3

Subjective Normativity and Action Guidance[1]

ANDREW SEPIELLI

It's often claimed that when we are uncertain about what's the case, we cannot guide our actions by[2] objective norms (e.g. objective utilitarianism) and must instead avail ourselves of subjective norms (e.g. subjective utilitarianism). I agree with this claim, once it is appropriately restricted. But I don't think anyone has satisfactorily explained why it is true.

Equipped with only unsatisfactory explanations, we risk seriously distorting our understanding of the notions implicated in the debate about subjective normativity and action guidance—to wit, the notions of a subject, of normativity, of action, and of guidance.[3] In this paper, then, I want to take the first steps towards an explanation of why subjective norms are uniquely suited to play the action-guiding

[1] Thanks to Adam Arico, Ruth Chang, Kelin Emmett, Benj Hellie, Tom Hurka, Leon Lukic, Christopher Maloney, Elinor Mason, Jennifer Nagel, Devlin Russell, Holly Smith, Mark Timmons, and two anonymous referees for helpful comments and discussions.

[2] It is natural to say, "Ms. X guided her action by such-and-such a norm", and also to say "Ms. X guided her action by such-and-such a belief state". "Guided by . . . " seems clearly to be equivocal. The sense in which I guide my action "by" a mental state is not the sense in which I guide my action "by" the content of a mental state. See Turri 2009. To reinforce this distinction, then, I will say that one guides one's actions *by* the contents of one's beliefs, and *with* the beliefs themselves.

[3] Hewing to the wrong explanation will also imperil our ability to solve more specific problems. For my own part, I'd like to defend a theory of how we ought to behave when we must act under *normative uncertainty*—that is, uncertainty about the reasons, "oughts", requirements, etc. provided by the non-normative facts. (Uncertainty among *moral theories* is a notable type of normative uncertainty.) Any such theory faces a nasty regress problem—"what if you're uncertain about which theory of what to do under normative uncertainty is correct?"—that we'll be unable to solve without having in hand the correct account of the role subjective norms play in the guidance of action. For more on normative uncertainty and the regress problem, see Sepielli 2010.

role they do. In other work, I present my positive proposal in detail.[4] My aim here is simply to lay the groundwork for this positive view by presenting a conceptual framework within which candidate explanations should operate, showing how certain of these explanations fail, and drawing some lessons from their failure.

I'll begin with a characterization of subjective normativity and its relation to objective normativity. Then, I'll clarify and restrict the claim about action guidance the truth of which I seek to explain. After that, I'll argue against four candidate explanations. I'll conclude by providing the beginnings of my own explanation.

OBJECTIVE AND SUBJECTIVE NORMATIVITY

Let us suppose that a pill will kill James if he takes it, but that James reasonably believes that the pill will cure some illness of his. I might say of James that he ought not to take the pill, on the grounds that it will kill him. You might say of James that he ought to take the pill, on the grounds that he reasonably believes that it will cure his illness. It is an assumption of this paper that we may both be right. I am right insofar as I am employing the objective sense of "ought"; you are right insofar as you are employing one of the subjective senses of "ought". Now, parties might disagree about which of these senses is more important, or about the role(s) played by each sense, or about which sense the ordinary, natural-language "ought" is identified with in this context or that. But I'll assume throughout that in employing our respective "oughts", we are not having a straightforward substantive disagreement about what someone in James's situation ought to do. That is, ours is not a disagreement of the sort the utilitarian and the deontologist are having when they argue about whether one ought to push the large man in front of the trolley.

There are different ways of spelling out the thesis that there are multiple "senses" of "ought". One possibility is that they express different OUGHT concepts;[5] there is the objective OUGHT concept,

[4] See Sepielli (work in progress-a).

[5] Following Jerry Fodor, I use words or phrases in all capital letters to refer to concepts. See Fodor 1998. I use words or phrases in quotation marks to refer to words or phrases that express those concepts. I'll shift back and forth between speaking of different senses of "ought"

and at least one subjective OUGHT concept. Another possibility is that, rather than there being separate objective OUGHT and subjective OUGHT concepts, OUGHT concepts contain a *parameter* that is set by context, assessor's information, speaker's intention, or something else.[6] Whether "ought" in some use is objective or subjective depends on how the parameter of the concept it expresses is filled in at the time of expression. I want to remain neutral between these spellings-out. I can say what I want to in this paper without endorsing either of them over the other.

Several people have suggested to me that there's no way of understanding the subjective sense(s) of "ought" in terms of any objective normative notion(s). No doubt this skepticism has been engendered by the history of failed attempts at doing so. Nonetheless, I'd like to make an attempt of my own. I'll begin with a rough "first pass" proposal; then I'll refine it into the view I wish to defend.

We should allow that the objective OUGHT may depend for its proper application on any feature of the world whatsoever. Some authors say that the objective OUGHT's proper application may depend on any feature of the world *except* the agent's beliefs, or the evidence available to the agent, presumably because these are the things upon which the proper application of the subjective OUGHT(s) depend. But this is a mistake. First, there are cases where what I objectively ought to do seems to depend on my beliefs. I ought not to say "P" if it would be a lie, but whether it is a lie depends on whether I believe that P. Second, as a methodological matter, we will want to leave open as a conceptual possibility that what I objectively ought to do may depend on anything whatsoever.[7]

Subjective OUGHTs depend for their proper application on the agent's beliefs or degrees of belief, or on the agent's evidence, or on the probabilities *regarding the features of the world upon which the proper application of the objective OUGHT depends*. So if it turns out that the utilitarian is right, and what one objectively ought to do depends on how much utility each of one's candidate actions would produce,

(referring to the word), and different senses of OUGHT (referring to the concept or family of concepts).

[6] See Dowell (work in progress); and Kolodny and MacFarlane (work in progress).

[7] On these points, see also Smith 2010.

then what one subjectively ought to do will depend on one's beliefs or degrees of belief about how much utility each action will produce, or on the evidence or the probabilities regarding the same.

There are different subjective OUGHTs because beliefs/degrees of belief, evidence, and various sorts of probability are all different things, and there is a subjective OUGHT that depends for its proper application on each. Following Derek Parfit, we may speak of the *belief-relative* sense of "ought", the *reasonable-belief-relative* sense, the *degree-of-belief-relative* (or *credence-relative*, or *subjective-probability-relative*) sense, the *evidence-relative* sense, and the *objective-probability-relative* sense, each of which depends for its proper application on the feature mentioned in its label.[8] We could ramify even further. There are, for example, different "interpretations" of objective probability—the long-run frequency interpretation, the propensity interpretation, the logical interpretation, etc.—and there could be an OUGHT corresponding to each interpretation.

Finally, there is a subjective OUGHT that I call the *minimal-probability-relative* OUGHT. As far as I know, nobody in moral philosophy has discussed this sense before. This is unfortunate, for it is, I believe, the sense most crucial for action guidance. What's more, as we'll see later, we'll be tempted into confusion when we ask other subjective senses—specifically, the credence-relative and evidence-relative senses—to play the role most naturally occupied by the minimal-probability-relative sense. Let me take a moment to explain what minimal probability is, and the category of norm that is relative to it.

One *reports* one's belief that P by saying, "I believe that P". One *expresses* one's belief that P by saying, simply, "P". Just as one can report a belief, one can report a degree of belief, or credence. One does this by saying, e.g., "I have a credence of .6 that P", or, less precisely, "I'm slightly more confident than not that P". But how does one *express* a credence? Not with a sentence like the ones just mentioned; that would elide the distinction between reporting and expressing. And not by saying, "There's a .6 objective probability that P". That would be expressing a full belief about objective probabilities. And not by saying, "The evidence supports a credence of .6 that P", or "There's evidence of

[8] Parfit 2011.

'degree .6' that P". These would be expressions of full beliefs about evidence.

I will call the statements we use to express, rather than report, our degrees of belief "minimal probability" statements. Those with "minimal probability" in their vocabulary might express a .6 credence that P by saying, "There's a .6 minimal probability that P". But of course this term is a neologism. I suspect, however, that ordinary people express the same thing with statements like: "Frank's probably already at the party"; or "Five-to-one odds Koopmans ain't even gonna show". I also suspect that most uses of "I'm not sure the bank is open on Saturdays" and the like are non-literal; someone who says this is typically not reporting her credence that the bank is open on Saturdays (in the same way she might report her friend's credence by saying, "My friend is not sure the bank is open on Saturdays"). Rather, she is expressing her credence in what, to us sticklers, is a grammatically misleading way. All of these quotidian statements are ways of expressing the same thing we might more precisely express using the language of minimal probabilities.

So what *are* minimal probabilities? Of particular concern is whether they are features of the world to which we commit ourselves whenever we express our credences, in the way that we commit ourselves to objective probabilities when we make statements about objective probabilities. That'd be strange. If I can express a full belief that there are ostriches without committing myself to any non-ostrich entities or properties, I should be able to express a credence that there are ostriches without committing myself to any non-ostrich entities or properties.

We can avoid the strange result by assigning semantic values to minimal probability statements, in part, *expressivistically*. Rather than explaining the semantic features of the "probably" in statements like "Frank's probably already at the party" by appeal to minimal probability properties in the world, we explain their semantic features by the mental states they're used to express—namely, credences. So suppose we want to explain why "There's a .8 minimal probability that P" and "There's a .8 minimal probability that ~P" are inconsistent. We would

do that by saying that the former expresses a credence of .8 that P, the latter expresses a .8 credence that ~P, and these are credences that it's not rational to hold together.[9]

There is a sense of OUGHT that is minimal-probability-relative. It is the sense one typically expresses when one says to oneself, "Hmmm, it's probably gonna rain. If it's probably gonna rain, I ought to bring an umbrella. So yeah, I'll bring an umbrella." In other words, "minimal-probability-relative" is just an unfamiliar label for a sense of OUGHT that, I contend, couldn't be more familiar or primordial for agents like us. We'll see later on that understanding this primordial sense of OUGHT is very helpful theoretically.

It will be helpful to draw one last distinction: between "oughts" relative to different *subsets* of beliefs or probabilities. Sticking with belief-relative normativity for a moment, we may distinguish between an "ought" that is belief-relative only with regard to one's *non-normative beliefs*; one that is belief-relative only with regard to one's *normative beliefs*; and one that is belief-relative with regard to all of one's beliefs.[10] Moral theories like subjective utilitarianism are belief-relative in the first sense. What one ought to do according to subjective utilitarianism is independent of one's beliefs about utilitarianism itself and its competitors. Theories of "what to do when you don't know what to do" (i.e. under uncertainty among normative theories), like those offered by Ted Lockhart,[11] Jacob Ross,[12] and myself,[13] are belief-relative in the second sense. What one ought to do according to such theories *does* depend on one's degrees of belief in utilitarianism, contractualism, and all the rest.

[9] My expressivist treatment of minimal probability is similar to Seth Yalcin's expressivist treatment of epistemic probability. See, e.g., Yalcin 2011. I explain the differences in my (work in progress-c).

[10] This is just one way of carving one's set of beliefs "at the joints". We might, for example, also have an "ought" that is relative to one's beliefs about *a priori* matters only, and another that is relative to one's beliefs about *a posteriori* matters only. We could also carve up belief-sets less naturally—with an "ought" that is belief-relative only with regard to one's beliefs about rodents, say.

[11] Lockhart 2000.

[12] Ross 2006.

[13] Sepielli 2009.

SUBJECTIVE NORMATIVITY AND TRYING

As I noted in the last section, this "first pass" characterization of the relationship between objective and subjective normativity isn't exactly right. Again, the proposal has been: subjective OUGHTs depend for their proper application on the (degrees of) belief/evidence/probabilities regarding the features upon which the proper application of the objective OUGHT depends.

This characterization fails because it is inadequately specific. For while the proper application of the subjective OUGHT does indeed depend upon these beliefs, probabilities, and so on, so does the proper application of the subjective FORBIDDEN, the subjective OUGHT *NOT*, the subjective LESS REASON TO . . . THAN TO . . . , and so forth. They depend in different ways, and we want to be able to say which way corresponds to the subjective OUGHT. We want to define the subjective OUGHT in such a manner that our definition picks out that very subjective normative notion, rather than any of the others.

We don't want to get *too* specific though. For example, we can't define what I subjectively ought to do as whatever maximizes expected objective value (hereafter: "EOV"), relative to some probability distribution or other.[14] I myself think it's very plausible that this *is* what we subjectively ought to do,[15] but it's a bad definition, for it rules out by fiat all other theories thereof. Notably, it rules out alternative views about how to respond to uncertainty or risk-maximin, risk-aversion or risk-seeking, the view that I subjectively ought to do what's most likely what I objectively ought to do, etc.

Instead, I offer as a basic proposal:

What I subjectively ought to do =*df.* the *best try* for me at doing what's objectively valuable.

This will require some unpacking. We are all acquainted with the concept TRYING; as agents of limited abilities, our most primordial use of it is in action. In acting under the description "trying to do A",

[14] By "value", here, I simply mean "strength of reasons", however those reasons are determined. So considerations other than, say, the consequences of actions could determine actions' "values", as I'm using the term.

[15] For arguments, see Sepielli 2010, chs. 2–3.

I aim at doing A just as I do when I act under the description "doing A".[16] As Benj Hellie puts it, both trying to do A and doing A involve a "strand of agency working towards having done" A.[17] What it is to aim at doing A, or for there to be such a strand of agency, is difficult to explicate, but for present purposes I want to assume the reader's understanding of this "highest common factor" of trying and doing.

The difference between trying and doing is essentially this: When I act under the description "trying to do A", I am consciously uncertain that I will succeed in doing A; but when I act under the simple description "doing A", I am not consciously uncertain that I will succeed in doing A. I can simply do A without doing something *else* that I conceive of as an attempt or a try. The distinction between conscious uncertainty and uncertainty that is not conscious will be important later on, so it's worth getting clearer on:

Conscious uncertainty is a conscious representational attitude that, necessarily, has a phenomenal character that includes a feeling of unsurety (or as we'll call it later, a "multidirectional" phenomenology). (By contrast, *conscious certainty* is a conscious representational attitude that, necessarily, has a phenomenal character that includes a feeling of surety (or as we'll call it later, a "unidirectional" phenomenology.)

Non-conscious uncertainty is a non-conscious representational attitude that, necessarily, has a functional profile characteristic of uncertainty. (By contrast, *non-conscious certainty* is a non-conscious representational attitude that, necessarily, has a functional profile characteristic of certainty.)[18]

I leave it open what the functional profile of non-conscious uncertainty is. Perhaps it is a set of dispositions to act or reason in certain ways; perhaps it is a disposition to feel *consciously* uncertain in response to some cue; perhaps it is something else. I take it that most of us are non-consciously uncertain about *most* propositions, insofar as we would not admit upon questioning to a credence of 1 in them, or would not, say, stake our lives on them for a payout of a single chocolate chip.

[16] On the notion of action under a description, see Anscombe 1957, section 11; as well as Anscombe 1979.

[17] Hellie (forthcoming).

[18] I am using "conscious", then, to mean what Ned Block has called "phenomenally conscious" rather than what he has called "access conscious". See Block 1995.

It would be wrong, then, to say that one who aims to do A, but is uncertain whether he will succeed, must consign himself to merely trying to do A. For one may do A under non-conscious uncertainty without employing the TRYING concept, so long as he is not consciously uncertain. There are two ways to be non-consciously uncertain of a proposition without being consciously uncertain of it. One way is to have no conscious doxastic attitude at all regarding it. We sometimes call this "taking something for granted". Another way is to be consciously certain about it. Both of these situations are common. I am non-consciously uncertain whether the door knob will stay on the door when I turn it, and yet I take for granted that it will. I am non-consciously uncertain that the morning bus will leave at 6:55, and yet in running through the day's plans in my head, I am consciously certain that it will.[19]

I may do without trying if either I take for granted my success or I am consciously certain of it. For example, I can simply act under the description "saying my mother's name". I do not have to conceive of myself as merely saying the name that is most likely my mother's, and then praying that I am correct. Unless we see this crucial point, we will end up with a psychologically unrealistic picture of human action—one on which we see ourselves as "rolling the dice" with our every movement.[20]

Now on to the second part of the formulation: "... at doing what's objectively valuable". It is very important that we state things this way rather than saying "... at doing what's objectively best". For the feature of being the best is a *non-scalar* feature, and so the best try at doing what has this feature is simply the action with the greatest chance of having the feature. But it is highly implausible that one subjectively ought to do the action that is mostly likely to be objectively best. Suppose action

[19] On the effects of what I'm calling "deliberative context" on conscious (un)certainty, see Gollwitzer 1990; and Kruglanski and Webster 1996. For a helpful philosophical discussion of these psychological phenomena, see Holton 2009. Thanks to Jennifer Nagel for showing me around the psychological literature.

[20] So I'm rejecting the old-fashioned picture on which (a) there are a small class of actions that are "basic" for every agent in every situation, and (b) we do everything else *by* doing one of these basic actions. My preferred picture accords with the psychological literature cited in fn. 18 *supra*, and with Candlish 1984; John McDowell's discussion of Chuck Knoblauch in McDowell 2009; and Hellie (forthcoming).

A has a .51 chance of having a value of 100 and a .49 chance of having a value of −1,000,000, and action B is certain to have a value of 99. A is more likely than B to be objectively best, but surely B is what one subjectively ought to do. At the very least, we will not want to enshrine it as a definitional truth that one subjectively ought to do what is most likely to be objectively best.

The feature of being valuable, though, is a *scalar* one. It comes in degrees. And the best try at doing what has a scalar feature needn't simply be the action that is most likely to have the highest degree of the feature. (For example, the best try at making money on the stock market is not simply to throw all of one's initial investment into high-risk, high-return penny stocks.) Rather, it need only be an action that maximizes some function of the degrees of the scalar feature and the action's probability of having each of those degrees, respectively. It is an open, substantive question exactly what such a function should look like. But on what seem like the only plausible answers, action B will come out as the best try at doing what's objectively valuable in our example.

A definition of the subjective OUGHT in terms of trying and some scalar objective normative feature seems, then, to be the best way to avoid both underspecificity (i.e. a definition that fails to pick out the subjective OUGHT uniquely) and overspecificity (i.e. a definition that rules out what should be live theories of what one subjectively ought to do).

Now, we noted earlier that there are different senses of OUGHT that we might call "subjective"—the belief-relative sense, the minimal-probability-relative sense, and so forth. The present account of subjective OUGHTs in terms of TRYING can capture this, because there are, correspondingly, different senses of TRY, GOOD TRY, and so on. There are senses of TRY that are relative to our beliefs, and senses that are relative to each of the different types of probability, respectively.

As saw earlier, we use TRY in its most primordial sense when we ourselves act under "trying . . . " descriptions—that is, when we are consciously uncertain that we will succeed in our aims. This is the sense of TRY we can use to define whichever kind of subjective OUGHT that is most fundamental to action guidance under uncertainty. This sense of TRY is, I shall claim, relative to minimal probabilities, and so we may use it to define the minimal-probability-relative OUGHT:

What I minimal-probability-relative ought to do=*df.* the best minimal-probability-relative try for me at doing what's objectively valuable.

But we also use the language of "trying" in a more third-personal way, to describe the behavior of others (or of our past or future selves) who may not be consciously uncertain of their success, but whose probability of success is, on some interpretation of probability, less than 1. Suppose James is consciously certain that he will succeed at some ball-involving carnival game. Then he will not, in playing the game, token the concept TRY. But unbeknownst to him, the game is actually very difficult. Then we might speak of him as trying—and perhaps of his try as "good" or "poor"—despite his own conscious attitudes regarding his success.

For suppose that the publicly available evidence suggests that, say, bouncing the ball off of the side wall is the surest strategy for winning the game. Then we might rightly say that James's bouncing the ball off of the side wall is a good try. But suppose that James believes that the ball is completely inelastic, and correspondingly believes that the surest strategy is, say, throwing the ball straight at the hole. If he expresses this credence to us, we might then say that, given his credence, throwing the ball straight at the hole is a good try. A good try, in the former sense, at doing what is objectively valuable is what one ought in the evidence-relative sense to do. A good try, in the latter sense, at doing is what is objectively valuable is what one ought in the credence-relative sense to do.

THE EXPLANANDUM, IN LIGHT OF THE FOREGOING

This understanding of subjective normativity will allow us to state our explanandum more precisely.

For consider action under *certainty*. Suppose a millionaire tells me that he will donate $1000 to a cause of my choosing. If I am certain that I have most (objective) reason to choose Oxfam, then I may simply guide my choosing Oxfam with this belief. Not only needn't I employ a subjective norm; it would be downright *weird* to. Imagine asking oneself, "What subjectively ought I to do given that it's certainly the case that I objectively ought to do A?"

Now consider an otherwise similar description of action under *uncertainty*. I have some credence that I have very strong (objective) reasons

to choose Oxfam, slightly weaker reasons to choose Greenpeace, and substantially weaker reasons to choose PlanUSA; I also have some credence that I have very weak reasons to choose Oxfam, stronger reasons to choose Greenpeace, and substantially stronger reasons to choose PlanUSA. In other words, I have a credence distribution over "objective reason"-propositions. I cannot move straight from this uncertainty to an action or to an intention to act such that my action would count as guided with this credence distribution. If I wish to guide my behavior with a doxastic attitude toward a normative proposition, I will have to form a conclusion about what I subjectively ought to do when confronted with these branching possibilities.

When objective normative propositions are placed in the "content box" of a full belief or certainty, we have an actionable belief state, but when they are placed in the content boxes of intermediate credences, we end up with this set of belief states that is impotent in guiding action, and need to avail ourselves of subjective norms—norms about "what to do when you don't know what to do". This is what I seek to explain.

Now, I said at the outset that this explanandum is true only if it is "appropriately restricted". The restriction I have in mind is that the uncertainty in question must be conscious. For I do not need to employ subjective norms when I am non-consciously uncertain, so long as, at the conscious level, I either take the relevant objective normative proposition(s) for granted or am consciously certain of them. I need only resort to subjective norms when I am consciously uncertain of them.[21]

[21] A referee suggested that since a phenomenal zombie would "need action guidance under uncertainty", but would have no use for a subjective OUGHT defined in terms of phenomenology, the subjective OUGHT I am discussing cannot be the one relevant to the guidance of action under uncertainty. I am unsure why phenomenal zombies call for a different treatment than those who I labeled as "taking for granted" the propositional contents of their non-conscious uncertainty, since both types of agents combine non-conscious uncertainty in these propositions with the lack of *any* conscious doxastic attitude towards them. In response to the referee's example, though: Consider a case where a phenomenal zombie's non-conscious uncertainty regarding objective normative propositions causes a bit of behavior. In keeping with what I said about cases of "taking for granted", I would want to say that this zombie guided her behavior with this uncertainty. Denying this, it seems to me, would commit us to denying the possibility of action guidance except in the very rare instance where an agent is utterly certain, in this non-conscious, dispositional sense, of some objective or subjective norm. (We saw earlier why this would be rare, and I discuss it again in the conclusion.) The latter denial strikes me as utterly implausible. However, if we do wish to say that the zombie cannot guide her behavior by her uncertainty, but could for some reason guide her behavior by a state of non-conscious certainty in a subjective norm, then I would question the propriety of the

This follows from our earlier analysis of "what I subjectively ought to do" as "the best try for me at acting in accordance with objective normativity", along with our distinction between "trying to do A" and "doing A" in terms of *conscious* uncertainty specifically. We said that I can act under the description "doing A" (rather than "trying to do A") so long as I am not consciously uncertain that I will succeed in doing A (even if I am non-consciously uncertain that I will succeed). So it follows that I can act under the description "doing what I objectively ought to do" even if I am non-consciously uncertain that I will succeed. I'll need to act under the description "trying to do A" only if I am consciously uncertain that I will succeed in doing A. So it follows that I will need to employ a subjective norm only if I am consciously uncertain about the objective normative status of my action.

the "metacognitive signal" explanation

With our conceptual framework on the table, and a clearer idea of what we seek to explain, let's have a look at some candidate explanations.

One draws inspiration from some recent work in empirical psychology, and the uptake of this work by a few philosophers. We might interpret the characteristic feeling of conscious uncertainty as a sort of "metacognitive signal"—a feeling that delivers information about the functioning of our own cognitive processes.[22] On this explanation, the feeling of unsurety is plausibly construed as a warning that, if we simply act on our uncertainty without forming a separate belief about *how* to act under that uncertainty, we are likely to act in a way that is improperly responsive to that uncertainty—that is irrational. We have

locution "*need* action guidance". For I would wonder what, on this conception of guidance, was the import of the contrast between guided and unguided action such that the former could be said to be "needed".

[22] See Alter et al. 2007. Peter Railton stresses the role of such signals in "shaping and guiding . . . coordinate suites of thought and action". See Railton 2009: 106. Railton discusses the example of a driver, Christine, whose fluent driving is interrupted when she notices an older driver ahead puttering along and decides to slow down to avoid startling him. He says, "In Christine, the negative affect generated by her . . . simulation of the other driver's situation tend[s] to inhibit her current course of action (blasting ahead), refocus attention, and prompt thoughts of less aggressive alternatives." In recent work, Jennifer Nagel emphasizes the role of the "feeling of knowing" some fact in prompting attempts to remember that fact. See Nagel 2010.

difficulty transitioning from a credence distribution over objective normative propositions to an action because, were we to do so, we would be acting in the face of an indication that such a transition is likely irrational.

I don't wish to deny that the feeling of unsurety may serve as this sort of signal. But I deny that it is merely such a signal, and that its status as a signal is underivative of its other roles. For one thing, an unadorned "metacognitive signal" hypothesis leaves out half the phenomenology of conscious uncertainty. When I am consciously uncertain, it is not as though I see some action as the thing to do, but something in the back of my mind tells me that I have erred in so seeing it. Instead, there is a feeling of being unsettled, of not having enough to go on, of the contents of my divided credences failing to direct me towards doing one candidate action rather than the others. The phenomenology is not simply one of something holding me back; it is also one of nothing pushing me forward. Moreover, the feeling of being held back seems derivative of the feeling of not being pushed forward. Candidate actions strike me as risky *because* undirected. And this is something that the bare "metacognitive signal" hypothesis fails to capture.

Furthermore, and relatedly, the metacognitive signaling answer fails to explain why guiding one's actions with states of conscious uncertainty is *impossible* as opposed to merely unlikely. For we may deliberate and act—we may "push on"—despite the presence of inhibitory metacognitive signals. This happens often. Sometimes we ignore metacognitive signals because we deem them inaccurate. Suppose, for example, that you conclude that the answer to an exam question is "William Lyon MacKenzie King", but then immediately have the feeling of having misremembered. If you were to discover that you had been slipped a drug, the effect of which is to induce this feeling of misremembering, it seems that you could simply and safely ignore this feeling, and guide your writing this answer by your belief that it is correct. Or we may ignore metacognitive signals because we think the extra cognitive effort needed to silence them would be too costly. So if the metacognitive signaling hypothesis is true, it would be possible for us to guide our actions with conscious uncertainty. It might feel uncomfortable, but there would be no *in principle* bar to our doing it. But as I will suggest later on, guiding our actions with conscious uncertainty is impossible, for such uncertainty, as its phenomenology suggests, simply doesn't

direct us toward a particular action to the exclusion of others. Suppose, for example, that you are consciously utterly uncertain what the correct answer to an exam question is. Learning that this feeling of unsurety was drug-induced wouldn't bring you any closer to writing an answer.

For all the "metacogntive signal" answer tells us, the phenomenology of conscious uncertainty could be anything. After all, the form a signal takes is irrelevant to its role as a signal, once we see what it's indeed a signal for. I could rig up my car so the gas gauge reads "empty" when the tank is full and "full" when the tank is empty; so long as I remembered that I did this, the ability of my gas gauge to serve a signal of my tank's fullness would be undisturbed. Similarly, a mere signal that I'm about to do something irrational could take the form of a painful sensation, a fearful sensation, a bubble-gummy taste in my mouth—anything. But the explanation for why I can't guide action with conscious uncertainty about what to do seems to inhere in that state's very particular phenomenology. As I suggest later and argue elsewhere,[23] this particular phenomenology is what *makes it the case* that I can't guide action with the state.

THE "INSIDE OF AN AGENT" EXPLANATION

A more widely endorsed explanation goes like this: We cannot guide our behavior by a state of uncertainty among objective normative propositions because such propositions advert to features that are in some sense outside the agent's grasp. We must instead guide our behavior by norms that advert to features *internal to the agent*, or at least *internal to the agent's ken*.

Frank Jackson suggests this explanation in a well-known discussion of objective consequentialism:

> the fact that a course of action would have the best results is not in itself a guide to action, for a guide to action must in some appropriate sense be present to the agent's mind. We need, if you like, a story from the inside of an agent . . . and having the best consequences is a story from the outside.[24]

[23] See Sepielli (work in progress-a).

[24] Jackson 1991: 466–7. Elinor Mason suggested to me that there are ways to interpret Jackson other than the way I do in this section. This is a fair point, but ultimately I'm more

"Inside of an agent" explanations should be rejected. To see why, first imagine that I am certain what will have the best results. If I am a consequentialist, I will think to myself, "Doing A will have the best results. If doing A will have the best results, then I (objectively) ought to do A. So I (objectively) ought to do A." I may then guide my doing A by this conclusion. There is no need to employ a norm that adverts to one of my mental states, or to the evidence/information to which I have access.

But now note that "the fact that a course of action would have the best results" is no more "present to the agent's mind" when the agent is certain than it is when the agent is uncertain. Certainty is no guarantee that my beliefs about the world match the way the world actually is. Still it seems that, so long as we are certain about the relevant propositions, we may guide our conduct by norms that advert to mind-independent features of the world. We don't need a "story from the inside of an agent" in that case. So if we cannot guide our actions under conscious *un*certainty by objective norms, it cannot be because of some general fact that we can only guide our actions by norms that advert to features of our own minds. It's an utterly banal fact that the *vehicles* of our reasoning are our own mental states. But this does not imply that the *contents* of those vehicles must make reference to our mental states.

Unless we see that this explanation fails, we will be tempted into distorted pictures of human agency under uncertainty about what to do. Any view on which we must have access to the truth of P to use P as a premise in our practical reasoning is such a picture, and such views are widespread. But while there is a real problem of action guidance under uncertainty *qua* uncertainty, there is absolutely no problem of action guidance under ignorance *qua* ignorance. Action guidance is imperiled not by a bad connection between mind and world, but rather by a divided mind.

WHY THIS EXPLANATION MAY BE TEMPTING (AND HOW TO RESIST THE TEMPTATION)

So why have so many been tempted by this faulty explanation? I want to suggest one source of temptation both because I suspect it's very

concerned with whether this sort of explanation is "in the air" than with whether Jackson in particular endorses it. For another seeming endorsement of this explanation, see Gibbard 2005.

widespread, and because it can be cured using a piece of apparatus I introduced earlier. My speculation is that we first arrive, *via* an argument-by-elimination, at views about the sorts of norms we *must* use to guide our actions under uncertainty, then we *read off* from the contents of these norms the conclusion that we may only guide our actions under uncertainty by norms that advert to the contents of our own mental states or to features of the world to which we have access.

For suppose you are a utilitarian who is consciously uncertain about how much utility various actions will produce. Then you cannot guide your behavior by the objective norm: "One (objectively) ought to maximize (actual) utility". So which options are left? How about, "One ought (in the objective-probability-relative sense) to maximize the quantity Σi p(Si) x u(A given Si), where p(Si) is the objective probability of state of affairs Si and u(A given Si) is the utility produced by doing A, if Si obtains"? This seems unsatisfactory. Why? For reasons of ontological parsimony, you might not believe in such things as objective probabilities. Or you might not believe there are any objective probabilities other than zero and 1, in which case you will no more be able to guide your actions by this norm than by the objective norm. (After all, you will be just as uncertain about which state of affairs *has a probability 1 of obtaining* as you would be about which state of affairs *obtains*.) Even if you do believe in intermediate objective probabilities, you might be uncertain what they are, and consequently, uncertain whether A, or B, or C maximizes quantities like the aforementioned. In that case, you would presumably need another objective-probability-relative norm to guide your action under *this* uncertainty, another to guide your action under uncertainty regarding *that* norm, and so on. Finally, the objective probabilities you believe there are might not correspond to your credences, in which case you will be unable to guide your actions under uncertainty by norms that advert to objective probabilities. I am not imagining that, e.g., you believe there is an objective probability of .4 that P, but have a credence of .3 that P anyway. I'm thinking instead of cases involving disjunctive objective probabilities. Suppose a factory produces cubical boxes with side lengths in the interval 0 m < x $\leq$1 m. Let us suppose that if the next box produced has a surface area of >3 sq m, then a six-sided die will be tossed, and if it has a surface area of <3 sq m, then a twenty-sided die will be tossed. What is the objective probability of the next die roll yielding a "4"?

There are reasons for thinking the answer is: "either 1/6, or 1/20, and that's all we can say about it".[25] But one's credence cannot be: either 1/6 or 1/20, with no more to say. Rather, it may be some "fuzzy" credence, representable *via* an interval or a family of functions. This is the sort of case where objective probability and credence can come apart.

Given that objective norms and objective-probability-relative norms won't always help us guide our actions, it can seem that credence-relative or evidence-relative norms are the only other options. These norms advert to features that differ from objective probabilities in all of the crucial respects, namely: (1) there obviously are intermediate credences or degrees of evidence, (2) it's less plausible that one would be uncertain about one's own credences or the evidence to which one has access, as one might be uncertain about objective probabilities,[26] and (3) one's credences may not, of course, diverge from *themselves*, and one's accessible evidence will tend to line up with one's credences. The credence-relative version of utilitarianism, for example, says that one ought to maximize *expected utility*—or in other words, that one ought to maximize the quantity Σi p(Si) x u(A given Si), where p(Si) is the *subjective* probability of state of affairs Si (i.e. the agent's credence that Si will obtain).

But now that we think we must guide our actions under uncertainty by credence- or evidence-relative norms, it is only natural to think, "Well, obviously this is because of the one essential feature of credence-(evidence-) relative norms—that they advert to the agent's credences (evidence)!" And now it's natural to think, "Well, what separates credences or evidence, on the one hand, from objective probabilities and actual utility produced, on the other? Clearly it's that credences and evidence are inside the agent or accessible to the agent, and these other things are outside the agent! So the lesson to be drawn is that we must guide our behavior by norms that advert to features that are inside the agent, not outside."

This argument-by-elimination assumes that norms that advert to "inside of an agent('s ken)" features like beliefs and accessible evidence are our only alternatives to objective norms and objective-probability-relative norms. But we know from earlier that this is a

[25] See Hajek 2010, section on "Bertrand's Paradox".
[26] But see Smith 2010.

mistake. For it overlooks an important space on the conceptual chessboard—the one occupied by minimal-probability-relative norms. There are minimal probabilities other than zero and 1, just as there are credences other than zero and 1. Indeed, because "There's a minimal probability of X that P" stands to "I have a credence of X that P" in just the way "P" stands to "I believe that P" (the first statement of each pair expresses the mental state that the second statement of each pair reports), I cannot coherently hold that I have intermediate credences without also holding that there are intermediate minimal probabilities.

I suggest a picture on which minimal-probability-relative norms are the ones by which we *fundamentally* guide our actions, and that guidance by, say, credence-relative norms is more derivative and less primordial. For guidance by minimal-probability-relative norms is guidance by the informational content of one's credences about the situation;[27] guidance by credence-relative norms is, technically, guidance by the informational content of one's *beliefs about those credences*. But of course the general usability of the latter content in reasoning piggybacks entirely on the general usability of the former content. It's only if I can guide my behavior by *what I think* that I can guide my behavior (in a more self-alienated way) by *the fact that I think it*.

My preferred picture puts action guidance under uncertainty on all fours with action guidance under certainty. Under certainty: I express my belief that the plane will depart at 4 PM by saying, "The plane will depart at 4 PM." And when I consciously consider what to do, I say the following to myself (in "mentalese"): "The plane will depart at 4 PM. If the plane will depart at 4 PM, then I ought (objectively) to be at the airport at 3 PM. So I'll be at the airport at 3PM." And now under uncertainty: I express my reasonably high credence that the plane will depart at 4PM by saying, "The plane will probably depart [read: it has a high *minimal probability* of departing] at 4 PM". And when I consciously consider what to do, I say the following to myself (again, in "mentalese"): "The plane will probably depart at 4 PM. If the plane will probably depart at 4 PM, then I ought (in the minimal probability-relative sense) to be at the airport at 3 PM. So I'll

[27] By the "informational content" of a mental state, I simply mean the proposition stood for by an expression of that mental state.

be at the airport at 3 PM." In both cases, the premises in my practical reasoning are the mentalese expressions of the relevant mental states.

By contrast, the proposal that the fundamental form of action guidance under uncertainty is by credence- or evidence-relative norms yields a sharp break between reasoning under certainty and reasoning under uncertainty. It implies that, when I'm certain about, e.g., whether giving the patient a pill will kill him or cure him, I can employ norms that advert to the effect on the patient; but when I am uncertain about the same, I can only employ norms that advert to internal features like my own beliefs or the evidence that is presently available to me.[28] But why should the move from certainty to uncertainty occasion this change of focus from the outer to the inner?

"Inside of an agent" explanations of our explanandum lose their appeal once we reject this unnatural shoehorning of belief- and evidence-relative norms into the role that's played, most fundamentally, by minimal-probability-relative norms. For once we cast minimal-probability-relative norms in this role, we see that action under uncertainty is not typically guided by norms that advert to features "inside [] an agent".

THE EXPLANATION FROM "JACKSON CASES"

Another well-known explanation for why we cannot guide our behavior under uncertainty by objective norms is that these norms cannot deliver the right results in cases like those Jackson presents in his "Decision Theoretic Consequentialism and the Nearest-and-Dearest Objection".

Jackson takes as a jumping-off point this remark of Peter Railton's: "... objective consequentialism sets a definite and distinctive criterion of right action ... and it becomes an empirical question which modes of decision-making should be employed and when".[29] As Jackson characterizes Railton's idea: "... the moral decision problem should be approached by setting oneself the goal of doing what is objectively

[28] Of course, one way of discovering whether I believe P is by asking myself whether P. See Evans 1982, ch. 7. But this is not the *only* way to discover what I believe. There are third-personal ways as well—the same ways I'd use to discover what someone *else* believes—brain scans, observations of behavior, etc. And these ways seem offensively beside the point when a patient's life is on the line!

[29] Railton 1984: 117.

right—the action that has in fact the best consequences—and then performing the action which the empirical evidence suggests is most likely to have this property".[30]

Jackson argues that Railton's approach "gives the wrong answers" in cases like this one:

> Jill is a physician who has to decide on the correct treatment for her patient, John, who has a minor but nor trivial skin complaint. She has three drugs to choose from: drug A, drug B, and drug C. Careful consideration of the literature has led her to the following opinions. Drug A is very likely to relieve the condition but will not completely cure it. One of the drugs B and C will completely cure the skin condition; the other though will kill the patient, and there is no way that she can tell which of the two is the perfect cure and which the killer drug.[31]

And this one:

> Jill has only two drugs, drug X and drug Y, at her disposal which have any chance of effecting a cure. Drug X has a 90% chance of curing the patient but also has a 10% chance of killing him; drug Y has a 50% chance of curing the patient but has no bad side effects.[32]

In the first case, Railton's proposal (as Jackson construes it) will counsel Jill to choose either drug B or drug C, since drug A has *no chance* of being what she objectively ought to do. In the second case, this proposal will counsel Jill to choose drug X, since that drug has a *higher* chance of being what she objectively ought to do than drug Y has.

But as Jackson points out, this seems like the wrong advice in both cases. A *morally conscientious agent* would choose drug A in the first case and drug Y in the second case. Therefore, Jackson concludes, the proposal that we guide our behavior under uncertainty by objective norms is flawed. We need to "go subjective" instead.

One might respond that I've changed the subject a bit in considering this point of Jackson's. We had been considering a *non-normative, purely psychological* problem for the proposal that one guide one's behavior under uncertainty by objective norms—namely, that it's impossible. But now it seems that we are considering a *normative*

[30] Jackson 1991: 467. [31] Ibid.: 462–3. [32] Ibid.: 467.

problem for this proposal—not that we *can't* guide our behavior under uncertainty by objective norms, but that doing so will lead us astray. And aren't these different problems? Isn't bad guidance still guidance?

Perhaps this response is right, and if so, read this part of the paper as an evaluation of an argument that subjective norms are essential for *morally conscientious* guided action under uncertainty, not for guided action under uncertainty, period. Such an argument, if successful, would still secure an important role for subjective normativity.

But I think this response a bit too quick, for the "psychological" and "normative" issues here are not so easily severable. The notion of action guidance is in part a normative notion. Here's why: As I suggest later and argue in other work, I guide my behavior with a conscious mental state only if that mental state renders my behavior intelligible from my perspective.[33] Rendering an action intelligible is a way of causally explaining it, but a mental state can causally explain an action without rendering it intelligible. Suppose, for example, that I am sure I ought to vote for Warren rather than for Brown, and that this state causes me in some aberrant case to vote for Brown. It is still not the case that this belief renders my vote intelligible from my perspective. Rather, a mental state can render something intelligible *only if* it *rationalizes* it—in other words, only if the behavior is right in the belief- or credence-relative sense, relative to that mental state.[34] So if, as in Jackson's cases, the action that is most likely to be objectively right given my credences is not *also* credence-relative right relative to these credences, then it cannot count as guided with these credences. So the normative problem Jackson alleges would, unless surmounted, end up imperiling the very guidedness of the action that in these cases of Jackson's is most likely to be objectively right.

But I am not convinced that Jackson has really scored a point against objective normativity or for subjective normativity. That's because it is

[33] Sepielli (work in progress-c).

[34] You may wonder whether guidance of an action, A, with a mental state, S, requires that A be *exactly* right relative to S. For can't I guide my doing A with S even if there's another possible action, B, that would be belief-relative better relative to S? As I argue in Sepielli (work in progress-b) and (work in progress-c), the answer is "no". In that case, S would at most *incompletely* intentionally explain my doing A, and thus my doing A would only be *partially* guided with S.

not obvious that the problem with the objective "ought" *qua* action-guide is with the "objective" part of it or with the "ought" part of it.

It's true that, if Jill does what she most likely objectively ought to do, she will fail to prescribe Drug A in the first case, and fail to prescribe Drug Y in the second case. But suppose that Jill believes in a form of objective consequentialism that specifies not only what she objectively ought to do, but also the strengths of objective reasons for various actions. Her belief in this scalar version of objective consequentialism gives Jill at least three other options.

Option #1: Consider Jackson's second case. Armed with her new, scalar form of consequentialism, Jill will have a credence of .9 that prescribing X has a fairly high value, and a credence of .1 that prescribing X has an extremely low value; she will also have a credence of .5 that prescribing Y has a fairly high value, and a credence of .5 that prescribing Y has a moderate value. Nothing that Jackson says forecloses the possibility that Jill may guide her prescribing drug Y *with this state of uncertainty*, without the need to form *any* further attitudes. If she does this, then she will have done the conscientious thing, without employing a subjective normative concept of any sort.

Option #2: Similarly, nothing that Jackson says forecloses the possibility of Jill's forming the *additional belief* that *prescribing Y has the highest expected objective value (EOV)*, and guiding her prescribing Y *with that belief*. Again, if she does this, she will have done what Jackson says is the morally conscientious thing, but without needing to employ a subjective norm. (To be clear about what's being envisaged: Jill is not guiding her action with the belief that she subjectively ought to do the action with the highest expected objective value; she is simply guiding it with the belief that prescribing Y has the highest expected objective value—no subjective normative concepts required.)

Option #3: Finally, nothing that Jackson says forecloses the possibility of Jill's forming the *additional belief* that *prescribing Y would be her best try, or best attempt, at doing what is objectively valuable*, and guiding her prescribing Y *with this belief*.[35]

[35] For a "scalar" response to Jackson, see Menzies and Oddie 1992. For more on scalar moral theories, see Norcross 2006.

Now, I don't think any of these options represents a way of successfully denying the explanandum of this paper—but this isn't because of anything having to do with Jackson cases. Option #3 fails because it's not a way of denying the explanandum at all. As I explained earlier, "S subjectively ought to do A" simply *means* "A would be S's best try at doing what is objectively valuable".

Nor does Option #2 represent a way of denying the explanandum. If I guide my behavior with the full belief, or certainty, that some action has the highest expected objective value, I have not guided it with a state of uncertainty regarding the objective value of the action, but rather with a full belief that the action is the one that maximizes some quantity: EOV.

Option #2 also fails because we cannot generally guide action with the belief that that action has the highest EOV. For it is an open question whether I ought, in any sense, to do the action with the highest EOV (why not be risk-averse instead?), just as it is an open question whether I ought to do the action that, say, yields the most equal distribution of resources (why not be a prioritarian instead?).[36] In other words, I can "step back" and sensibly wonder about the import of some action's having the highest EOV. I cannot, by contrast, sensibly wonder about the import of some action's being the one that I have most reason to do.

So far, we have no grounds for ruling out Option #1. As I've been claiming, this present argument of Jackson's does nothing to disparage it, because Jackson artificially restricts his focus to uncertainty among objective OUGHTs. And we've seen that neither the "metacognitive signal" nor the "inside of an agent" explanation seems to work. To rule out Option #1, we will need an alternative explanation. It's to such an explanation that I'll now briefly turn.

[36] In other work, I do argue that in selected "deliberative contexts", we can guide our behavior with beliefs about EOV, and even with beliefs with non-normative propositions as their content. See Sepielli (work in progress-a). We have no need in these contexts to appeal to beliefs with normative contents, or to desires understood as distinct mental entities. (See McDowell 1978 for a similar view, as well as the distinction between desires as distinct entities, and desires as ascribable *by courtesy* whenever motivation by beliefs occurs.) My point here is simply that we cannot *generally* do this. OUGHT-beliefs are, we might say, "deliberative-context-independent" guides to action. EOV-beliefs, EQUAL-DISTRIBUTION-OF-RESOURCES beliefs, and so on, are not.

TOWARDS A BETTER EXPLANATION

Why aren't objective norms sufficient for action-guidance under uncertainty? Why must we step back from uncertainty about objective reasons and form a judgment about what, subjectively, we ought to do relative to the probabilities represented in that uncertainty?

My view is that there is a certain sort of connection between action guidance and phenomenology: One cannot guide one's actions by conscious mental states with a "multidirectional" phenomenology, which, again, is the phenomenology constitutive of conscious uncertainty. This is a very weak claim, for it is compatible with denying many other alleged connections between guidance and phenomenology. For all I claim, there may be no phenomenology common to all mental states upon which one *may* guide one's action; some such states may have no phenomenal character at all. Specifically, as I've tried to emphasize throughout, it is no part of my position that one can only guide one's actions by a mental state with a "unidirectional" phenomenology.

An action cannot count as guided if it is prospectively unintelligible from the agent's perspective. Intelligibility from the agent's perspective is amenability to a particular sort of explanation—one that helps me understand my own action *as my action*. For a contemplated action to be unintelligible in this sense is compatible with its being explicable in all sorts of other ways. I might explain why I will do it by adverting to my nerves and neurons, to my upbringing, to my subconscious resolution of the Oedipal Complex, and even to the states of my mind characterized dispositionally. But all of these are distinctly third-personal sorts of explanation, and as such, none of them help me understand my action as such; hence, none of them render any of my contemplated actions intelligible in the sense I have in mind.

Insofar as one's conscious normative judgments regarding a contemplated action are multidirectional, that action will simply *feel* unintelligible. One will feel that that action does not make any more sense than any of the other contemplated actions towards which the multidirectional states "point". And I shall want to claim that, if an action of mine *feels* unintelligible, it *is* unintelligible from the agential perspective[37]—

[37] A referee suggests that it is implausible that if an action is intelligible, then it feels intelligible. I agree, and this accords with what I've been saying throughout. My claim is the

just as a visual state's having a partly "yellow" phenomenal character is sufficient for its representing something as being yellow, or a tactile sensation's being a hot feeling is sufficient for its representing the touched object as hot. This particular tight connection between feeling and being is a defining feature of the particular sort of intelligibility that is intelligibility from the agent's own perspective.[38] Were I to do the contemplated action in the face of conscious uncertainty, there would no doubt be some possible explanation I or others could give of it, but such an explanation would do nothing to dissolve its unintelligibility from the inside. And no action can be guided that is unintelligible from the inside.

Now, earlier, I said that a mental state renders an action intelligible only if it rationalizes that action. But the converse is not true. For a state of uncertainty over propositions about the objective values of actions might be sufficient for there to be an action with, say, the highest EOV. And if it turns out that, on the correct theory of rationality, the highest-EOV action is the rational one, then this uncertainty will be sufficient to rationalize the highest-EOV act. But it does not render this action intelligible from the agent's perspective. (As a parallel: Suppose that moral rationalism is true, and that it turns out—whether I believe it or not—that the correct moral theory is utilitarianism. This means that my belief that doing A maximizes utility will rationalize—render rational—my doing A. But this belief will not, all by itself, render intelligible my doing A.)

My account differs from the "Jackson cases" explanation in a way that the rationalizing/rendering-intelligible distinction helps to emphasize. I do not say that one cannot guide one's behavior with a state of uncertainty because such uncertainty will yield irrational action. Rather, I locate the inability of uncertainty to serve as an action guide elsewhere. My account differs from the "inside of an agent" explanation in that mine draws on the fact that states of uncertainty are states of a divided mind, while the "inside of an agent" explanation draws instead on the

very different one that if an action *feels un*intelligible, then it *is* unintelligible from the agent's own perspective.

[38] For more on the relationship between a state's phenomenal character and its intentional features, see Horgan and Tienson 2002; Kriegel 2003 and forthcoming; Pitt 2004; and Mendelovici 2010.

fact that states of uncertainty represent an imperfect connection between mind and world. Finally, my account differs from the "metacognitive signal" explanation in a more subtle way. Both accounts rely in some way on the phenomenology of conscious uncertainty. But I take this phenomenology to ground the unintelligibility-conferring features of this state, and in turn, the inability of the state to guide action. The "metacognitive signal" explanation is entirely silent on such features, and treats the phenomenology constitutive of conscious uncertainty as a mere signal of the normative "riskiness" of candidate actions—a signal that in another metaphysically possible world is played by, e.g., an itchy nose.

GOING FORWARD

Subjective norms are supposed to help us guide our actions when we are uncertain about what we objectively ought to do. But of course, we might be uncertain which of those subjective norms are correct. It is a perfectly good, substantive question whether we ought to maximize expected objective value, or be slightly risk-averse, or simply act on the most likely view about what we objectively ought to do, or do something else, under conditions of uncertainty. When we are unsure which subjective norms are correct, we might appeal to further, higher-order subjective norms in order to guide our actions. But again, we might be uncertain which of *these* norms is correct, and so on. Since there seems to be no principled stopping point to this uncertainty, it may seem mysterious how norm-guided action is possible at all.

I think this mystery and others like it can be dissolved, but not unless we think about subjective normativity and action guidance in the right way. We need a decent account of what subjective norms are, and how they relate to objective ones. We need a clear statement of the conditions under which subjective norms become necessary as guides to action. And we need the correct theory of *why* subjective norms become necessary under those conditions. I have tried to provide these things in this paper. I have given an account of subjective normativity that appeals to the notion of trying. I have said that subjective norms are necessary as action guides under uncertainty, but only when that uncertainty is of a phenomenally conscious sort. And I have suggested a theory of why subjective norms are necessary for this purpose, after

rejecting some competing theories of the same. My hope is that, in doing so, I have put subjective normativity on slightly surer footing and helped to foster an appreciation of its importance.

REFERENCES

Alter, Adam L., Oppenheimer, Daniel M., Epley, Nicholas, and Eyre, Rebecca N. (2007) 'Overcoming Intuition: Metacognitive Difficulty Activates Analytic Reasoning', *Journal of Experimental Psychology: General*, 136: 569–76.

Anscombe, G. E. M. (1957) *Intention* (Oxford: Blackwell).

——(1979) 'Under a Description', *Nous*, 13: 219–33.

Block, Ned (1995) 'On a Confusion about a Function of Consciousness', *Behavioral and Brain Sciences*, 18: 227–87.

Candlish, Stewart (1984) 'Inner and Outer Basic Action', *Proceedings of the Aristotelian Society*, 84: 83–102.

Dowell, Janice (work in progress) 'Flexible Contextualism about "Ought"'.

Evans, Gareth (1982) *The Varieties of Reference* (Oxford: Clarendon Press).

Fodor, Jerry A. (1998) *Concepts: Where Cognitive Science Went Wrong* (New York: Oxford University Press).

Gibbard, Allan (2005) 'Truth and Correct Belief', *Philosophical Issues*, 15: 338–50.

Gollwitzer, P. M. (1990) 'Action Phases and Mindsets', in E. T. Higgins and R. M. Sorrentino (eds.), *Handbook of Motivation and Cognition*, Vol. 2 (New York: Guilford Press), 53–92.

Hajek, Alan (2010) 'Interpretations of Probability', in E. N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy*.

Hellie, Benj (forthcoming) *Conscious Life* (Oxford University Press).

Holton, Richard (2009) *Willing, Wanting, Waiting* (New York: Oxford University Press).

Horgan, Terence and Tienson, John (2002) 'The Intentionality of Phenomenology and the Phenomenology of Intentionality', in D. Chalmers (ed.), *Philosophy of Mind: Classic and Contemporary Readings* (New York: Oxford University Press), 520–32.

Jackson, Frank (1991) 'Decision-Theoretic Consequentialism and the Nearest and Dearest Objection', *Ethics*, 101: 461–82.

Kolodny, Niko and MacFarlane, John (work in progress) 'Ought: Between Objective and Subjective'.

Kriegel, Uriah (2003) 'Is Intentionality Dependent on Consciousness?', *Philosophical Studies*, 116: 271–307.

——(forthcoming) 'The Phenomenal Intentionality Research Program', in U. Kriegel (ed.), *Phenomenal Intentionality: New Essays* (New York: Oxford University Press).

Kruglanski, Arie W. and Webster, Donna M. (1996) 'Motivated Closing of the Mind: "Seizing" and "Freezing"', *Psychological Review*, 103: 263–83.

Lockhart, Ted (2000) *Moral Uncertainty and its Consequences* (New York: Oxford University Press).

McDowell, John (1978) 'Are Moral Requirements Hypothetical Imperatives?' *Proceedings of the Aristotelian Society, Supplementary Volume*, 52: 13–29.

——(2009) 'What Myth?', in *The Engaged Intellect* (Cambridge, MA: Harvard University Press), 308–23.

Mendelovici, Angela (2010) 'Mental Representation and Closely Conflated Topics' (Ph.D. Thesis, Princeton University).

Menzies, Peter and Oddie, Graham (1992) 'An Objectivist's Guide to Subjective Moral Value', *Ethics*, 102: 512–33.

Nagel, Jennifer (2010) 'Epistemic Anxiety and Adaptive Invariantism', *Philosophical Perspectives*, 24: 407–35.

Norcross, Alastair (2006) 'Reasons Without Demands: Rethinking Rightness', in J. Dreier (ed.), *Contemporary Debates in Moral Theory* (Oxford: Blackwell), 38–54.

Parfit, Derek (2011) *On What Matters* (Oxford University Press).

Pitt, David (2004) 'The Phenomenology of Cognition, or, "What Is It Like to Think That P?"', *Philosophy and Phenomenological Research*, 69: 1–36.

Railton, Peter (1984) 'Alienation, Consequentialism, and the Demands of Morality', *Philosophy and Public Affairs*, 13: 134–71.

——(2009) 'Practical Competence and Fluent Agency', in D. Sobel and S. Wall (eds.), *Reasons for Action* (New York: Oxford University Press), 81–115.

Ross, Jacob (2006) 'Rejecting Ethical Deflationism', *Ethics*, 116: 742–68.

Schroeder, Mark (2008) *Being For: Evaluating the Semantic Program of Expressivism* (New York: Oxford University Press).

Sepielli, Andrew (2009) 'What to Do When You Don't Know What to Do', in R. Schafer-Landau (ed.), *Oxford Studies in Metaethics*, Vol. 4 (Oxford: Oxford University Press), 5–28.

——(2010) 'Along an Imperfectly-Lighted Path' (Ph.D. Thesis, Rutgers University).

——(work in progress-a) 'Why We Need Subjective Normativity'.

——(work in progress-b) 'Subjective Normativity, Action Guidance, and Regress'.

——(work in progress-c) 'What to Do When You Don't Know What to Do When You Don't Know What to Do'.

Smith, Holly (2010) 'Subjective Rightness', *Social Philosophy and Policy*, 27: 64–110.

Turri, John (2009) 'The Ontology of Epistemic Reasons', *Nous*, 43: 490–512.

Yalcin, Seth (2011) 'Nonfactualism about Epistemic Modality', in A. Egan and B. Weatherson (eds.), *Epistemic Modality* (Oxford University Press), 295–332.