

# Normative Uncertainty and the Dependence Problem

Abelard Podgorski

## I. Introduction

One of the many inconveniences of being finite and fallible human agents is the need to make choices through the fog of ignorance. Penny the poker player must decide whether to play or fold without knowing for certain whether her opponent is holding aces or bluffing. Danny the doctor must decide whether to give a patient a risky drug without knowing for certain whether it will cure them or kill them. Ginny the gourmet must decide whether to order chicken or falafel without knowing for certain whether animals deserve moral consideration. It is natural to think that what these agents ought to do, in some sense, is sensitive to their uncertainty.

Uncertainty, however, comes in at least two kinds. Danny and Penny are uncertain about *descriptive* facts, like the causal effects of their action. Ginny, on the other hand, is uncertain about *normative* facts, like which moral theory is correct. And while a great deal of philosophical attention in the past century has been devoted to the development of norms of action that are sensitive to descriptive uncertainty, it is only relatively recently that the role of distinctively normative uncertainty has been seriously and systematically investigated.<sup>1</sup> Indeed, while it is widely agreed that there is some sense in which what agents like Penny or Danny ought to do, both morally and

---

<sup>1</sup> Though it is worth noting that the late medieval scholastic tradition was deeply engaged with the topic of moral uncertainty. An overview of some of these debates can be found in Schüssler (2005).

prudentially, is sensitive to their uncertainty, it is controversial whether Ginny's uncertainty matters in any analogous way.

A number of theorists (Lockhart 2000, Ross 2006, Sepielli 2009, Macaskill 2014) have tried to provide accounts of decisionmaking under normative uncertainty. But there has been a backlash against this entire project from a contingent of philosophers who argue that *only* your descriptive uncertainty matters for what you ought to do (Weatherson 2014, Harman 2015, Hedden 2016).

In this paper, I draw attention to a puzzling kind of case which the controversy over normative uncertainty has neglected – where our uncertainty about descriptive facts *depends* on our uncertainty about normative facts. I will show that most existing views about decisionmaking, both those that take into account normative uncertainty and those that do not, give unacceptable recommendations in these cases. But the problem, I'll go on to argue, is much worse for views that don't take normative uncertainty into account at all. To avoid the counterexamples, views must be sensitive not only to both kinds of uncertainty, but also to the *relationship* between them. As I'll show, there is a relatively painless strategy for incorporating this kind of sensitivity into existing views that respect normative uncertainty, but no similarly easy solution in reach for those who claim that only descriptive uncertainty matters. Ultimately, then, the problem is ammunition against the recent skepticism of theorizing about normative uncertainty. Indeed, it raises doubts that there are *any* interesting norms that are sensitive to merely descriptive uncertainty. The backlash against theorizing about normative uncertainty, I will argue, is misguided.

Before we get into the main argument, however, it is important to set the stage and introduce some helpful terminology.

## Objective and Subjective Norms

Let us distinguish three kinds of norms governing actions, based on their relationship to an agent's uncertainty. Under the heading of "uncertainty", I will lump together both the agent's *actual* uncertainty, and their *rational* (or evidentially-supported) uncertainty. This distinction will become significant later in the paper, but for the moment we may treat them jointly.<sup>2</sup>

A norm is *objective* if its prescriptions depend on the way the world is, and not on the agent's uncertainty (rational or otherwise) about the world. *Objectivist hedonist utilitarianism*, which says that an agent ought to perform whichever action maximizes the total amount of pleasure, is an example of an objective (moral) norm. Its prescriptions depend only on the actual distribution of pleasure resulting from potential actions, and not on an agent's beliefs or uncertainty about that distribution.

A norm is *descriptively subjective* or *D-subjective* if its prescriptions depend on an agent's uncertainty about descriptive facts but not on their uncertainty about normative facts (such as what is fundamentally of value or which moral theory is true). *Expectabilist hedonist utilitarianism*, which says that an agent ought to perform whichever action maximizes the *expected* pleasure generated by their action<sup>3</sup>, is an example of a D-subjective norm. Its prescriptions depend on the agent's uncertainty about descriptive facts about pleasure, but not on the agent's uncertainty about normative facts, including about whether pleasure is valuable.

---

2 This means that we are treating jointly what Parfit (2011) distinguishes as the 'belief-relative' and 'evidence-relative' senses of "ought".

3 The technical notion of expectation will be explained in more detail later.

A norm is *normatively subjective* or *N-subjective* if its prescriptions depend on an agent's uncertainty about distinctly normative facts.<sup>4</sup> “*My Favourite Theory*”, which instructs an agent to perform whichever action is recommended by the moral theory in which they are the most confident, is an example of an N-subjective norm. An agent's descriptive uncertainty may matter, on this view, for what they ought to do (since it affects the recommendations of different moral theories), but so does their uncertainty about what the correct moral theory is.

Having made these distinctions, it remains open whether there is any interesting sense of “ought” which is best accounted for by norms of each type. That is the primary topic of this paper. More specifically, our concern is with the two types of subjective norms. It is widely accepted that there are subjective uses of “ought”. The debate I want to enter is over whether these subjective oughts are best explained by merely *D-subjective* norms or whether, at least sometimes, we need to appeal to *N-subjective* norms.

## **N-subjective Norms: For and Against**

Defenders of N-subjective norms typically argue that the same sorts of considerations that motivate taking descriptive uncertainty into account also motivate taking normative uncertainty into account. If true, this puts a D-subjectivist in a bind – without these, it would be no longer clear why

---

<sup>4</sup> Norms that are sensitive to both normative and descriptive uncertainty, then, count as normatively subjective. Nobody, as far as I know, has defended norms that are sensitive to normative but not descriptive uncertainty.

we need subjective norms at all. Three such considerations feature regularly in defenses of N-subjective norms.<sup>5</sup>

First, it is suggested, we need norms that are sensitive to uncertainty of both kinds in order to explain what is problematic about doing things that are intuitively objectionably *risky*, from the agent's point of view (MacAskill 2016, Moller 2011). If Penny calls her opponent, despite believing that he probably has a strong hand, or Danny gives his patient the drug, despite believing it quite likely that it will kill the patient, there is something problematic about their actions – Penny is being reckless with her money, and Danny with his patient's life. Similarly, N-subjectivists argue, if Ginny believes it to be quite likely that eating meat is deeply wrong, she too would be taking a kind of moral risk, from her perspective, by ordering the chicken. Just as we need a D-subjective norm to explain what is objectionable about Penny and Danny's actions, we need an N-subjective norm to explain what is objectionable about Ginny's action.

The second main reason adduced in favor of subjective norms concerns *action-guidance*. (Ross 2006, Sepielli 2012) One important role for norms to play, the thought goes, is that of guiding our behavior. We consider what we ought to do, and, having formed a judgment, act accordingly. Purely objective norms cannot adequately play this role. These objective norms will tell Penny that she should call if and only if her opponent is in fact bluffing, and will tell Danny to give his patient the drug if and only if it will in fact cure them. But this is no help at all to Penny or Danny, because they are not in a position to reach the correct objective “ought” judgment (besides taking a wild,

---

<sup>5</sup> For a characteristic example of a D-subjectivist who appeals to these sorts of motivations in favor of their own view, see Hedden (2012).

irrational stab in the dark), and thus guide their behavior by the norm. D-subjective norms, since they require as input only Penny and Danny's beliefs or credences about the relevant facts, can guide their actions despite their uncertainty. Similarly, N-subjectivists suggest, Ginny is not in a position to guide herself by what she ought to do in a D-subjective sense, since she is not sure which moral theory is correct. So she must use a norm which takes into account her beliefs or credences about the moral status of animals.<sup>6</sup>

Thirdly, it is argued that we need subjective norms to serve as the proper basis for *blame* and related normative criticism (Guerrero 2007). Whether an agent has violated an *objective* norm seems to have very little connection with whether it is appropriate to blame them. If Penny calls the bluff and wins the pot, and if Danny gives his patient the risky drug and cures them, they have done as they objectively ought but nevertheless may deserve blame. And it seems inappropriate to blame someone for failing to live up to an objective norm when they were ignorant of some of the relevant facts. If Danny thinks there's a good chance the drug will kill his patient, he cannot be blamed for failing to administer it. Danny's ignorance (or at least his rational ignorance) matters for his blameworthiness, and so if we want a norm with a close connection to our practices of praise and blame, it must take the agent's ignorance into account. Once again, the N-subjectivist suggests, the same goes for normative ignorance. It would be inappropriate to blame agents on the basis of purely D-subjective norms, when they are (rationally) ignorant of the relevant normative facts.

---

<sup>6</sup> This is one version of the action-guidance argument. There is significant disagreement in the literature over exactly how to properly characterize the action-guidance requirement (see e.g. Way and Whiting (2017)), and there are concerns about whether even subjective norms could satisfy such requirements, in light of arguments about the opacity of our own mental states (such as Timothy Williamson's (2000) case against luminosity).

For their part, opponents of N-subjective norms raise several challenges. Brian Weatherson (2014) argues that being guided by an N-subjective norm amounts to being motivated by rightness *de dicto*, and that this is a kind of moral *fetishism*. We ought to be motivated by our beliefs about happiness and suffering and the other things that *make* actions right and wrong, not by our beliefs about rightness and wrongness themselves.<sup>7</sup> Elizabeth Harman (2015) argues that there is an important asymmetry between normative and descriptive ignorance – descriptive ignorance exculpates from blame, but normative ignorance does not. If these criticisms are right, then N-subjective norms are poor candidates for the roles of guidance and criticism that we saw were the motivations for subjective, rather than purely objective, norms in the first place.

In addition, there are serious technical obstacles to developing an N-subjective theory. Most attractive candidates for N-subjective norms require us to compare degrees of wrongness between moral theories, and it has proven difficult to give an account of how to do this in a consistent and non-arbitrary manner, especially when the theories involved differ structurally in substantial ways.<sup>8</sup> Finally, a potential regress lurks – an agent may, after all, be uncertain about the right N-subjective norms, requiring further norms that are sensitive to this higher level uncertainty, and about which the agent might again be uncertain. (Sepielli 2009)

I provide this background to give the reader a sense of the debate as it stands and to preclude some disappointed expectations – although I will be coming down on the side of the N-

---

7 See Sepielli (2016) for a response on behalf of the N-subjectivist.

8 The worry is pressed by Gracely (1996), Nissan-Rozen (2015), and Hedden (2016). MacAskill (2016) offers a solution that treats cases of moral uncertainty by analogy to voting problems.

subjectivist, it is not my aim in this paper to respond to the challenges above, which are forceful and would require more space than is available here to address (though some of them will make an appearance later in the paper). Rather, I will be developing a distinct argument in favor of N-subjectivism.

## **II. The Dependence Problem**

### **Normative-Descriptive Dependence**

My strategy will be to present a counterexample to most existing views of decisionmaking under uncertainty, both D-subjective and N-subjective, and then to throw out a lifebuoy that only the N-subjectivist can plausibly grab. The distinctive feature of this kind of counterexample, and I believe the reason it has not been given attention in the literature, is that it involves a dependence relationship between an agent's credence in a normative theory and their credence in descriptive facts. Before I present the case, it will be helpful to discuss this phenomenon more generally to allay some potential skepticism.

For most normative theories and most descriptive facts, our credences in them are independent of each other. Learning that utilitarianism is true would not change my credence that some policy would raise GDP, and learning that some policy would raise GDP, while it might raise my confidence in the rightness of the policy, would not change my credence in utilitarianism, or any other moral theory. With respect to these descriptive facts, our normative uncertainty is in this sense quarantined. The existing discussion of normative uncertainty implicitly proceeds as though all

uncertainty in normative theory is quarantined from all relevant descriptive facts, at least for rational agents.

I think it is not too hard to see why that might initially seem like a plausible assumption. On one picture of descriptive and normative epistemology, our ways of discovering and evaluating these different types of claims are fundamentally distinct. For descriptive claims, we have observation, science, and the experimental method. For normative theories, we have some process of bringing our moral intuitions into reflective equilibrium, or pure rational insight. And, one might think, never the twain shall meet. This is one interpretation of the infamous “is-ought gap”.

But I think the quarantine picture quickly falls apart on inspection. First, there are all sorts of empirical facts about human beings and their intuitions that plausibly matter to our assessment of moral theories (see, for example Appiah 2008). In cases of moral testimony, our moral beliefs are sensitive to our descriptive beliefs about what others say or think.<sup>9</sup> But most relevantly for our purposes, our best understanding of some natural entities is as beings responsive to reasons and other normative facts, and so our expectations about their behavior will be sensitive to our beliefs about those normative facts.

Perhaps the most straightforward examples of this kind of normative-descriptive epistemic dependence involve religious beliefs. Suppose I am convinced that an omnibenevolent, omniscient, and omnipotent being exists and designed the world. Since the being is omnibenevolent, it has designed the world to be the best of all possible worlds, and since the being is omniscient, it knows

---

<sup>9</sup> Though the appropriateness of forming beliefs by moral testimony is controversial. See e.g. Driver (2006), Hopkins (2007), McGrath (2009), Enoch (2014).

the correct theory of value. I, however, am uncertain about the correct theory of value. I am not sure, for instance, whether the world would be intrinsically better if it contained a wide variety of intelligent species. If I were to be convinced of a theory of value that entailed that a variety of intelligent species would make the world better, this would raise my confidence in the descriptive claim that there is intelligent life elsewhere in the universe. Or perhaps I am convinced that the divine being rewards good people and punishes evil people in the afterlife, but am unsure about which moral theory is correct, and therefore unsure whether someone who tells lies in order to spare others suffering is good or evil. My descriptive uncertainty about whether that person will spend the afterlife dancing with the angels or as the main course in some infernal barbecue depends on my normative uncertainty about which moral theory is true.

Although the religious examples are particularly stark, there are more ordinary cases with a similar flavor. Suppose I know my friend Ulrikke is a wiser and better person than I. I also know she spends her Saturdays volunteering for what she judges to be the worthiest cause. I am uncertain between different moral theories and therefore about what the worthiest cause is, and do not know what she believes about the issue, but I think that Ulrikke is likely (though unlike our divine being, not perfectly certain) to get such things right. My credence about where Ulrikke will be on Saturday will, in such a case, depend on my credence in different normative theories; if I learn that a moral theory is true on which the most worthy cause is relieving the suffering of the poor, I might expect to find her at the homeless shelter, while if I learn that a moral theory is true on which the most worthy cause is assisting the proletariat's revolution against the owners of the means of production, I might expect to find her plotting underground with the local communists. Of course, there are

descriptive facts about Ulrikke, which, if I knew them, would make my normative beliefs irrelevant to my predictions of her behavior – facts about what she believes, for example. But in my present position, my best predictions of her behavior will go through my normative beliefs.

I aim to suggest with these examples that there is nothing particularly alien, obviously irrational, or otherwise degenerate about cases of normative-descriptive dependence as such. Although the examples I will use next are a little more outlandish, they should not be rejected merely because they contain such dependence.

### **The Birthday Girl**

To see the problem raised by dependence, it will be helpful to consider and contrast two cases:

#### ***Angel's Present***

Heloise wakes up on the morning of her birthday to a knock on the door. On her doorstep stands an angel, bearing a gift for her. If she opens its gift, the angel informs her warmly, something very good will happen. Heloise knows, we will postulate, that the angel has a perfect understanding of what is good and what is bad, and that it is honest in describing its gift. She takes the box and places it on her coffee table, and considers whether to open it.

### *Devil's Present*

Heloise wakes up on the morning of her birthday to a knock on the door. On her doorstep stands a devil, bearing a gift for her. If she opens its gift, the devil tells her, cackling through a toothy grin, something very bad will happen. Heloise knows, we will postulate, that the devil has a perfect understanding of what is good and what is bad, and that it is honest in describing its gift. She takes the box and places it on her coffee table, and considers whether to open it.

There is a twist, but I will pause here for us to get a sense for how straightforward Heloise's decision should be in each case, before we clutter things up with some additional (and, I will suggest, irrelevant) detail. I don't expect Heloise's decisions strike the reader with the flavor of a moral dilemma. The *easy answer* is that Heloise ought, in every interesting sense of the term, open the angel's present, and leave the devil's present closed. I'll say more in favor of this judgment in a moment, but it is worth appreciating its raw intuitive force. Now let us add the following to the description of both cases:

Heloise holds the box to her ear and shakes it a little, to try and guess what is inside. She hears an unmistakable squeal – a sound that is made by only two rare creatures – the azzip and the buzzip. Opening the box will, she knows, release one of these two creatures. Now Heloise, it turns out, is an ethics professor, and has narrowed the correct theory of value down to two candidates, Theory A and Theory B, with her credence split

evenly between them. According to Theory A, an azzip being released would be good,<sup>10</sup> but a buzzip being released would be very bad. According to Theory B, a buzzip being released would be good, but an azzip being released would be very bad. And let us suppose that in fact, Theory A is true.<sup>11</sup>

We can capture Heloise's normative uncertainty (in both cases) in the diagram below, with the numbers representing the relative value and disvalue of the azzip and buzzip being released according to each theory.

### Normative Uncertainty

<u>Theory A</u>	<u>Theory B</u>
Azzip: +1	Azzip: -2
Buzzip: -2	Buzzip: +1

50%

50%

In both cases, Heloise's descriptive uncertainty depends on her normative uncertainty. Heloise is certain that the box contains an azzip or a buzzip, but she is uncertain which. And her uncertainty depends on her uncertainty about which theory of value is true. We can represent her descriptive uncertainty (in both cases) as follows:

---

10 We can run these cases in terms of moral value or prudential value, interpreting the relevant “ought” claims in the corresponding way – what follows should apply equally well to both sorts of normativity.

11 The reader may be worried that I am treating the values at stake as consequentialist. I do this for a few reasons – first, it allows us to set aside independent (though significant) difficulties that arise when discussing normative uncertainty between theories with very different structures. And second, the dominant approaches both to descriptive and to normative uncertainty are the easiest to apply uncontroversially in a consequentialist framework, and this allows the structural problem I want to describe to present itself most clearly. I do not think that the consequentialist assumption is essential, however – similar cases can be described with a deontological flavor. It would, in any case, be shocking in its own right to find that a reasonable approach to normative uncertainty was inconsistent with the truth of any consequentialist theory.

### Descriptive Uncertainty

Box Contains Azzip	Box Contains Buzzip
--------------------------	---------------------------

50%

50%

“Theory A” and “Theory B” function in this example as placeholders. If we want to make the case more concrete, we could postulate that Theory A is hedonistic utilitarianism and Theory B is desire-satisfaction utilitarianism. If so, the effect of releasing the azzip is one net unit of pleasure and two net units of desire frustration, and the effect of releasing the buzzip is one net unit of desire satisfaction and two net units of suffering.

### In Defense of the Easy Answer

Clearly, what Heloise *objectively* ought to do, in light of all the facts, is open the angel’s present, in the first case, and leave the devil’s present alone, in the second. The angel’s present has an azzip in it, and in fact, it will be good if an azzip is released. The devil’s present has a buzzip in it, and in fact, it will be bad if a buzzip will be released.

But I think it is almost as obvious that what Heloise *subjectively* ought to do is open the angel’s present and leave the devil’s present alone. There are weak and strong versions of this claim.

*The Easy Answer (Weak):* There is *some* interesting subjective sense in which Heloise ought to open the angel’s present and not to open the devil’s present.

*The Easy Answer (Strong)*: In *every* interesting subjective sense, Heloise ought to open the angel's present and not to open the devil's' present.

I think both the weak and the strong claims are true. For the primary aim of this paper, defending N-subjectivism, only the weaker claim is required, so that will be my focus. However, I believe the things to be said in its favor make the stronger claim difficult to deny as well.

Recall the motivations for subjective normativity – capturing intuitively objectionable riskiness, providing a ground for blame, and guiding an agent's actions from the inside. Opening the angel's present is the antithesis of an intuitively risky choice – it's a win-win, no matter which theory is true, and the devil's present is a lose-lose. We would blame Heloise if she opened the devil's present, or failed to open the angel's. And the judgment that she ought to open the angel's present, but not the devil's, is the one we expect her to form and to guide her action – no information necessary for her to be guided in this way is missing to her.

And even aside from these intuitive considerations, I think there are other reasons to think that the easy answer must be correct.

First, it follows from a very minimal and plausible constraint concerning the relationship between objective and subjective norms:

*KO*->*S*: If an agent knows for certain that she objectively ought to  $\varphi$  (for some non-disjunctive action  $\varphi$ )<sup>12</sup>, then she subjectively ought to  $\varphi$ .

---

<sup>12</sup> Disjunctive actions raise complications that make this qualifier necessary. On some views, it follows from the fact that one objectively ought to  $\varphi$  that one objectively ought to  $(\varphi \text{ or } \psi)$  for any  $\psi$ . But an agent who knows the

Recall, the reason to think we needed subjective norms above and beyond objective norms was that we were not always in a position to know what we objectively ought to do. When we *are* in a position to know our objective duties, of course we subjectively should follow them. Heloise knows that if Theory A is true, she objectively ought to open the angel's present and not the devil's. She knows if Theory B is true, she objectively ought to open the angel's present and not the devil's. And she knows that either Theory A or Theory B is true.

And there are several very closely related attractive principles that have the same implication.

*Idealization:* If an agent knows that were she fully informed and reasonable and otherwise in her present position, it would be the case that she subjectively ought to  $\varphi$ , then she subjectively ought to  $\varphi$ .

*Deference:* If an agent knows that a fully informed and reasonable version of herself would advise her to  $\varphi$ , then she subjectively ought to  $\varphi$ .

*Regret:* If an agent knows that she would regret not doing  $\varphi$  if she were fully informed and reasonable, then she subjectively ought to  $\varphi$ .

I will not go through each of them individually, but it should be fairly clear in each case that Heloise meets the antecedent conditions with respect to opening the angel's gift, and leaving the devil's gift unopened. Depending on the correct views about the relationship between regret, advice, and the various senses of "ought", some of these principles might follow more or less trivially from

---

disjunctive ought without knowing whether it is true because they objectively ought to  $\varphi$  or because they objectively ought to  $\psi$  might still be subjectively obligated to do some third thing. The well-known Miners Puzzle (Parfit unpublished) is an example of such a case.

others. I do not want to commit to all of them being correct in as simple a form as I've given above – some of them may have degenerate exceptions common among subjunctive formulations of normative principles.<sup>13</sup> But they seem plausible at least as presumptive rules of thumb, and the fact that all of them point in the same direction is significant.

## The Failure of D-subjectivism

If it seems like I am belaboring an obvious point, it is only because surprisingly many views fail to capture the easy answer.

What should Heloise do, according to the D-subjectivist? There are an infinite number of possible D-subjective theories, but for illustration, let us begin by considering the most well-developed and widely accepted framework of the D-subjective type – expected value maximization. On this picture, each possible descriptive outcome of one's action has a certain objective value, determined by the correct normative theory. What an agent subjectively ought to do is perform the action that has the highest *expected* value, which is given by the following:<sup>14</sup>

---

13 For example, a fully informed version of myself might never need to collect more information, given their omniscience, but we should not infer from this that I ought not do so. This threatens *Idealization* as stated.

14 It is worth noting that though the expected utility framework fits very naturally with purely consequentialist and agent-neutral theories of value, in principle the value of an outcome can depend on agent-relative or act-relative features of the situation as well.

*Maximize Expected Value (MEV)*: One ought to do an action that maximizes EV, where  $EV(\varphi) = \sum_i P(O_i|\varphi)V(O_i)$ ,  $P(O_i|\varphi)$  is the agent's credence in  $O_i$  conditional on  $\varphi$ , and  $V(O_i)$  is the objective value of that outcome, according to the true theory.<sup>15</sup>

This is a D-subjective picture, because its recommendation depends on the agent's descriptive credences about the likely outcomes of her action, but on the *true* theory (Theory A) about what is valuable, regardless of what she thinks is valuable. In the concrete version of the case where Theory A is hedonistic utilitarianism, *MEV* tells Heloise to maximize expected pleasure. Heloise's expected value for opening the angel's box is 0.5 (her credence that the box contains an azzip) x 1 (the objective value of the outcome if the box contains an azzip) + 0.5 (her credence that the box contains a buzzip) x -2 (the objective value of the outcome if the box contains a buzzip), or -0.5. Her expected value for leaving the box unopened is 0. Therefore, Heloise ought to leave the angel's gift unopened. Since her descriptive credences about what is in the boxes are exactly the same in the case of the devil's gift, she likewise ought to leave that gift unopened.

This is bad – the easy answer was that she should open the angel's gift. The view gets the intuitively correct answer for the *devil's* gift, but this only highlights the depth of the problem. No matter what values the true theory assigns to the azzip and the buzzip, the D-subjective expected value of opening the devil's gift is always going to be the *same* as the D-subjective expected value of opening the angel's gift, since Heloise's descriptive credences about the outcomes of opening the box are identical in both cases, and that is the only variable that makes a difference. Moreover, this is

---

<sup>15</sup> This is the formula for what is called *evidential* expected value. The differences between causal and evidential expected value theories will not concern us here.

not limited to the expected value theory – it is endemic to D-subjectivism. *Any* D-subjectivist view makes what Heloise ought to do in different cases depend only on differences in her descriptive credences about the outcomes of each action, and as a result, any D-subjectivist view can get the right result for the angel's gift only by getting the wrong result for the devil's gift, and vice versa. No amount of tinkering with risk aversion or other alternatives to the expectational framework will help. Because this problem is a consequence of the dependence of Heloise's descriptive beliefs on her normative beliefs, we can call this the dependence problem.

### **The Failure of N-subjectivism?**

Before we consider some possible responses on behalf of the D-subjectivist, we should make sure that the dependence problem does not equally haunt the N-subjectivist. And it turns out that we should not break out the champagne and declare victory too early, for most N-subjective norms in the literature *also* fail to accommodate the easy answer. Take the N-subjective norm Hedden (2016) calls *Maximize Inter-Theoretic Expectation*:

*Maximize Inter-Theoretic Expectation (MITE)*: One ought to do an action that maximizes ITE, where  $ITE(\varphi) = \sum_i P(T_i)EV_i(\varphi)$ , where  $EV_i(\varphi)$  is the D-subjective expected value of  $\varphi$  under normative theory  $T_i$ .

This norm, which Hedden judges to be so much more attractive than alternative N-subjective norms that he calls it “the only game in town”, and which is defended in some form by Lockhart (2000), Ross (2006), and MacAskill (2014)<sup>16</sup>, tells us to get the intertheoretic expectation of an action by

---

<sup>16</sup> Hedden attributes *MITE* to Sepielli (2009) as well, but as I will note later, Sepielli's view is importantly different.

multiplying the probability that each theory of value is right with the associated D-subjective expected value of the action under that theory, and pick the action that maximizes that expectation.

In *Angel's Present*, since the D-subjective expected value of opening the present, given Heloise's credences, is -.5 on both Theory A and Theory B (we already calculated this in the previous section for Theory A, and the two theories are symmetric), her ITE is  $.5(-.5) + .5(-.5)$ , or -.5. So MITE tells us, incorrectly, that Heloise ought to leave the present unopened. And again, no matter what the relative values of the outcomes, because her credences in descriptive outcomes and normative theories are the same in *Angel's Present* and *Devil's Present*, it will not be able to distinguish them.

It is not just expectational theories which have this consequence. Like MITE, most N-subjective norms in the literature operate by performing some operation on intermediate D-subjective norms in which the agent has confidence. Take *My Favourite Theory*, a version of which is defended by Gustafsson and Torpman (2014):

*My Favourite Theory (MFT)*: One ought to do an action that is permitted by a (D-subjective) theory in which you have the highest credence.<sup>17</sup>

Gustafsson and Torpman do not use the language of "D-subjectivity" themselves, but the theory in question must be a D-subjective one, or else their view ends up being insensitive to descriptive uncertainty. And no matter what D-subjective theory Heloise has the highest credence in, it must, as we saw earlier, give the same verdict on *Angel's Present* and *Devil's Present*.

---

<sup>17</sup> Gustafsson and Torpman's final version of this principle is slightly more complex, but in ways that will not concern us here.

Even the following very weak dominance principle (implied by both views above) is problematic:

*Dominance:* If one ought to  $\varphi$  according to every D-subjective theory in which one has credence, then one N-subjectively ought to  $\varphi$ .<sup>18</sup>

This seems awfully plausible on its face, once we begin to take N-subjectivity seriously. But in Heloise's case, opening the angel's present will dominate on each D-subjective theory if and only if opening the devil's present does.

## A Diagnosis

It may seem, then, that *Angel's Present* and *Devil's Present* are a problem for everyone. But there is a way out. First, let us diagnose the source of the failure. All the theories we have discussed have been unable to avoid the implication that Heloise ought to do the same thing in *Angel's Present* as she does in *Devil's Present*. The reason that this is so hard is that a) her descriptive uncertainty about the outcomes is the same in both cases, and b) her normative uncertainty between theories is the same in both cases. The D-subjectivist views fail because they claim what Heloise ought to do is a function of her descriptive uncertainty about the outcomes. The N-subjectivist views we considered fail because they claim what Heloise ought to do is a function of her descriptive uncertainty about the outcomes and her normative uncertainty between theories, taken independently. None of them are sensitive to the *relationship* between her normative and her descriptive uncertainty.

---

<sup>18</sup> Macaskill (2014, pg. 21) and Ross (2006) both use something stronger than this principle: they suggest that if one D-subjectively ought to  $\varphi$  according to some theories in which one has credence, and all other theories rank  $\varphi$ ing equal to every other option, then one N-subjectively ought to  $\varphi$ .

The D-subjectivist is out of luck – since they do not acknowledge the relevance of normative uncertainty at all, they can't take into account the relationship between normative and descriptive uncertainty. But the N-subjectivist is in a better position. Indeed, there are multiple ways for an N-subjectivist view to capture this kind of sensitivity. The view defended by Sepielli (2009)<sup>19</sup> turns out to be one. But I want to describe a more general strategy which can be applied as a fix to all the N-subjective norms we have so far discussed.

### **The Conditionalizing Strategy**

The relationship between our normative and descriptive uncertainty can be represented in terms of *conditional probabilities*. When our normative uncertainty is quarantined from our descriptive uncertainty, then the probability of all our descriptive claims conditional on the truth of any normative theory is the same as their unconditional probability. In cases like *Angel's Present*, the conditional probabilities and the unconditional ones come apart. The unconditional probability Heloise assigns to the box containing an azzip is 0.5, but the probability conditional on the truth of Theory A is 1. The problem with the N-subjective norms we considered earlier is that they don't allow these conditional probabilities to play a role. But this is easy to correct. Whenever we get to the part of the decision procedure where we draw out the implications of one of the normative theories in which we have credence, we should be using our descriptive credences *conditional on that theory*, not our *unconditional* descriptive credences. For example, we can modify *MITE* as follows:

---

<sup>19</sup> Sepielli's view advises *maximizing expected objective value*: Perform the action with the highest EOV, where the EOV of an action is calculated by multiplying the subjective probability that some (objective) practical comparative is true by the objective value of that action if it is true, doing the same for all the other practical comparatives, and adding up the results.

*Maximize Inter-Theoretic Conditionalized Expectation (MITCE)*: Do the action that maximizes ITCE, where  $ITCE(\varphi) = \sum_i P(T_i)CEV_i(\varphi)$ , and  $CEV_i(\varphi)$  is the D-subjective expected value of  $\varphi$  under  $T_i$  using the agent's credences *after conditionalizing* on  $T_i$ .

The difference between MITE and MITCE is that MITE holds fixed the agent's present descriptive credences when calculating the D-subjective expected value of an action under a theory, and MITCE uses the agent's credences *after* conditionalizing on the theory itself. This gives opening the angel's present an ITCE of  $0.5(1) + 0.5(1) = 1$ . Opening the devil's present, on the other hand, has an ITCE of  $0.5(-2) + 0.5(-2) = -2$ . So Heloise ought to open the angel's present, and not the devil's present.

One nice thing about the conditionalizing strategy, compared to other ways of incorporating normative-descriptive dependence (such as using the expectational framework but treating the truth of a normative theory as part of the *outcome* of an action) is that one can apply it to all kinds of potential norms, and not just those with an expectational structure. We can fix *MFT* in a parallel way:

*My Favourite Conditionalized Theory (MFCT)*: Do an action that is permitted by a D-subjective theory in which one is most confident, after conditionalizing on that theory.

In the same vein:

*Conditionalized Dominance*: If one ought to  $\varphi$  according to every D-subjective theory in which one has credence, after conditionalizing on that theory, then one N-subjectively ought to  $\varphi$ .

The reader can check for themselves that these views no longer get the wrong results in any of the cases we have considered. As far as I can tell, there is no reason for an N-subjectivist to resist these modifications. When normative uncertainty is quarantined, the conditionalized versions of the principles are equivalent to the originals. The only reason they have not been incorporated is that people have simply not been considering cases of normative-descriptive dependence.<sup>20</sup>

### III. D-Subjectivist Responses

#### D-Subjectivism and De Re Moral Concern

One option available to the D-subjectivist, of course, is simply to bite the bullet. This would involve rejecting the easy answer to Heloise's dilemmas, and rejecting attractive principles like  $KO \rightarrow S$ , *Idealization*, *Deference*, and *Regret*. Normally, this would seem less like a response than a capitulation, but the D-subjectivist might have something further to say in favor of it here. This is because at least one existing motivation for D-subjectivism already implies we should be skeptical about principles like  $KO \rightarrow S$ , and about the kind of reasoning driving us towards the easy answer, even before we start considering cases of dependence.

According to the *de re* D-subjectivist, a morally decent agent is not motivated by thoughts about rightness or goodness at all. Rather, they are motivated by the considerations that *make* actions right or wrong, such as facts about happiness or suffering. When these things are at stake, the

---

<sup>20</sup> The view is too complex to describe in detail here, but MacAskill's (2016) voting-theoretic solution to the problem of incomparability can equally accommodate the conditionalizing strategy, by taking each theory's choice-worthiness rankings post-conditionalization to determine its voting preferences.

defenders of this strategy suggest, being motivated by one's thoughts about bare goodness or rightness amounts to a sort of moral *fetishism*. (Smith 1994, Arpaly 2002, Weatherson 2013)

This charge is typically motivated by reflection on cases involving *inverse akrasia*, where an agent acts contrary to their misguided beliefs about what they ought to do, while being responsive to the things that in fact give us moral reasons. The standard case of this sort is that of Huck Finn, who, due to his upbringing in antebellum Mississippi, believes that he is obligated to turn his friend Jim, an escaped slave, in to his former master. When Huck, out of friendship and compassion, refuses to do so, many intuitively judge him praiseworthy for what he does, despite doing what he has every reason to believe (it is stipulated) is wrong.

Someone who endorses this line might claim that Heloise is effectively in Huck Finn's position. Unlike him, she doesn't have a *false* belief about what she objectively ought to do, but, the argument goes, Huck's case shows she should not be motivated by those beliefs at all, whether or not they are true. And from her perspective, both the devil's and the angel's boxes have a high chance of producing a great deal of suffering – this should make her reluctant to open either box, no matter her moral beliefs. Moreover, if this is correct, then it is not surprising that principles like *KO*->*S*, which require an agent to be a responsive to her *moral* understanding, will fail to hold.

I now think this is probably the D-subjectivist's best chance at a response. In part, whether one finds it satisfying will depend on how heavily one is willing to lean on the charge of moral fetishism, which is controversial.<sup>21</sup> The main thing I want to bring out here, however, is that in order

---

21 See for example Wolf (1986), Olson (2002) and Carbonell (2013) for defenses of de dicto moral concern.

to furnish a complete response to the dependence problem, one has to draw an *especially* strong lesson from cases like Huck's, much stronger than the examples themselves can bear on their own.

First, though perhaps least worryingly, the inverse akrasia cases typically discussed involve *moral* motivation, but the dependence problem arises for other normative domains as well. So if the *de re* response is to rescue D-subjectivism in general, the charge of fetishism has to be extended to agents that respond to *de dicto* concern with prudence, rationality, and so on. It is at least not immediately obvious that the same lesson can be drawn across them all.

Second, there is considerable space between the claim that a moral agent must be motivated *only* by her *de dicto* beliefs about rightness, which Huck Finn cases arguably refute, and the claim that an agent's beliefs about rightness are *irrelevant* to what she ought to do, which is what is needed to take the bite out of the dependence problem. The latter claim is quite extreme – it seems, for example, to make a mystery of the practical relevance of moral deliberation and inquiry altogether.<sup>22</sup> As long as it makes sense for an agent to grant any weight whatsoever to *de dicto* moral considerations, a version of the dependence problem arises. Indeed, we can imagine a version of the *Angel's Present* and *Devil's Present* cases where the agent does not believe that suffering or any other morally relevant consideration *de re* is at stake, because her credence is split between false moral theories which do not acknowledge those values. The D-subjectivist will have to claim that even then, when there is no room for compassion or other *de re* concern to motivate the agent one way or another, she should be completely indifferent to her knowledge that opening the angel's present

---

<sup>22</sup> Sepielli (2016) presses this point against the charge of fetishism.

would be objectively good, and the devil's present objectively bad. This is more than the standard intuitions can support.

Finally, and I believe worst of all, to undermine the dependence problem, the D-subjectivist would have to reject as fetishistic motivation not just by things like bare rightness and goodness itself, but also by *rightmaking* normative considerations like fairness, cruelty, respect, loyalty, honor, and well-being.<sup>23</sup> This is because versions of *Angel's Present* and *Devil's Present* can easily be constructed in terms of these normative concepts. When Heloise's descriptive beliefs depend on her uncertainty between theories of well-being, or fairness, or respect, it can turn out that Heloise is certain not just about the objective goodness and rightness of her options, but that opening the devil's present would be harmful, unfair, and disrespectful and that opening the angel's present would be benevolent, fair, and respectful, without any difference in her descriptive credences about the contents of the boxes. And it is much harder to swallow the suggestion that there is something fetishistic about being motivated by any of these. After all, the natural way to describe Huck Finn's own motivations are in terms of normative concepts like these - at the very least, by the consideration of what would be *harmful* or *disloyal* to his friend Jim. Even the committed defenders of inverse akrasia do not take their argument this far.<sup>24</sup>

---

23 Many of these are paradigmatic examples of what philosophers call "thick" normative concepts in contrast with "thin" concepts like rightness and goodness, but I avoid this here to sidestep controversy over what the thickness/thinness distinction amounts to, and because it is not obvious that thickness is what is relevant for our purposes. See Smith (2013) for a discussion of this distinction.

24 Weatherson (2013) acknowledges, for example, that the charge of fetishism does not seem to apply to all normative concern, though he does not seem to appreciate just how many plausible objects of appropriate concern are normative in nature. Weatherson has in mind considerations of vices or virtues an action might manifest, but even concern with benefits and harms is normatively loaded.

Ultimately, then, even if the appeal to cases of inverse akrasia grounds an objection to certain N-subjectivist principles which give exclusive weight to rightness or goodness, it is not enough for a defense of a fully D-subjective picture.

### **Truth-Conditionalized D-Subjectivism**

Is there anything left for the D-subjectivist to say if they are not prepared to commit to the extreme *de re* strategy above? The success of the conditionalizing strategy for the N-subjectivist should naturally make us wonder whether there is a similar move available to the D-subjectivist. The use of unconditional descriptive credences guaranteed that we could not distinguish taking the angel's present from taking the devil's present. What if instead of using our unconditional descriptive credences, we used our conditional credences – not on the normative theories in which we have credence (as in the repaired N-subjectivist norms) but on the *true* theories?<sup>25</sup>

This kind of proposal, which we can call *Truth-Conditionalized D-Subjectivism*, would not be exactly D-subjectivist in letter, since it makes what an agent ought to do sensitive not to her descriptive beliefs at all, but to certain hybrid conditional credences. But it is arguably D-subjectivist in spirit, since it implies that two agents who differ only in their uncertainty between normative theories will never differ in what they ought to do, and in cases where our descriptive credences are quarantined from our normative credences, it has the same recommendations as traditional D-subjective views.

---

<sup>25</sup> I would like to thank two anonymous referees at *Mind* for proposing this strategy on behalf of the D-subjectivist.

Unlike standard D-subjective views, however, a truth-conditionalized norm can accommodate the easy answer to the Birthday Girl cases. To see this, consider the natural expectational version of such a view:

*Maximize Truth-Conditionalized Expected Value (MTCEV)*: One ought to do an action that maximizes TCEV, where  $TCEV(\varphi) = \sum_i P(O_i | \varphi \ \& \ T) V(O_i)$ , where T is the conjunction of all true normative theories, and  $V(O_i)$  is the objective value of that outcome, according to T.

For Heloise, the probability of an azzip being released, conditional on her opening the angel's present *and the true Theory A* is 1. And the objective value of an azzip being released is 1. The probability of a buzzip being released, conditional on her opening the devil's present *and the true Theory A* is 1. And the objective value of a buzzip being released is -2. So *MCTEV* advises Heloise to open the angel's present and leave the devil's present closed.

So the view gets better results than the standard D-subjective views in cases of normative-descriptive dependence. Moreover, in cases where there is no such dependence,  $P(O_i | \varphi \ \& \ T)$  will be the same as  $P(O_i | \varphi)$  and the view has the same recommendations as *MEV*. So far, so good.

However, I think there is a fatal problem with the truth-conditional approach; the behavior of an agent acting in accordance with it is utterly unintelligible. Consider the following case:

### ***Chuck Finn***

Chuck Finn, like a certain famous cousin of his, rationally believes a false moral view. He has become convinced that inflicting suffering on bad people is more important than preventing the suffering of good people. As a matter of fact (we will stipulate), preventing the suffering

of good people is more important than inflicting suffering on bad people. A mad philosopher has presented Chuck with three buttons, of which Chuck can press only one. Button A, Chuck has every reason to believe, will inflict X amount of suffering on a bad person. Button B, Chuck has every reason to believe, will prevent X amount of suffering to a good person. Button C, Chuck has every reason to believe, will drop a load of manure on his own head. However, thanks to some very complicated dependence relationships between his beliefs, coming to believe the correct moral view (which, it must be emphasized, Chuck cannot rationally do) would make Chuck confident that button C would prevent X amount of suffering to two good people (and would not have any relevant impact on his other beliefs).

There is an intelligible story about how Chuck might end up choosing button A, as standard N-subjective views will recommend. In this story, Chuck cares deeply about doing the right thing. He is rationally confident that pressing button A is the right thing to do, and so he presses button A.

There is an intelligible story about how Chuck might end up choosing button B, as standard D-subjective views will recommend. In this story, Chuck, like his cousin, deeply cares *de re* about the suffering of good people. He is rationally confident that pressing button B will prevent the suffering of a good person. Despite his judgment that morality requires him to press button A, his compassion wins out, and he presses button B.

I think there is no intelligible story whatsoever about how he chooses button C, as the truth-conditional D-subjective view will recommend. It could not be out of a concern to do what is right. It could not be out of a desire for any of the things that matter *de re*. If he became convinced of the

true theory, of course, we could understand his choosing C (as all three kinds of theories would then recommend). But given that he does not, and rationally could not, believe the true theory, this behavior has nothing at all going for it that could latch on to Chuck's motivations.

It is not just that there would be something objectionable about Chuck choosing button C. There are fair concerns one might raise about whether Chuck in the first story really has his heart in the right place, or whether Chuck in the second story is violating a rational requirement by acting against his better judgment, or whether that story is compatible with him being *guided* by the relevant norm. But pressing button C is in a way worse than vicious, and worse than irrational – it is inexplicable. There is not even any vice, or any familiar form of irrationality, which could make sense of that decision.

It is fine for a theory about the *objective* ought to make judgments about what a rational agent ought to do which do not match the behavior that would be intelligible from them. But describing behavior that would be intelligible from an agent seems like a minimal constraint on an account of the *subjective* ought. So I think the truth-conditional strategy will not work.

## **Normative Omniscience**

There is one last strategy available to the D-subjectivist that is worth considering. To explain it, we will have to reintroduce a distinction I set aside at the beginning of this paper.

There is a division among subjectivist theories between those that take the beliefs or credences relevant to what an agent ought to do to be their *actual* credences, and those that take the relevant beliefs or credences to be those that are *rational* or *justified* or *evidentially supported*. This

division cross-cuts the N-subjectivist/D-subjectivist distinction, and everything I have said so far can be applied in terms of actual or rational uncertainty interchangeably. The last response relies on taking a stand on this point.

What I will call the *normative omniscience strategy* rests on two claims:

- 1) What an agent subjectively ought to do depends on the credences that are rational (or justified, or evidentially supported) for them.
- 2) *Normative Omniscience*: It is never rational for an agent to be uncertain about the truth of a normative theory.

If these two claims are true, it would undermine the possibility of an *Angel's Present* style counterexample. It would be impossible for Heloise to be rationally uncertain between Theory A and Theory B, and therefore impossible for her to be rationally uncertain about whether there is an azzip or a buzzip in the box. Whatever her *actual* uncertainty is, she is rationally required to be certain that Theory A is true and therefore that there is an azzip in the box, and this view will correctly predict that she should open the angel's present (and, for parallel reasons, leave the devil's closed). Moreover, while this kind of view will have exactly the same implications in general as the truth-conditional strategy above, by denying that it could ever be rational for Chuck Finn to believe a false moral view it is able to insist that there is an intelligible story where he chooses button C – one where he first corrects his rational mistake.<sup>26</sup>

---

<sup>26</sup> This is a view seemingly favored by Hedden (2016) and Wedgwood (Unpublished).

Again, this view deviates slightly from the letter of D-subjective views, since it does not deny so much as trivialize the claim that normative beliefs play a role in what one ought to do. But just as D-subjectivists would balk at a view that matched theirs in letter but demanded that rational agents be omniscient about all descriptive truths, N-subjectivists will balk at the same move for normative truths.

How plausible is the normative omniscience strategy? One thing to note is that there are actually two versions of the first claim – a strong and a weak one. On the weak version, there is *some* subjective sense of ought which depends on rational credences. On the strong claim, *all* subjective senses of ought depend on rational credences. The strong claim is what is needed to defend full-on skepticism about N-subjective norms with this strategy, and I think it is significantly less plausible than the weak one. But I will not challenge either here. My main worries are with *Normative Omniscience*.

First of all, it clearly fares extremely badly at the level of commonsense intuition. Normative theory is *hard*. Getting to the truth about morality plausibly requires having access to roughly accurate moral intuitions, which not every rational agent may have. Moreover, our intuitions conflict, both internally and with those of others. Even if moral claims are in some sense a priori, they do not seem to be trivial, formal, or merely conceptual. So we do not normally judge agents who are not normatively omniscient irrational for that reason alone.

Second, the main arguments people have given in favor of claims in the vicinity of *Normative Omniscience* are not general enough to give the D-subjectivist what they need. Michael Titelbaum (2015) defends what he calls the *fixed-point thesis*: that it is irrational to have false beliefs about

rationality. Titelbaum's own assessment of the prima facie plausibility of his claim is "surprising – if not incredible". And yet it is weaker than normative omniscience, and therefore weaker than what the D-subjectivist requires, along at least two dimensions. First, normative omniscience concerns *all* normativity, not just rationality. Heloise's uncertainty, after all, might be about *moral value*. Second, the fixed point thesis says only that *false beliefs* about rationality are irrational, not that *credences less than certainty* about rationality are irrational. And this is not incidental – his argument relies on a principle of akrasia which has no obvious analogue true of credences. Heloise doesn't have any false beliefs. So Titelbaum's incredible claim is not nearly incredible enough.

Elizabeth Harman suggests that agents cannot be justified in believing (2011) or being certain of (2015) false moral views, at least when one doesn't believe on the basis of testimony. But her argument seems to be that false moral belief not based on testimony must be based on "fallacious, though seductive lines of thought" (2011, 461). But having misleading moral intuitions is not in the same category as committing a logical fallacy. Moreover, nothing about our cases demands that Heloise's credences not be based on testimony, and Harman's claims are once again too weak, since Heloise does not have a belief in any false moral view.

Finally, the more extreme our view of the demands of rationality, the less plausible it becomes that *all* interesting subjective senses of "ought" depend on rational, rather than actual, credences.

## Conclusion

I conclude, then, that the dependence problem ultimately supports the N-subjectivist over the D-subjectivist. I have not addressed the many challenges raised against N-subjective norms, as a full defense would require. But I have tried to give reason to hold out hope that a defense is possible. Decision under normative uncertainty is a research programme that is still in its infancy, and it would be a mistake to give up on it too soon, if there seem to be powerful reasons in its favor. At the same time, the argument suggests a note of caution to those developing an account of normative uncertainty: many principles attractive at first blush will fall apart if we do not attend to the ways our normative and descriptive beliefs are integrated.

## References

- Arpaly, N. 2002. 'Moral Worth.' *Journal of Philosophy* 99(5), 223-245.
- Carbonell, V. 2013. 'De Dicto Desires and Morality as Fetish.' *Philosophical Studies*. 163(2), 457-477.
- Driver, J. 2006. 'Autonomy and the Asymmetry Problem for Moral Expertise.' *Philosophical Studies* 128, 619-644.
- Enoch, D. 2014. 'A Defense of Moral Deference.' *Journal of Philosophy* 111, 1-30..
- Hopkins, R. 2007. 'What is Wrong With Moral Testimony?' *Philosophy and Phenomenological Research* 74, 611-634.
- Gracely, E. 1996. 'On the Noncomparability of Judgments Made by Different Ethical Theories.' *Metaphilosophy* 27, 327-332. *Journal of Philosophy* 111, 1-30.
- Gustafsson, J. and Torpman, O. 2014. 'In Defence of My Favourite Theory.' *Pacific Philosophical Quarterly* 95, 159-74.
- Guerrero, A. 2007. 'Don't Know Don't Kill: Moral Ignorance, Culpability, and Caution.' *Philosophical Studies* 136(1), 59-97.
- Harman, E. 2011. 'Does Moral Ignorance Exculpate?' *Ratio* 24, 443-68.
- Harman, E. 2015. 'The Irrelevance of Moral Uncertainty.' *Oxford Studies in Metaethics* 10.
- Hedden, B. 2012. 'Options and the Subjective Ought' *Philosophical Studies* 158(2), 343-360.

- Hedden, B. 2016. 'Does MITE Make Right?' *Oxford Studies in Metaethics* 11, 102-128.
- Lockhart, T. 2000. *Moral Uncertainty and its Consequences*. Oxford University Press.
- MacAskill, W. 2014. 'Normative Uncertainty.' Dissertation.
- MacAskill, W. 2016. 'Normative Uncertainty as a Voting Problem.' *Mind* 125(500), 967-1004.
- McGrath, S. 2009. 'The Puzzle of Pure Moral Deference.' *Philosophical Perspectives* 23, 321-344.
- Moller, D. 2011. 'Abortion and Moral Risk.' *Philosophy* 86(3), 425-443.
- Nissan-Rozen, I. 2015. 'Against Moral Hedging.' *Economics and Philosophy* 3, 1-21
- Parfit, D. Unpublished. 'What We Together Do.'
- Parfit, D. 2011. *On What Matters* Vol I. Oxford University Press.
- Ross, J. 2006. 'Rejecting Ethical Deflationism.' *Ethics* 116, 742-768.
- Sepielli, A. 2009. 'What to Do When You Don't Know What to Do.' *Oxford Studies in Metaethics* 4, 5-28.
- Sepielli A. 2012. 'Subjective Normativity and Action Guidance.' *Oxford Studies in Metaethics* 2, 45-73.
- Sepielli, A. 2013. 'Moral Uncertainty and the Principle of Equity Among Moral Theories.'  
*Philosophy and Phenomenological Research* 86, 580-9.
- Sepielli, A. 2016. 'Moral Uncertainty and Fetishistic Motivation.' *Philosophical Studies* 173 (11), 2951-2968

- Schüssler, R. 2005. 'The Anatomy of Probabilism.' In Kraye, J. and Saarinen, R. (Eds) *Moral Philosophy on the Threshold of Modernity*
- Smith, H. 2010. 'Subjective Rightness.' *Social Philosophy and Policy* 27, 64-110.
- Smith, M. 1994. *The Moral Problem*. Wiley-Blackwell.
- Smith M. 2013. 'On the Nature and Significance of the Distinction between Thick and Thin Ethical Concepts.' In Kirchin, S. (Ed.) *Thick Concepts*. Oxford University Press.
- Titelbaum, M. 2015. 'Rationality's Fixed Point.' *Oxford Studies in Metaethics* 5, 253-294.
- Way, J. and Whiting, D. 2017. 'Perspectivism and the Argument from Guidance.' *Ethical Theory and Moral Practice* 20(2), 361-374
- Weatherson, B. 2013. 'Running Risks Morally.' *Philosophical Studies* 167, 1-23.
- Wedgwood, Ralph. Unpublished. 'Moral Disagreement and Inexcusable Irrationality.'
- Williamson, T. 2000. *Knowledge and its Limits*. Oxford University Press.
- Wolf, S. 1982. 'Moral Saints.' *Journal of Philosophy*, 79(8), 419–439.