Australasian Journal of Philosophy*.
- Zimmerman, Michael. 2010. Living with Uncertainty. Cambridge University Press.