Who's Afraid of Normative Externalism?

ZOË JOHNSON KING
University of Michigan

To appear in *Meaning, Decision, and Norms: Themes from the Work of Allan Gibbard*, Billy Dunaway and David Plunkett (eds). Maize Books, an imprint of Michigan publishing.

Note to the reader: This is a draft. Nonetheless, please feel free to share and cite it with gleeful abandon.

1. Moral uncertainty

This paper is about what someone should do when she is not only unsure what first-order moral theory is true, but also uncertain about whether this moral uncertainty is itself morally relevant.

I am not sure what first-order moral theory is true. Of course, I have some sense of which things matter morally. But I don't take myself to have figured out precisely how many and which things matter morally, and the relative degrees to which each of these things matters morally, and the metaphysical relationships (causal or constitutive) that obtain between them. But the true first-order moral theory, fully spelled out, will settle all these matters. So, there are a great many possible moral theories that are consistent with my rough sense of which things matter morally. I am unsure which of these theories is true.

This paper addresses the question of how someone like me – who is not sure what first-order moral theory is true – should act. The question has been a hotbed of philosophical activity in recent years. One central insight that emerges from the extant literature is that we can use the tools of decision theory to discuss and generate possible answers to this question. We can do this because we can construe my moral uncertainty as a kind of uncertainty about what the world is like: it is uncertainty about what the world is like *morally*. With this construal in hand, we can model such uncertainty using the table familiar to us all from undergraduate courses on decision theory. Simplified drastically, it looks something like this:

	T1 (0.6)	**T2 (0.2)**	**T3 (0.2)**	**Expected Value**
A1	10	2	4	7.2
A2	4	8	6	5.2
A3	2	6	50	12.4

Here the columns represent each moral theory in which the agent has some positive credence. (I have some positive credence in far more than three moral theories, but, for ease of exposition in drawing the tables,

I'll just discuss an imaginary possible agent who divides her credence exhaustively between exactly three theories.) The rows represent acts that she might perform. And each box contains a number representing the moral value of performing the corresponding act according to the corresponding theory. For example, the table shows that theory 1 says that performing act 1 would have 10 units of moral value.

The phrase "units of moral value" here might raise a skeptical eyebrow. It is difficult to explain what this can mean in a way that remains neutral between first-order moral theories. And it is even more difficult to devise a way to render different moral theories' assessments of acts commensurable, such that they can all be placed on a single scale – which is necessary for the numbers in the table to even make sense. These are some major problems that have been discussed at length in the literature on moral uncertainty.[1] I do not think that these problems can be easily solved. But I will set them aside for the purposes of this paper, since I am interested in an entirely different problem. So, for present purposes, I will proceed as if we have found a clear way to make inter-theoretic comparisons of moral value.

There are lots of different possible decision-rules saying how someone should act given the kind of moral uncertainty depicted in the table above. These decision-rules for moral uncertainty correspond to rules in traditional decision theory saying how an agent should act given uncertainty about mundane empirical matters, like whether it will rain, or whether an egg that she is considering adding to her omelet is rotten.[2] The most widely-discussed such rule is this:

> Maximize Expected Objective Value (MAXEOV): Morally uncertain agents should *maximize expected objective value* – i.e., perform the act with highest expected objective moral value.

MAXEOV is explicitly inspired by the traditional decision theorist's refrain that an uncertain agent ought to maximize expected value. According to MAXEOV, a *morally* uncertain agent ought to maximize expected *moral* value. This is accomplished by calculating the expected moral value of each act – a weighted sum of the amounts of moral value that it has according to each moral theory in which the agent has some positive credence, weighted by the agent's credence in each of the corresponding theories – and then performing the act with the highest total expected moral value. In the case represented in the table above, this rule says that the uncertain agent should perform A3. Defenses of MAXEOV appear in Lockhart (2000), Sepielli (2009), and Enoch (2014), among others.

Myriad other possible decision-rules for moral uncertainty would mimic other principles from traditional decision theory. For example, there is a possible decision-rule stating that morally uncertain agents should use maximin, or maximax, or minimax regret, or that they should maximize risk-weighted expected moral value in the style of Lara Buchak (2013), et cetera. Thankfully, these rules do not yet all have advocates in the literature on moral uncertainty.

Another widely-discussed rule is this:

---

[1] Some important contributions to this literature are Gracely (1996), Lockhart (2000), Ross (2006), Sepielli (2009, 2013a, 2013b), Hedden (2016), and Hicks (forthcoming). N.B. It might be possible to solve the problem of intertheoretic value comparisons by "consequentializing" all moral theories (on which see e.g. Portmore 2007, Dreier 2011, Hurley 2013). But I suspect that even this would not enable us to place the evaluations of all moral theories on a single scale.

[2] For the omelet example, and a classic statement of traditional decision theory, see Savage (1954).

> My Favorite Theory (MFT): A morally uncertain agent should just consult the moral theory in which she has the highest credence, and should perform the act that is ranked best by this theory, ignoring all the others.

MFT holds that a morally uncertain agent should perform the act recommended by the moral theory in which she has the highest credence. In the case represented in the table above, this rule says that the agent should perform A1. This approach is attractive because it completely avoids having to somehow compare the degrees of moral value that acts have according to different moral theories – a clear advantage of MFT over MAXEOV, and over all the other possible decision-theoretic rules just mentioned. Nonetheless, MFT is not very popular, because it generates some highly counterintuitive results. Most notably, if an agent has a plurality of credence in a theory that ranks action A1 the highest, but all of the other theories in which she has some positive credence hold that A1 would be a serious moral disaster, MFT cheerfully recommends A1. This holds even if there is a "safe" alternative act, A2, which all theories agree would be very good. This seems like a bad result.[3] Still, MFT has received some sophisticated recent defenses, notably from Johann Gustafsson and Olle Torpman (2014).

One last decision-rule will play a starring role in this paper. Here it is:

> Do The Right Thing (DTRT): Morally uncertain agents should do whatever is in fact the right thing to do, regardless of their credences.

DTRT is very different from the other decision-rules. According to DTRT, an agent's moral uncertainty is morally irrelevant. What a morally uncertain agent should do is the same as what someone who is certain of the true first-order moral theory should do, which is also the same as what someone who is certain of a false moral theory should do: she should do whatever is in fact required by the theory that is in fact the true first-order moral theory. DTRT has been defended by Elizabeth Harman (2015) and Brian Weatherson (2013). It follows from the approach termed "Normative Externalism" and discussed by Weatherson in his (*ms*), according to which the true moral norms apply to agents regardless of our epistemic states with respect to these norms, and thus are in an important sense "external" to us. On this view, an agent faced with the kind of moral uncertainty depicted in the table above should ignore the table entirely and do whatever is in fact morally right. Moreover, on this view, debates between "internalist" views like MAXEOV and MFT about how agents' credences in moral theories make a difference to what she should do are totally wrongheaded. They are like debates about what blood type R2D2 is. R2D2 is a robot, so he doesn't have any blood. Thus an inquiry into R2D2's blood type would be a total waste of time. Similarly, according to DTRT, agents' credences in moral theories make no difference to what they should do. Thus an inquiry into what kind of difference they make is a total waste of time.

[3] One could render this result less implausible by adding a restriction to MFT that requires uncertain agents to follow the dictates of their favorite theory only if their credence in this theory is above 0.5 – this would avoid recommending actions that the agent thinks are more likely than not to be a serious moral disaster. But this version of MFT still entails that someone should perform A1 if she has credence 0.5000001 in a theory that holds that A1 is slightly better than a safe alternative, and credence 0.4999999 in a theory that holds that A1 would be a serious moral disaster. This verdict still seems unacceptably morally risky to many people.

I am inclined to think that something like MAXEOV must be correct. But, for a long time, I was troubled by the possibility that Normative Externalism might be true. This is partly because, as just discussed, if this position is true then the debates about moral uncertainty that seem to me most interesting and important are in fact a total waste of time. But that is not all. I was also troubled by the fact that, while I am inclined to think that something like MAXEOV must be correct, I certainly do not assign credence 1 to its correctness. The other decision-rules just mentioned all have arguments in their favor that are good enough for me to assign *some* positive credence to each of them being correct. So, in addition to being uncertain about what first-order moral theory is true, I am in a state of *higher-order uncertainty*: uncertainty about what a morally uncertain agent (such as myself) should do. As an internalist, I am inclined to think that the higher-order theories in which I have some positive credence should each be taken into account in some way when I decide what to do. However, it turns out to be quite difficult to know how to take account of the possibility that Normative Externalism is true – and that I am wasting my time, because the correct decision-rule for morally uncertain agents is just DTRT – within a MAXEOV framework. That is what this paper is about.

2. Higher-order uncertainty

As just discussed, I am not sure which theory about how people should act when they are morally uncertain is true. But here is one thing of which I am fairly certain: I am fairly certain that the term "should" in the phrase "how people should act when they are morally uncertain" is a *moral* "should". I find it hard to see what other "should" it could be – it does not, for instance, seem to be prudential, or epistemic, or aesthetic, given that the arguments for and against MAXEOV and its rivals trade on what seem to be moral intuitions rather than intuitions pertaining to any of these other normative domains. Nor does it seem plausible that the "should" in question is an all-things-considered "should", akin to the Gibbardian "primitive ought" of practical deliberation (2003, p.**). The primitive "ought" is the one in the final "ought I to φ?" question that an agent asks herself before acting; it is the one such that it is an incoherent combination of attitudes to judge that one ought to φ and yet not intend to φ. Here is Gibbard (*ibid.*, p.153):

> [T]o think something the thing to do is to plan to do it. To think, for instance, that the thing now to do is to defy the bully who torments me is to plan to defy him. And planning right now to defy him right now, to do it at this very moment, amounts to setting out to do it. My theory thus yields internalism in a strong form: if I think that something is now the thing to do, then I do it. My hypothesis about ordinary *ought* judgments is that they are judgments of what to do, of what is the thing to do. I don't, then, think that I ought right now to defy the bully unless I do defy him. If I fail to defy him, then as a matter of the very concept of *ought*, I don't believe I ought to.

This does not seem to be the case for the "should" in claims of the form "the true principle governing how people should act when they are morally uncertain entails that, in light of my moral uncertainty, I should φ". One need not plan right now to φ in order to count as accepting a claim of this form. On the contrary, it seems perfectly intelligible to make such a claim and then nevertheless ask, "…but shall I φ?". (And it seems perfectly intelligible to answer "no, I shan't", if there are weighty prudential or epistemic or aesthetic considerations count against φ-ing.) The fact that such questions always remain open tells us that it is not a conceptual truth that the verdicts of principles governing morally uncertain agents always settle the

practical question of what to do. So, the "should" in principles governing morally uncertain agents is not akin to the Gibbardian primitive "ought". The "should" here is not a primitive "should". Given that the arguments for and against MAXEOV and its rivals trade on what seem to be moral intuitions, then, I take the "should" in such principles to be a *moral* "should".

This means that the decision-rules surveyed above are all partially self-referential. They apply to morally uncertain agents, and they are themselves (putative) moral principles. So, each one applies, in part, to uncertainty about whether it is indeed the correct moral principle governing morally uncertain agents.

I am therefore in a state of *higher-order moral uncertainty*. Not only am I uncertain about which first-order moral theory is true, but I am also uncertain about what I (morally) should do in light of this fact.

We can draw up a decision-theoretic table representing the higher-order moral uncertainty of people like me. It would look something like this:

	MAXEOV (0.95)	**MFT (0.025)**	**DTRT (0.025)**	**Expected Value**
A1	7.2	10	?	?
A2	5.2	4	?	?
A3	12.4	2	?	?

Here, again, the columns represent each (higher-order) moral theory in which the agent has some credence, the rows represent her available acts, and each box contains a number representing the moral value of performing the corresponding act according to the corresponding theory. (Again, for ease of exposition I will discuss an imaginary agent who divides her credence exhaustively between exactly three higher-order theories, though I myself have some positive credence in more theories than just these.)

Some of the columns in this table are easier to fill out than others. It is easy to fill out the MAXEOV column; the moral value of each act according to MAXEOV is just the expected objective value of the act. This is fairly easy to calculate using one's first-order decision-table. It is also easy to fill out MFT's column; the value of each act according to MFT is just the value that it has according to the first-order moral theory in which one has the highest credence, which is fairly easy to see from one's first-order decision-table.[4] But it is impossible for a morally uncertain agent such as myself to fill out the DTRT column of her higher-order decision-table. I can describe the values that this column should contain: it should contain the moral value

[4] This overstates things a bit. Just as I do not know what the true first-order moral theory is, I do not know *precisely* what credence I have in various first-order moral theories, since my credences are not completely luminous to me. So, technically, the numbers in the MAXEOV column should be narrow ranges of the sort that I discuss when describing the upper-and-lower-bounds strategy later in this section. Nonetheless, in the vast majority of cases, the ranges will be sufficiently narrow to generate a clear ordering over my available acts. And, if I do know precisely what my first-order credences are, then I can complete the MAXEOV column precisely. By contrast, in *all* cases, uncertain agents will be unable to fill out the DTRT column of their decision-table with actual numbers, and – as we will see below – the ranges will frequently be too broad to generate a clear ordering over my available acts. (The solution that I propose in §3 could be used to accommodate uncertainty about the values of acts according to MAXEOV and MFT, as well as DTRT.) Thanks to Billy Dunaway for pointing all of this out to me.

of each act according to the *true* first-order moral theory. But I do not know what the true first-order moral theory is – therein lies my moral uncertainty. Since I do not know what the true first-order moral theory is, I do not know how much moral value each of my available acts has according to the true theory. So, I am not in a position to fill out the DTRT column of my higher-order decision-table with actual numbers.

These observations troubled me for a long time. I was troubled because I thought that they highlighted a problem for MAXEOV. Like all theories about how morally uncertain agents should act, MAXEOV is partly self-referential – it applies, in part, to uncertainty about whether it is indeed the correct principle governing morally uncertain agents. But MAXEOV simply does not work when applied to the higher-order decision-tables of agents, like me, who have some positive credence in Normative Externalism. I cannot fill out the DTRT column of my higher-order decision-table, as I am unsure what first-order moral theory is true. But this means that I cannot apply a maximizing algorithm. I cannot calculate weighted sums of the values in each row of my decision-table if one box in each row contains a question mark instead of a number. The maximizing algorithm crashes; it needs precise numbers in each box as input. My credence in DTRT thus has a devastating impact on the higher-order application of MAXEOV. Indeed, the problem is not just that I am no longer able to identify a choice set among my available acts. It's much worse than that. I am unable to rank the acts at all, nor even to assess the higher-order expected objective value of any one of them.

This does not mean that MAXEOV cannot handle *any* uncertainty as to whether it is the correct principle governing morally uncertain agents. On the contrary, it is possible to calculate the higher-order expected objective value of one's available acts if the higher-order theories in which one has some credence are all internalist. If all higher-order theories in which I have some positive credence take the value of my available acts to depend in some way on the contents of my first-order decision-table, then I am okay. For example, an agent who divided her credence entirely between MAXEOV and MFT would have no trouble applying MAXEOV to her higher-order decision-table, since these theories both assign precise numbers representing the value of the agent's available acts in light of her first-order uncertainty. (The numbers are the first-order expected objective value of each act, and the value of each act according to the first-order theory in which the agent has the highest credence, respectively.) It is only when someone has some positive credence in an externalist principle, like DTRT, that she cannot apply MAXEOV to her higher-order uncertainty.

Nonetheless, I found this troubling. I think it is plausible that the true theory about how morally uncertain agents should act should apply to both first-order and higher-order uncertainty. This would offer an attractively unified picture of normative reality. Certainly, if one theory were true for first-order moral uncertainty and a totally different theory were true for higher-order uncertainty, this would be a striking fact that calls out for explanation.[5] And I cannot think of a satisfying explanation for why MAXEOV should be the true theory for first-order uncertainty but not for higher-order uncertainty – certainly, "the math doesn't work" does not seem like a satisfying explanation. But I also think it is plausible that the true theory about how morally uncertain agents should act should be able to withstand agents' credence in alternative theories. After all, the point of these theories is to guide the uncertain. It would therefore be ironic if such a theory were unable to identify a choice set, nor to rank the available acts, nor even to assess the value of

[5] Here I deviate from Sepielli (2013b), who suggests that each agent simply resolve her uncertainty at level N by appeal to whichever theory she is most confident of at level N+1, if there is one. I am not so laissez-faire. I think it would be strange for different theories to be true at different levels. So, if someone takes this to be the case, I hope that she has a satisfying explanation of why the levels are different.

any individual act, for all uncertain agents who have some positive credence in a particular alternative theory. And this seems to be the case here.

One possibility for the friend of MAXEOV is to dismiss DTRT out of hand. Perhaps we should exclude agents' credence in DTRT from their higher-order decision-tables, or revise our maximizing algorithm to ignore it (renormalizing the agents' credence in other theories accordingly). It has been suggested to me that maximizers may defend this move by way of an analogy between DTRT and moral nihilism. Some philosophers hold that, in a first-order decision-table, it is permissible to ignore the possibility that moral nihilism is true and that nothing has any moral value whatsoever. After all (so the argument goes), decision theorists typically ignore states of the world in which all acts result in the same outcome, since these outcomes "cancel out" when we assess the expected utility of each act. But nihilism is the possibility that nothing has any moral value. So, it is a possibility according to which all acts result in the same outcome, namely no moral value. Thus, it can be safely ignored.[6] The defender of MAXEOV might try suggesting that DTRT is relevantly analogous to moral nihilism. Nihilists answer the question, "Which first-order moral theory is true?" by saying, "None of them". And, similarly, normative externalists answer the question, "Which decision-theoretic principle governs morally uncertain agents?" by saying, "None of them". Thus, like first-order moral nihilism, perhaps normative externalism can be safely ignored.[7]

I think that this is the wrong way to think about DTRT. The possibility that Normative Externalism is true is not the possibility that all acts have the same moral value. Rather, it is the possibility that the first-order moral uncertainty of agents like me is morally irrelevant, and that what we should do is determined by the true moral theory alone, rather than by our credences in moral theories. So, as long as the true moral theory holds that different acts have different values, DTRT holds this too. This is therefore not a principle according to which all acts are equally good. It is rather a principle such that agents who are first-order uncertain cannot tell which acts it evaluates as better than others, or by how much.

Relatedly, it is a mistake to see DTRT as a principle according to which there is no answer to the question of what someone should do in light of her moral uncertainty. On the contrary, DTRT does issue verdicts as to what someone should do in light of her moral uncertainty – but they are the same verdicts as those that the principle issues for agents who have credence 1 or credence 0 in the true first-order moral theory. To put the same point another way: on this view, as on MAXEOV or MFT, there is a function from an agent's epistemic state and the moral facts to a choice set. But the striking thing about DTRT is that this function's output stays the same as the agent's credences in moral propositions change, so long the moral facts (and morally relevant non-moral facts) are held fixed.[8] To emphasize: the trouble with DTRT is not that it fails to evaluate or to compare the available acts of first-order uncertain agents, but rather that first-order uncertain agents cannot tell what these evaluations and comparisons are.

---

[6] This suggestion comes from Ross (2006), and is criticized in MacAskill (2013). The idea that a state of the world in which all acts yield the same outcome can be ignored is the idea that motivates the Allais paradox; see Allais (1953).

[7] Thanks to Andrew Sepielli for raising this possibility to me in conversation.

[8] Elizabeth Harman (2015) makes this point especially clearly. Defending a version of Normative Externalism that she calls "Actualism", Harman repeatedly emphasizes that she offers Actualism as a theory about what uncertain agents *subjectively* ought to do (p.58). As she puts it, "Actualism is a proposed answer to the very same question the [defenders of MAXEOV] are interested in, namely: how should a person act, taking into account her beliefs and credences (including her moral beliefs and credences), given that one sometimes must act while experiencing moral uncertainty?" (p.70). Since this is Harman's own characterization of her view, I think it is a fair characterization of the view.

So, simply dismissing Normative Externalism out of hand seems unmotivated. Another option is for me to make an educated guess as to what goes in the DTRT column of my higher-order decision-table. I can use my first-order decision-table to think about what the values in the "DTRT" column of the higher-order decision-table might, for all I know, turn out to be. Granted, I cannot say confidently what the values in this column are, since I am not sure what first-order moral theory is true. But I do have a partition of possible first-order theories between which I divide my credence. And each of these theories assigns a precise degree of objective moral value to each of my available acts. (Recall that we are assuming, for present purposes, that there is a way to make inter-theoretic comparisons of value such that a first-order decision-table can intelligibly be completed.) So, I can place *upper and lower bounds* on the degree of moral value that each of my available acts might turn out to have. The upper bound is the highest degree of objective value assigned to the act by a moral theory in which I have some positive credence, and the lower bound is the lowest such degree of value. Although I am not sure precisely what degree of moral value the act has, I am sure that it falls somewhere within this range.[9]

Given the first-order decision-table presented at the beginning of §1, this strategy will yield a higher-order decision-table that looks like this:

	MAXEOV (0.95)	**MFT (0.025)**	**DTRT (0.025)**	**Expected Value**
A1	7.2	10	[2 – 10]	[7.14 – 7.34]
A2	5.2	4	[4 – 8]	[5.14 – 5.24]
A3	12.4	2	[2 – 50]	[11.88 – 13.08]

This is progress. With a range within which the value of each act according to DTRT must fall, it is possible to identify a range within which the act's higher-order expected objective value according to MAXEOV must fall. This is accomplished by calculating two weighted sums, one that takes the value of the act for DTRT to be the lowest value in the range, and one that takes it to be the highest in the range. Since the agent is sure that the value of the act according to DTRT falls somewhere within the range, she can establish that a higher-order application of MAXEOV would yield a degree of higher-order expected objective value somewhere within the bounds specified in the final column. Moreover, in some cases – like the one above – this provides enough information to rank some of the acts, since some acts' ranges of possible degrees of expected objective value can be wholly above or below some of the others. In the table above, the ranges are sufficient to totally order the acts (A3 > A1 > A2), and to identify a singleton choice set ({A3}).

However, this ordering is *only* possible when one act's range of possible degrees of expected objective value is wholly above another's. And this will not always be the case. Indeed, it can very easily fail to hold. For example, if the agent's first-order decision-table is exactly as presented at the beginning of §1 except that the moral value of A3 according to the third theory in which the agent has some credence is 25 rather than 50, then her higher-order decision-table will be as follows:

[9] Thanks to Brian Weatherson for raising this possibility to me in conversation.

	MAXEOV (0.95)	MFT (0.025)	DTRT (0.025)	Expected Value
A1	7.2	10	[2 – 10]	[7.14 – 7.34]
A2	5.2	4	[4 – 8]	[5.14 – 5.24]
A3	7.4	2	[2 – 25]	[7.13 – 7.705]

Now it is no longer possible to rank acts A1 and A3, since A1's range of potential expected objective value is wholly contained within A3's. There is no answer to the question, "Which is greater, *between 7.14 and 7.34* or *between 7.13 and 7.705*?" – the question is ill-formed. So, the uncertain agent cannot order these actions, nor can she identify a choice set from among them. Of course, the defender of MAXEOV could propose a rule for what to do in such cases: the rule might be maximax, or maximin, or a rule telling us to choose the act whose range of potential expected objective value has the highest midpoint, or whatever. But the very fact that such a wide range of rules like this are available makes the choice of any particular one seem arbitrary and unmotivated. So, in the absence of a positive argument for any particular rule, I suggest that we try a different tack.

3. A solution

In this section, I will present what I think is the best strategy for defenders of MAXEOV to take account of the credence that agents like me have in DTRT.

The strategy begins with the same observation as the upper-and-lower-bounds strategy above: an uncertain agent like me has a first-order decision-table displaying a partition of possible first-order theories between which she divides her credence, each of which assigns a precise degree of objective value to each available act. As we have just seen, this means that we can identify upper and lower bounds to what might, for all the agent knows, be the true moral value of each act – and thus its value according to DTRT. My preferred strategy notes that we can do even better. For the agent also assigns a precise credence to each of these first-order theories' being true. She is still not sure which of them is true, and thus which assigns the values to acts that DTRT endorses. But, for each theory, she has some precise credence that this theory specifies the values of her acts according to DTRT. DTRT is the decision-rule of Normative Externalism, which instructs uncertain agents to do whatever is in fact required of them by the true first-order moral theory. So, the agent's credence that T1 assigns the values to acts that DTRT endorses is just her credence that T1 is in fact the true first-order moral theory. And so similarly for all other first-order moral theories in which the agent has some credence; her degree of confidence that DTRT evaluates and compares acts in accordance with the dictates of this theory should just be her credence that this theory is true.

With this in in mind, here is the strategy. The uncertain agent should think of the possibility that Normative Externalism is true, not as one possibility, but as many possibilities – as many as there are first-order theories to which she assigns some positive credence. There is no way for Normative Externalism to be true without some particular first-order theory being true. And the agent has at her disposal a partition of possible first-order moral theories that she thinks might be true. So, she should divide the credence assigned to DTRT in her higher-order decision-table between these possibilities. That is to say: rather than

including one single "DTRT" column in the table, she should include as many columns as there are first-order moral theories to which she assigns some positive credence, with these columns representing the combined possibilities that DTRT is the correct rule for uncertain agents *and* T1 is the true first-order moral theory, that DTRT is the correct rule for uncertain agents *and* T2 is the true first-order moral theory, and so on. Her credence in each combined possibility can be easily calculated; it is the product of her credence in DTRT and her credence in the first-order theory.[10]

The result will be a higher-order decision-table that looks something like this:

	MAXEOV (0.95)	**MFT (0.025)**	**DTRT & T1 (0.015)**	**DTRT & T2 (0.005)**	**DTRT & T3 (0.005)**	**Expected Value**
A1	7.2	10	10	2	4	7.27
A2	5.2	4	4	8	6	5.17
A3	12.4	2	2	6	50	11.88

Call this strategy "the repartitioning strategy".

As is clear from this table, the repartitioning strategy has some merits. In contrast to the upper-and-lower-bounds strategy, the repartitioning strategy will always enable the uncertain agent to complete her higher-order decision-table with precise numbers in every box. This ensures that it is always possible to apply a maximizing algorithm to the higher-order decision-table, notwithstanding the agent's credence in DTRT. Moreover, the agent's credence in DTRT is not simply being set aside or ignored, as in the strategy that treats Normative Externalism like moral nihilism. On the repartitioning strategy, the agent's credence in DTRT is fully taken into account, by being repartitioned into her credences in the combined possibilities that serve as inputs to the maximizing algorithm. Moreover, the repartitioning strategy enables uncertain agents to calculate the higher-order expected objective value of each of their available acts in a manner that will always yield a total ordering (in the example above, A3 > A1 > A2) and a choice set (in this case, {A3}). The repartitioning strategy thus solves MAXEOV's problems. So, I think it is the best strategy for defenders of MAXEOV to take account of the credence that agents like me have in DTRT.

The repartitioning strategy has a surprising implication. We can begin to grasp this by considering what would happen if the agent whose uncertainty is depicted in the table above were to become somewhat less confident in MAXEOV and somewhat more confident in DTRT. Suppose, for instance, that she shifts from credence 0.95 in MAXEOV and credence 0.025 in DTRT to credence 0.775 in MAXEOV and credence 0.2 in DTRT. Her higher-order decision-table would change as follows:

[10] This assumes that DTRT is probabilistically independent of each first-order theory in which the agent has some credence. I take this to be a reasonable assumption, since the existing arguments for and against Normative Externalism are all supposed to hold independently of which first-order theory is true. Nonetheless, I confess to being unsure as to whether it really is plausible that DTRT is probabilistically independent of all first-order theories. For instance, maybe DTRT is more plausible if objective consequentialism is the true first-order theory than if subjective consequentialism is the true first-order theory, just because this would be a less weird combination of lower-order and higher-order facts about the normative relevance of an agent's credences.

	MAXEOV (0.775)	MFT (0.025)	DTRT & T1 (0.12)	DTRT & T2 (0.04)	DTRT & T3 (0.04)	Expected Value
A1	7.2	10	10	2	4	7.27
A2	5.2	4	4	8	6	5.17
A3	12.4	2	2	6	50	11.88

Now suppose that she shifts to credence 0.525 in MAXEOV and credence 0.45 in DTRT. Her table would then look like this:

	MAXEOV (0.525)	MFT (0.025)	DTRT & T1 (0.27)	DTRT & T2 (0.09)	DTRT & T3 (0.09)	Expected Value
A1	7.2	10	10	2	4	7.27
A2	5.2	4	4	8	6	5.17
A3	12.4	2	2	6	50	11.88

Finally, suppose that she assigns credence 0 to MAXEOV and credence 0.975 to DTRT. Her higher-order decision-table would then look like this:

	MFT (0.025)	DTRT & T1 (0.585)	DTRT & T2 (0.195)	DTRT & T3 (0.195)	Expected Value
A1	10	10	2	4	7.27
A2	4	4	8	6	5.17
A3	2	2	6	50	11.88

The general point is as follows. On the repartitioning strategy, the uncertain agent can shift any amount of credence back and forth between MAXEOV and DTRT without this making a difference to her choice set, or to the ordering of her available acts, or even to the higher-order expected objective value of any act. This holds in full generality; no matter which first-order theories the agent has some credence in, and no matter which other higher-order theories (besides MAXEOV and DTRT) she has some credence in, the result will still hold that the agent can shift an arbitrary amount of credence from MAXEOV to DTRT and back again without anything in the "Total" column of her higher-order decision-table changing at all.

This is not just some crazy coincidence. It happens because the part of the higher-order calculation whereby MAXEOV takes account of my credence in DTRT, on the repartitioning strategy, is much like a simple first-order calculation of expected objective value. I take each first-order theory in which I have some credence,

take the objective value of the act according to that theory, and then multiply this number by the product of my credence in the first-order theory and my credence in DTRT. I do this for each first-order theory in which I have some credence. I then sum the results. This part of my calculation of the higher-order expected objective value of the act is mathematically equivalent to simply taking its first-order expected objective value and multiplying it by my credence in DTRT. To verify that this is so, compare these equations:

(i) $(a * m * q) + (b * n * q) + (c * o * q)$
(ii) $((a * m) + (b * n) + (c * o)) * q$

Here the variables *a, b*, and *c* stand for my credences in T1, T2, and T3, while *m, n,* and *o* stand for the value of an act according to each theory (respectively), and *q* stands for my credence in DTRT. The first equation is the part of my higher-order calculation that takes account of my credence in DTRT, on the repartitioning strategy. The second is a first-order calculation of expected objective value, multiplied by my credence in DTRT. These equations are straightforwardly equivalent. So, on the repartitioning strategy, the part of my higher-order calculation that takes account of my credence in DTRT is equivalent to simply taking the act's first-order expected objective value and multiplying it by my credence in DTRT. Now, consider the part of my higher-order calculation that takes account of my credence in MAXEOV. This part involves taking the act's first-order expected objective value – the value of the act according to MAXEOV – and multiplying it by my credence in MAXEOV. It is no wonder, then, that the uncertain agent employing a repartitioning strategy can shift any amount of credence back and forth between MAXEOV and DTRT without the higher-order expected objective value of any act changing at all. Whatever credence she assigns to MAXEOV and to DTRT, she will then effectively just multiply each of these credences by the first-order expected objective value of the act and sum the results. Shifting credence back and forth between MAXEOV and DTRT will not change this sum.

This is a curious result. The repartitioning strategy ensures that, in all cases, my credence in DTRT may as well be credence in MAXEOV. So it turns out that, after all, I *can* safely ignore the possibility that Normative Externalism is true – just as the analogy with nihilism suggested that I do. But I am no longer ignoring this possibility based on a misunderstanding of Normative Externalism, as the analogy with nihilism did. On the contrary, I am now ignoring this possibility on principled grounds. I am ignoring it because the most promising way to accommodate it in my calculations of expected objective value implies that my credence in this theory may as well just be credence in MAXEOV. So, I can acknowledge that Normative Externalism might be true, without this acknowledgement ever making any difference to what I should do, nor to the ordering over my available acts, nor even to the higher-order expected objective value of any act. In short, on this strategy, my acknowledging that Normative Externalism might be true doesn't change anything.

This might make the repartitioning strategy seem unfair to Normative Externalism. Why accommodate this theoretical possibility in a way that ensures that it never makes a difference to what I should do? But that is too quick. The points above all apply only to my shifting credence between DTRT *and MAXEOV*. If there are other higher-order theories in which I have some credence – such as MFT – then shifting some credence from *these theories* to DTRT does change the higher-order expected objective value of my available acts. And this can make a difference to what I should do, depending on how much credence shifts and on what my first-order decision-table looked like. So, my credence in DTRT does matter, sort of.

I say "sort of" because shifting a certain amount of credence from another theory to DTRT will always have the same impact on an act's higher-order expected objective value as shifting the same amount of credence from this other theory to MAXEOV. It remains the case that my credence in DTRT may as well be credence in MAXEOV. But my credence in DTRT does matter, sort of, because it matters how much credence I assign to DTRT *rather than MFT*, and rather than any other higher-order theory, although it does not matter how much credence I assign to DTRT rather than MAXEOV.

There is another way in which the repartitioning strategy genuinely is unfair to Normative Externalism. It is unfair because Normative Externalism holds that uncertain agents should perform the act required by what is in fact the true first-order moral theory, even if they have *no* credence in the true theory. The DTRT decision-rule evaluates my available acts according to the true first-order moral theory, whether or not this theory appears in one of the columns of my first-order decision table. So, if I were to fully acknowledge the possibility that Normative Externalism is true, I would have to take account of the possibility that the true moral theory is one to which I assign credence zero. This is not a possibility that the repartitioning strategy is able to accommodate.

But this does not seem such a terrible result. After all, we are considering the possibility that the true moral theory is one to which I currently assign credence zero. By stipulation, my credence in this possibility must be zero. And possibilities to which I assign credence zero are not usually ones that I am required to take into consideration when deciding what to do. For example, the possibility that my typing this sentence will summon an evil square circle that destroys the universe does not lead me to hesitate in typing this sentence, notwithstanding the fact that it would be very bad if the universe was destroyed, as I assign this possibility credence zero. Within a decision-theoretic approach to thinking about moral uncertainty, there is no way to accommodate hypotheses in which I have no credence. But this is simply part and parcel of the decision-theoretic approach to thinking about how uncertain agents should act. So, while the repartitioning strategy cannot capture everything involved in the possibility that Normative Externalism is true, it seems to be the best that MAXEOV can do.

This last point should be appealing to those attracted to a Gibbardian view of decision-making as a matter of deciding what to do in light of one's plan-laden beliefs. (Though, to be clear, I think that the strategy I have described is the best strategy for defenders of MAXEOV regardless of how they feel about Gibbard.) On a Gibbardian view, the possibility that Normative Externalism is true is not the sort of possibility that we can plan for. We plan what to do in light of possible subjective states that we might be in, where these are states that result from conditionalizing a set of initial credences on a set of total evidence. This means that we do not plan for certain sets of objective facts to hold – including the fact that a certain first-order moral theory is true – but for us to acquire evidence indicating that there is a certain probability that they hold. Here is Gibbard (2003, p.57):

> A plan must be one the agent can carry out with the information at her disposal. "Buy low, sell high,", for example, is not a plan, if one has no way of telling whether prices have reached their peaks or their troughs. An *occasion*, as I have characterized it, contains much that the agent has no way of knowing, but one's plans must respond to features of the occasion available to the agent. Alternatives must be subjectively characterized, so that the

> same alternatives are available on subjectively equivalent occasions. And a plan must permit the same alternatives on subjectively equivalent occasions.

If Gibbard is right that a plan must permit the same alternatives on subjectively equivalent occasions, then a morally uncertain agent cannot plan to perform different acts depending on what first-order moral theory is in fact true. And, if one's plans must respond to features of the occasion available to the agent, then they cannot respond to the objective moral facts. The best we can do as planners is to respond to our credences in various hypotheses as to what the objective moral facts might be, having conditionalized appropriately on our total evidence. Thus, the repartitioning strategy – or something very much like it – is the best we can do as planners in taking account of the possibility that Normative Externalism is true. We cannot be expected to respond to the facts (if they are facts) that this higher-order theory and some first-order moral theory are true. We can respond only to our credences in the combined hypotheses that they are true. Again, then, the repartitioning strategy seems to be the best that MAXEOV can do.

If my argument in this paper is correct, it means good news for expected-objective-value-maximizers. We simply do not have to worry about the possibility that Normative Externalism is true. We can respond to arguments for Externalism by listening patiently and then going about our days as if the arguments had never happened. And we can do this safe in the knowledge that, to the extent that the arguments we have heard have moved some of our credence in the direction of DTRT, the best way to incorporate this into our decision-making renders it equivalent to simply becoming even more confident of MAXEOV. Thus, maximizers can focus on their real enemies: other internalist decision-rules that propose rival ways for morally uncertain agents to take account of their credences in various first-order moral theories. Proponents of MAXEOV can pretty much just ignore DTRT, safe in the knowledge that, by doing so, they are acknowledging its possible truth as best they can.

4. Conclusion

Who's afraid of Normative Externalism?

Not me!

REFERENCES

Allais, Maurice (1953). "Le comportement de l'homme rationnel devant le risque: critique des postulats et axiomes de l'école Américaine". *Econometrica*. 21: 503–546.

Buchak, Lara (2013). *Risk and Rationality*. Oxford, UK: Oxford University Press.

Dreier, James (2011). "In Defense of Consequentializing." In Mark Timmons (ed.), *Oxford Studies in Normative Ethics, Volume 1,* pp.**-**. Oxford, UK: Oxford University Press.

Enoch, David (2014). "A Defense of Moral Deference." *Journal of Philosophy* 111: 229-258.

Gracely, E. J. (1996). "On the Noncomparability of Judgments Made by Different Ethical Theories". *Metaphilosophy* 27(3), pp. 327–332.

Gustafsson, Johann & Olle Torpman (2014). "In Defence of My Favorite Theory." *Pacific Philosophical Quarterly* 95:159-174.

Harman, Elizabeth (2015). "The Irrelevance of Moral Uncertainty". In Russ Shafer-Landau (ed.), *Oxford Studies in Metaethics, Volume 10*, pp.53-79. Oxford, UK: Oxford University Press.

Hedden, Brian (2016) "Does MITE Make Right? Decision-Making Under Normative Uncertainty". In Russ Shafer-Landau (ed.), *Oxford Studies in Metaethics, Volume 11*, pp. 102-128. Oxford, UK: Oxford University Press.

Hicks, Amelia (forthcoming). "Moral Uncertainty and Value Comparison". In Russ Shafer-Landau (ed.), *Oxford Studies in Metaethics, Volume 13*, pp. 161-183. Oxford, UK: Oxford University Press.

Hurley, Paul (2013). "Consequentializing and Deontologizing: Clogging the Consequentialist Vacuum". In Mark Timmons (ed.), *Oxford Studies in Normative Ethics, Volume 3*, pp. 123-153. Oxford, UK: Oxford University Press.

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

MacAskill, William (2013). "The Infectiousness of Nihilism". *Ethics* 123: 508-520.

Portmore, Douglas W. (2007). "Consequentializing Moral Theories". *Pacific Philosophical Quarterly* 88: 39–73.

Ross, Jacob. (2006). "Rejecting Ethical Deflationism". *Ethics* 116: 742–768.

Savage, Leonard (1954). *The Foundations of Statistics*. New York, NY: John Wiley and Sons.

Sepielli, Andrew (2009). "What to Do When You Don't Know What to Do." In Russ Shafer-Landau (ed.), *Oxford Studies in Metaethics, Volume 4*, pp. 5-28. Oxford, UK: Oxford University Press.

Sepielli, Andrew (2013a). "Moral Uncertainty and the Principle of Equity Among Moral Theories". *Philosophy and Phenomenological Research* 86: 580-589.

Sepielli, Andrew (2013b). "What to Do When You Don't Know What to Do When You Don't Know What to Do…" *Noûs* 48: 521–44.

Weatherson, Brian (2013). Running Risks Morally. *Philosophical Studies* 167: 1-23.

Weatherson, Brian (*ms*). *Normative Externalism*.