Moral Uncertainty and Value Comparison

Amelia Hicks

[Working draft—please do not cite without permission]

Abstract: Several philosophers have recently argued that decision-theoretic frameworks for rational choice under risk fail to provide prescriptions for choice in cases of *moral uncertainty* (where moral uncertainty is an epistemic state in which one's credences are divided between moral propositions). They conclude from this that there are no rational norms that are "sensitive" to moral uncertainty. This conclusion is surprising; if it's correct, then there's no rational requirement that moral uncertainty affect one's moral deliberation, even if one cares about acting in accordance with moral norms.

In this paper, I argue that one has a rational obligation to take one's moral uncertainty into account in the course of deliberation (at least in some cases). I first provide positive motivation for the view that one's moral beliefs can affect what it's rational for one to choose. I then address the problem of value comparison, which shows that when we're uncertain between competing moral propositions, we cannot determine the expected moral value of our actions. I argue that we should not infer from the problem of value comparison that there are no rational norms governing choice under moral uncertainty; even if there is no way of determining the "expected moral value" of one's actions in cases of moral uncertainty, a morally-motivated decision-maker can still have preferences over lotteries that entail the existence of rational requirements for choice.

Introduction

Consider the following possible line of reasoning:

> I know how animals are farmed, and that their treatment on those farms is, very often, horrifying. I also know that the likelihood of me being able to change any of that—especially through some sort of boycott—is very low. But I'm still not

> sure what to do, because I'm uncertain of which of the following two types of moral views is true. On the one hand, it might be permissible for me to make no changes to my life under these circumstances, given the low probability of any changes being effective. On the other hand, it might be impermissible for me to make no changes, in light of the fact that those changes would not be burdensome to me, and *could* make *some* difference. My decision, then, will be to try to do something. Not doing anything might be permissible, but it might also be very wrong; whereas doing something is surely permissible.

One important feature of this example is that the decision-maker's uncertainty about what to do stems from uncertainty about which of several moral propositions is true; they're uncertain between a moral proposition expressing a more permissive moral view, and a moral proposition that expresses a less permissive moral view. (We could change the example so that the decision-maker is uncertain about a different set of moral propositions. For example, they could be uncertain between a consequentialist theory and a theory according to which "merely symbolic" actions are intrinsically morally valuable. Or they could be uncertain between a theory that accords some moral worth to the lives of animals, and another theory that doesn't.)

In this paper, I defend the view that the type of reasoning in the above example is, at least sometimes, rationally required. More generally, I defend the view that if one cares about acting in accordance with moral norms[1] and is aware of one's moral uncertainty, then one's moral uncertainty can affect which action it's rational to perform. I argue that the main objection to this type of reasoning, *the problem of value comparison*, does not actually pose a challenge to this view.

In sections 1 and 2, I begin by providing some basic background and clarifying the concept of moral uncertainty. In section 3, I argue that when one cares about acting in accordance with moral norms and one is presented with an action that "dominates" all other available actions, then one should choose in accordance with the principle of dominance. In section 4, I consider what one's rational obligations are when one is morally uncertain but *no* available action dominates ("no-dominance" cases). I argue that, in light of the rational requirement to reason in accordance with dominance, it's implausible to think that in no-dominance cases one *never* has *any* rational obligation to attend to one's moral uncertainty. I then present the problem of value comparison, which allegedly shows that one is never rationally required to take one's moral uncertainty into account in the course of moral deliberation. In

[1] I won't address moral objections to caring de dicto about acting in accordance with moral norms. For a response to those sorts of worries, see Andrew Sepielli, "Moral Uncertainty and Fetishistic Motivation," *Philosophical Studies*, forthcoming.

section 5, I argue that the problem of value comparison, as raised in the moral uncertainty literature, is an instance of a general puzzle about rationality. Other instances of the puzzle do not demonstrate that we lack rational requirements that we originally thought we had. Thus, this instance of the puzzle shouldn't convince us that we lack a rational requirement to attend to our moral uncertainty. In section 6, I address the objection that my response to the problem of value comparison relies on an overly optimistic view of the capacities of everyday decision-makers. Finally, I conclude by surveying several interesting questions raised by my position.

1 Background

Beginning in section 3, I will assume some familiarity with basic decision theory and basic measurement theory. Here, in section 1, I clarify those concepts I'll be assuming familiarity with. Readers already familiar with these concepts may want to skip to section 2.

1.1 Actions, States, Outcomes, and Utility

I will frequently use the terms *action* (*act*), *state of nature* (*state*), *outcome*, and *utility.*

The *actions* (or *acts*) I will be discussing are those that the decision-maker is choosing between. (A rough definition of an action will suffice for our purposes.) *States of nature* are those states of the world that can affect the outcomes (and thus the utility; see section 1.3 below) of the actions one is choosing between. I discuss constraints on the states of nature in the following subsection. An *outcome* is determined by (and is sometimes treated as identical to) an act/state pair. To illustrate these three concepts, consider a case in which you must decide whether to take your umbrella with you to work. At the time of your decision, the actions available to you are to *take the umbrella* and to *not take the umbrella*; the possible states of nature are that there will be *some rain* and that there will be *no rain*; and the possible outcomes are (plausibly) *staying dry while being encumbered, staying dry while being unencumbered, and getting wet while being unencumbered.* We can represent your decision using a decision matrix:

	Rain	**No Rain**
Take Umbrella	Dry + Encumbered	Dry + Encumbered
Don't Take Umbrella	Wet + Unencumbered	Dry + Unencumbered

Given this description of the choice situation, it's unclear which action (taking or leaving the umbrella) is rational for you, the decision-maker, to perform. In order to determine which course of action is rational, we need to know something about your attitudes towards the various possible outcomes. I will discuss this further in sections 1.3 and 1.5.

1.2 Dominance (Weak and Strong)

Everyone accepts that if one is uncertain between mutually exclusive, jointly exhaustive, act-independent propositions (propositions that together express a partition of the states of nature), then one ought to reason in accordance with the *rule of dominance*.[2]

There are two forms of dominance. According to *weak dominance*, if action A weakly dominates action B, then one is rationally required to choose A. A weakly dominates B if and only if (1) A does not yield a worse outcome than B under any of the states of nature and (2) there is at least one state of nature under which A yields a better outcome than B. For example, let's say that I'm trying to decide whether to take an umbrella with me for my walk to work. If it rains, then carrying the umbrella with me is better than leaving it at home. And if it doesn't rain, the umbrella is nevertheless light enough that I won't notice it. Thus, because it will either rain or not rain (those being the two possible states of nature), carrying the umbrella weakly dominates not bringing an umbrella with me; carrying the umbrella is sure to yield a result that is either as great as or better than not bringing it with me.

According to *strict* or *strong dominance*, if action A strictly dominates action B, then one is rationally required to choose A. A strictly dominates B if and only if A yields a better outcone than B under each of the states of nature. For example, let's say that I'm trying to decide whether to walk my dog. Either my dog is in the sort of mood in which she really needs to be walked, or she's in the sort of mood in which she would enjoy walking but need not walk. If she needs to be walked, then I'm much better off walking her (otherwise, she'll destroy my house); but even if she doesn't need to be walked, I'm still better off walking her, because it will benefit me to get some exercise and fresh air. In this type of case, walking my dog strictly dominates not walking my dog, because walking my dog is sure to yield a greater payoff, no matter my dog's mood.

When applying either version of dominance, there are three constraints on the propositions used to express the states of nature. First, they must be mutually exclusive. If they

[2]Some have treated paradoxes such as Newcomb's Paradox or the Two Envelopes Paradox as counterexamples to dominance. However, these paradoxes merely illustrate that the rule of dominance cannot be applied in every type of decision; they do not illustrate that *appropriately restricted* applications of dominance yield irrational choices. This paper focuses on decisions that fall within the range of cases to which dominance can be appropriately applied.

aren't mutually exclusive, then the choice situation has been misdescribed; one would, in a sense, be choosing between a false dichotomy (or a false trichotomy, etc., depending on how many propositions one is uncertain between). For example, if one is planning to bake a tart and it's possible that both pears and apples are available, one would not want to say that the only relevant states of nature are *apples are available* and *pears are available*. After all, if *both* are available, that may very well affect the type of tart that it's best to bake. Second, the states of nature must be jointly exhaustive. If the states of nature are not jointly exhaustive, then one simply doesn't know all of the possible situations one might be in, and thus is unable to determine the possible outcomes of one's actions. For example, if I've neglected the possibility that my dog is in fact in *no* mood to walk—maybe she needs to be taken to the hospital instead—then I've neglected a state of nature that could significantly affect the outcomes of my actions. And, third, the propositions must be independent of the actions one is choosing between—they must be states of the world that won't be affected by the choice being made. To see why we need this constraint, imagine the following line of reasoning in which the states of nature are *not* act-independent: the two states of nature are *I'll eat well tonight* and *I'll go hungry tonight*, and the two actions available to me are cooking something and cooking nothing. I could then reason in this way: "If I go hungry tonight, then there's no point in cooking anything, since I'll just be hungry anyway. But if I eat well tonight, it would be much nicer to not have to cook anything. Thus, I shouldn't cook anything, because not cooking weakly dominates cooking." Clearly this is a terrible way of reasoning, because whether or not I eat well tonight will at least partially depend on whether I choose to cook. There is some debate about the type of independence required by this last constraint. However, that debate does not affect anything in this paper.

1.3 Expected Utility

To determine an action's *expected utility*, one (a) determines that action's utility under each state of nature, and then (b) for each state of nature, one multiplies the action's utility under that state by one's credence level in that state, and then (c) sums those products.

When one reasons in accordance with dominance, one thereby chooses the action with the highest expected utility.[3] However, it's possible to perform an action with the highest expected utility even when no available action dominates the others. Let's take our example from before, while assuming that the decision-maker assigns certain utilities to the possible outcomes:

[3]Again, when the application of dominance is appropriately restricted.

	Rain	**No Rain**
Take Umbrella	7	7
Don't Take Umbrella	1	10

Here we've assigned utilities to the outcomes, so we have some sense of how the decision-maker values those outcomes. And, it's clear from the decision-maker's evaluations of outcomes that no available action dominates the other. However, if we also know how probable the decision-maker believes rain to be, then we can still determine which of the two actions has the highest expected utility. Let's say that the decision-maker has a 0.1 credence that it will rain, and a 0.9 credence that it will not.

	Rain (0.1)	**No Rain (0.9)**
Take Umbrella	7	7
Don't Take Umbrella	1	10

In that case, *not* taking the umbrella has the higher expected utility (9.1, as opposed to 7 for taking the umbrella), even though in some sense that's the riskier action.

Thus, to determine the expected utility of an action, the decision-maker must not only have preferences that determine the values of the possible outcomes (more on how we can represent this evaluation of outcomes in section 1.5), but also have precise credence-levels in the different possible states of nature.

1.4 Levels of Measurement

In the next subsection, we'll learn that a set of preferences (that satisfies certain constraints) can be represented using a cardinal utility function, a function that can numerically represent the utilities of possible outcomes. However, to understand what sort of information such a function gives us, we need familiarity with levels of measurement and different types of rankings. The following scales increase from weakest to strongest: ordinal, interval, ratio, and absolute. When items are ranked on a weaker scale, we have less information about the ranked items, and thus can meaningfully represent the ranking of those items in more ways.

The weakest scale is an *ordinal* scale, on which items are simply ranked relative to each other, but on which we lack information about the *distances* between ranked items. For example, I could have the following ranking of ice cream flavors:

$$\textit{chocolate chip cookie dough} \succ \textit{butter pecan} \succ \textit{bubblegum}$$

But if all you have access to is this ranking, you do not know *how much more* I prefer butter pecan to bubblegum, or chocolate chip cookie dough to butter pecan. (Note that some items ranked on an ordinal scale can be equal to each other. For example, even if I'm indifferent between butter pecan and Mackinac Island Fudge, those two items still count as ordinally ranked.)

The second weakest scale is an *interval* scale, which not only expresses an ordinal ranking but also provides information about the intervals between the ranked items. However, if items are ranked only on an interval scale, you still lack information about the ratios between the ranked items—this is because an interval scale does not have a meaningful fixed zero-point. Most measurements of heat are on an interval scale. For example, in Fahrenheit, 90F is 5 degrees hotter than 85F. However, it isn't meaningful to say that 90F is *twice as hot* as 45F. One way of expressing this is to say that a ranking on an interval scale is "unique up to positive affine (i.e., linear) transformation"; we could transform the Fahrenheit scale using any positive linear function, and get a scale that expresses the very same information (albeit using different numbers).

The third weakest scale is a *ratio* scale. This is a scale that expresses not only information about the intervals between ranked items, but also information about the ratios between them; this is because a ratio scale does have a meaningful fixed zero-point. Measurements of distance use ratio scales; for example, we can meaningfully say that two miles is twice as far as one mile. Because a ratio scale is stronger than an ordinal or an interval scale, there are fewer permissible transformations that can be performed on a ranking on a ratio scale. However, measurements on a ratio scale can still be transformed in some ways; notice, for example, that we can express the same distance using inches or centimeters.

There is another type of scale—a maximally strong, *absolute* scale, for which there are no permissible transformations—that I will discuss later in this paper.

1.5 Rationality and Utility Maximizing

Decision theory is often described as the theory of rational choice. However, the type of rationality studied by decision theorists is somewhat "thin"; to be rational is to be *representable* in a decision-theoretic framework, which amounts to satisfying a set of consistency constraints. One of the most important insights of 20th century decision theory is that an agent who satisfies those constraints (and who is thus rational in a thin sense) can be

represented as a *utility maximizer.*

This insight comes from the development of *representation theorems.* There are several different representation theorems that one can prove by working with slightly different sets of axioms. But these representation theorems have several core features in common, so I won't distinguish between them here. For simplicity's sake, I'll use the Von Neumann-Morgenstern representation theorem as a guide.

Representation theorems tell us that when an agent has a set of preferences that satisfies certain requirements, then there is a cardinal utility function that "represents" that agent; that is, the agent can be represented as choosing so as to maximize some value, which we can call expected utility ("utility" here simply refers to a value given by a function that represents the decision-maker's preferences). The set of preferences must be an ordinal ranking over lotteries, where a lottery is a set of outcomes, each of which is paired with a probability. (Note that an outcome is a degenerate lottery; it's a set of a single outcome that's paired with the probability 1.) Moreover, that ordinal ranking over lotteries must satisfy the axioms of completeness, transitivity, independence, and continuity; although I won't go into detail about these axioms, they arguably jointly represent a consistency requirement for the decision-maker's preferences. Without going into the details of the proof here, the representation theorem tells us that if someone has a set of preferences (an ordinal ranking over lotteries) that meets these requirements, then there is a cardinal utility function that represents that person. More specifically, that function supplies us with information about how the decision-maker values lotteries *on an interval scale.* This allows us to assign *numerical* values to outcomes that represent how valuable each outcome is to the decision-maker. Those numbers represent the values that we can treat the decision-maker as maximizing.

To summarize so far: representation theorems tell us that if a decision-maker has a rational (in the thin sense) preference set, then we can assign numerical values (on an interval scale) to outcomes, and those values will represent the utility of those outcomes for the decision-maker. For example, we might be able to derive the following utility evaluations of outcomes:

	Rain (0.1)	**No Rain (0.9)**
Take Umbrella	7	7
Don't Take Umbrella	1	10

using a representation theorem. However, representation theorems can sometimes be used in another way: to supply information about an outcome itself (by supplying information about

the states of nature). Let's work with a different example. Imagine that you're having your friend over, and you'd like to offer her a snack. Unfortunately, you're not sure whether she *only* likes pistachios or whether she doesn't like them at all (both possibilities seem equally likely to you), and you need to decide whether to buy some to serve to her. (Assume that if you buy them, then there will be no other snacks.) We can represent your decision like this:

	Likes Pistachios (0.5)	**Doesn't Like Pistachios (0.5)**
Buy Pistachios	Happy Friend	Unhappy Friend
Don't Buy Pistachios	Unhappy Friend	Happy Friend

One question we might be interested in is *how happy* the friend would be if she likes pistachios and you buy them, compared to how happy she would be if she doesn't like pistachios and you don't buy them; after all, it's entirely possible that while she would be *very happy* if she likes them and you buy them, she would be only mildly happy if she doesn't like them and you don't buy them. If both possible versions of your friend are rational (by the lights of decision theory—that is, in both cases, she can ordinally rank lotteries in a way that satisfies all of the relevant constraints), then we can apply a representation to both possible versions of her, and thereby represent her utility levels for each outcome. Maybe they would look something like this:

	Likes Pistachios (0.5)	**Doesn't Like Pistachios (0.5)**
Buy Pistachios	10	1
Don't Buy Pistachios	3	7

Note that in this case the numbers are not representing the utility that you, the decision-maker, assign to outcomes; rather, they're representing how the friend would value certain of your actions. To figure out how to rationally decide what to do, you would have to determine how *you* value those outcomes, where each outcome includes your friend's evaluation of your action. And the utilities you ascribe to the various outcomes need not align with your friend's evaluations of your possible actions. This last point will turn out to be important for the main argument in this paper.

2 What's Moral Uncertainty?

"Moral uncertainty" can refer to different types of mental states. Most generally, moral uncertainty refers to a state in which one doesn't know what to do, morally speaking.

But there can be different reasons why one does not know what to do, morally speaking. In order to arrive at a moral decision—a decision about what's *morally* to be done—in an actual case, one must not only know about the *facts* of the case, but also about the proper *moral framework* to apply to that case.[4] So, one might be morally uncertain because one lacks crucial *empirical, factual* or *descriptive* information; for example, I might not know whether it's permissible for me to lie to my friend, because I lack crucial information about the consequences of such a lie. But one might also be morally uncertain because one lacks crucial *moral* information; so, for example, I might not know whether it's permissible for me to lie to my friend because I'm uncertain of whether such a lie would be intrinsically impermissible (regardless of its possible good consequences). Thus, we can distinguish between *descriptively-based moral uncertainty* and *morally-based moral uncertainty.*

We can further distinguish between different types of morally-based moral uncertainty. One might not know what's morally right to do because one is unsure of the correct normative theory. This is the type of moral uncertainty experienced by someone who isn't sure whether some version of consequentialism or some version of Kantian deontology is correct. Call this *theory uncertainty.* One might also not know what's morally right because one is unsure of what's morally valuable. This is the type of moral uncertainty experienced by a consequentialist who is unsure of which alleged "goods" to maximize. Call this *value uncertainty.*[5] And one might be morally uncertain because one doesn't know what one's moral obligations are. Call this *obligation uncertainty.* There are many reasons one might experience obligation uncertainty. It's the type of moral uncertainty experienced by a committed Kantian who isn't sure what their theory entails about a particular case (and not for lack of factual information). It is also the type of moral uncertainty that a moral particularist could experience in those cases in which they aren't able to determine what they ought to do.[6]

[4]This may be an over-intellectualized description of how people often arrive at moral verdicts; almost certainly one can responsibly arrive at a moral verdict without consciously distinguishing between these two types of information. However, if we assume that there is a distinction between descriptive and normative information, then it's clear that one needs both types in order to responsibly arrive at a moral verdict, even if the process by which one arrives at that verdict isn't explicit.

[5]Note that value uncertainty is plausibly a consequence of a type of theory uncertainty, since it's uncertainty about the extension of the correct axiological theory (where that axiological theory could be used to "fill out" a normative theory).

[6]W.D. Ross—an intuitionist who's not a particularist—also describes obligation uncertainty when he writes, "Where a possible act is seen to have two characteristics, in virtue of one of which it is *prima facie* right, and in virtue of the other *prima facie* wrong, we are (I think) well aware that we are not certain

In this paper, when I use the phrase "moral uncertainty," it will refer to *all forms of morally-based moral uncertainty*, unless I specify otherwise. Descriptively-based moral uncertainty is not at issue in this paper. This is because it's not controversial that one should take one's *descriptively*-based moral uncertainty into account—that's an issue on which my opponents and I agree. It will shortly become clear why I think my account of moral uncertainty applies to all forms of morally-based moral uncertainty. However, the reason is (briefly) this: I will argue that we should treat morally-based moral uncertainty in the same way that we treat descriptively-based moral uncertainty, because the *source* of moral uncertainty (in the broad sense) is irrelevant. Given that I think the source of moral uncertainty is irrelevant, it follows that the source of morally-based moral uncertainty is also irrelevant.

3 Moral Uncertainty and Dominance

Again, everyone accepts that one should reason in accordance with dominance (both versions) when one is uncertain between *descriptive* or *empirical* propositions, such as propositions about the weather or my dog's current mood (see section 1.2). (That is, everyone agrees that one should reason in accordance with dominance when the propositions that express the partition of the states of nature are descriptive propositions.) In this section, I simply want to argue that one should also choose in accordance with dominance when one is uncertain between moral propositions: when one is morally uncertain but one available action dominates the others, then one is rationally required to choose the dominating action (assuming that one cares about acting in accordance with moral norms).

The reason why dominance-reasoning extends to cases of moral uncertainty is that dominance is a formal rule for decision-making, and as such it is insensitive (for the most part) to the *content* of the propositions that express the relevant states of nature (such as *it will rain*, *it will not rain*, *my dog's in a walking mood*, *my dog's in a need-not-be-walked mood*). As we've already seen, when applying the rule of dominance there are only three constraints on the states of nature: that they're mutually exclusive, jointly exhaustive, and act-independent. Instances of moral uncertainty—cases in which one's credences are divided between moral propositions—can satisfy all three of these constraints. Clearly one's credences could be divided between mutually exclusive moral propositions (such as *it's sometimes permissible to kill an adult human* and *it's never permissible to kill an adult human*). And clearly those mu-

whether we ought or ought not to do it; that whether we do it or not, we are taking a moral risk.... For, to go no further in the analysis, it is enough to point out that any particular act will in all probability in the course of time contribute to the bringing about of good or of evil for many human beings, and thus have a *prima facie* rightness or wrongness of which we know nothing." See Ross' *The Right and the Good* (Indianapolis: Hackett Publishing Company, 1930), 30-1.

tually exclusive propositions could be jointly exhaustive (as with the last pair of propositions in parentheses). And clearly those mutually exclusive, jointly exhaustive propositions could be act-independent. Very few moral realists believe that you can *make* a moral proposition true simply by performing certain types of actions.[7]

Thus, we should accept the conclusion that we ought to reason in accordance with dominance when morally uncertain. If (a) I'm uncertain between mutually exclusive, jointly exhaustive, act-independent moral propositions, p and q and (b) I must choose between performing action A and performing action B, but (c) action A will be morally as good as or better than B no matter whether p or q is true (and under at least one state of nature will yield an outcome better than B), then I ought to choose A over B.

4 Moral Uncertainty Without Dominance: Can We Compare Values?

One might think that this conclusion is itself useful for the purposes of reasoning about real-world moral problems when one is morally motivated; but it isn't. For example, if we grant that one can have moral obligations to oneself, then it will turn out that there are very few cases in which one action even weakly dominates all others. Take the example provided at the beginning of the introduction to this paper. The states of nature in that example were, essentially, *that a more permissive theory is true* (according to which it's permissible to forego making changes that have only a low probability of producing positive change) and *that a less permissive theory is true* (according to which one is obligated to take non-burdensome steps in order to even slightly increase the probability of producing positive change). One assumption built into that example was that making changes would in *no way* be burdensome to me; but in the real world, that assumption is implausible. Given the implausibility of that assumption, and assuming the moral value of at least some instances of prudence, it's no longer clear that, in that example, taking action dominates not taking action. The less permissive theory doesn't specify what to do when the only changes one can make would be burdensome. And we can easily imagine a specification of the less permissive theory according to which one should *not* make burdensome changes. It's then no longer

[7]The foregoing argument explains one source of confusion in Nissan-Rozen's paper "Against Moral Hedging," in which (in one section) he argues that dominance reasoning under moral uncertainty conflicts with the *actual* rule of dominance. His argument goes astray when he introduces an additional constraint on the propositions that express the states of nature: he asserts that they must be *descriptive*, by which he means non-normative. All that Nissan-Rozen shows is that dominance reasoning *could* recommend one action with respect to one partition, but not with respect to another partition. But that only demonstrates that we would need to work with a finer-grained partition. See Ittay Nissan-Rozen's "Against Moral Hedging," *Economics and Philosophy* Vol. 3 (2015), 1-21.

the case that making changes is sure to yield a "moral payoff" as great as or greater than not making changes.

These "no-dominance" cases of moral uncertainty—cases in which no action dominates the others—will occur regularly, and not just because of the value of (some) instances of prudence. One might have to decide whether or not to kill a person in defense of another person, while one's credences are divided between a theory that says doing so would be impermissible and another theory that says that *not* doing so would be impermissible. In this sort of case, too, no available option even weakly dominates the others. No-dominance cases abound because the real-world decisions that are worth spending time thinking about—*tough* decisions—are those in which no available course of action is sure to be at least as good as any of the other available courses of action. Thus, the conclusion that we ought to reason in accordance with dominance when we're morally uncertain tells us very little about how we actually ought to reason in the types of situations we're likely to face.[8]

What should we make of cases of moral uncertainty in which no action dominates the others ("no-dominance" cases)? I will argue that we have good reason to think that one has a rational obligation to take one's moral uncertainty into account in (at least some) no-dominance cases; however, it will turn out that the way in which one ought to do this is highly dependent on one's specific credences and preferences. Moreover, in *some* no-dominance cases, an agent might not have any rational obligation to take their moral uncertainty into account. But nevertheless, we should not infer from this complexity (or from the lack of rational obligations *in some instances* of moral uncertainty) that *all* no-dominance cases are cases in which one lacks any rational obligation to attend to one's moral uncertainty.

I argue by disjunction elimination. First, it could be that one lacks any rational obligation to attend to one's moral uncertainty in no-dominance cases. Second, it could be that one does, at least sometimes, have a moral obligation to consider one's moral uncertainty in no-dominance cases. I argue against the first possibility, and then defend the second possibility from the *problem of value comparison*.

4.1 The First Possibility: We Have No Rational Obligation in No-Dominance Cases

One natural thought is that the choice a morally-motivated person rationally ought to make in no-dominance cases is determined by (a) their credence-levels in the competing moral propositions, and by (b) how morally valuable, for each of those propositions, their available

[8] Jacob Ross makes a similar point; see "Rejecting Ethical Deflationism," *Ethics*, Vol. 116, No. 4 (July 2006), p. 753.

actions would be if that proposition were true.[9] This natural thought suggests that we extend utility theory to cover cases of "moral risk" in which no available action dominates the others.[10]

As we'll see in section 4.2, several philosophers reject this natural thought because of the problem of value comparison. They then infer from the failure of utility theory to adequately address decision-making under moral uncertainty that one's rational obligations are not affected by moral uncertainty; they think that one's rational obligations are determined by one's descriptive beliefs, not by one's moral beliefs,[11] and they thereby endorse the first possibility:

> *Possibility 1:* in no-dominance cases of moral uncertainty, one never has a rational obligation to take one's moral uncertainty into account in the course of moral deliberation (even if one cares about satisfying moral norms).

In this section, I want to argue against the first possibility; we have good reason to reject the view that one's moral beliefs do not affect one's rational obligations.[12]

My argument against the first possibility proceeds from the assumption that we have a rational obligation to consider our moral uncertainty in cases of *dominance*. If I'm uncertain between two moral propositions, but no matter which proposition is true it's morally as

[9] Jacob Ross also discusses this move to decision theoretic frameworks for choice under risk when dealing with no-dominance cases; see "Rejecting Ethical Deflationism," p. 13.

[10] Notice that one can endorse this natural thought without being committed to the view that utility theory describes the decision-making process that one should consciously engage in; that is, one can endorse the natural thought and at the same time believe that utility theory only shows us what a decision-maker ought to choose in light of their preferences and credences (without making any recommendation for the type of deliberative process they should use).

Here I've described the natural thought as a thought about moral risk—that is, I've described it as a view about what one should do when one is uncertain between various moral propositions and can assign probabilities to those propositions. However, one could also hold the natural thought about cases of moral uncertainty in which one can't assign probabilities to the propositions one is uncertain between (we could call these cases of "moral ignorance," as opposed to "moral risk"). In this paper I focus on moral risk. However, the conclusions of my paper should straightforwardly apply to cases of moral ignorance. Because moral beliefs are not significantly different from other types of beliefs, whatever decision criterion (maximin, minimax regret, the principle of insufficient reason, tempered regret, etc.) for choosing under ignorance is appropriate in other cases will be appropriate for cases of moral ignorance. For a general discussion and comparison of some of these rules, see Luce and Raiffa's *Games and Decisions: Introduction and Critical Survey* (New York: Dover, 1989), 275-324. For a discussion of tempered regret, see Acker's "Tempered Regrets Under Total Ignorance," *Theory and Decision*, Volume 42, Issue 3 (1997), pp. 207-213.

[11] For an example of this inference, see Brian Hedden, "Does MITE Make Right? On Decision-Making Under Normative Uncertainty," *Oxford Studies in Metaethics, Volume 11*, (New York: Oxford University Press, 2015), p. 24.

Note that the page numbers provided for Hedden's paper are for the unpublished version; pagination is not yet available for the version printed in *Oxford Studies in Metaethics, Volume 11*.

[12] To use Hedden's language, I will argue that we have good reason to believe that there is a "super-subjective ought." See Hedden, pp. 1-2.

good as or better for me to choose action A, then I should choose action A. (Provided that I care about moral norms. If I don't care about moral norms, then my moral uncertainty does not affect the rationality of my action. To see this, consider a comparison to the person who's reasoning about whether or not to take their umbrella with them, but who doesn't care about whether or not they get rained on.)

I can begin to sketch my worry about this first possibility by pointing out that we don't typically believe that there's a drastic difference between dominance and no-dominance cases when dealing with *other* forms of uncertainty. So, consider the case in which I'm trying to decide whether to carry my umbrella with me, and it would be *very mildly* annoying to carry it, in the event that it doesn't rain. No one would say that there are no longer *any* rational norms that could govern my choice. After all, the probability of rain might be high, and I might greatly prefer to not get rained on; in light of those considerations, it might be worth risking a mild annoyance. And, moreover, I would be irrational if I were to risk a very high probability of something I hate only for the sake of avoiding an improbable slight annoyance. So, my first point is that, given that we don't believe that rational norms go out the window in no-dominance cases involving other forms of uncertainty, we would need a *very* persuasive reason for thinking that that happens with moral uncertainty.

We can make this first point more persuasive if we look at *why* we don't normally dispense with all rational norms in non-moral no-dominance cases. Some decision theorists would want to explain the rationality of dominance reasoning in terms of a broader framework for rational decision-making; such a framework would extend to no-dominance cases.[13] Other decision theorists take an axiomatic approach; they begin with dominance as an axiom describing a constraint on rational choice, and then build up their decision-theoretic framework by adding more axioms consistent with dominance.[14] However, whether or not one takes the "broad framework" approach or the "axiomatic" approach, *no one just stops with dominance.*

Because of this, I infer that the burden of proof falls on someone who holds that we *should* stop with dominance in the case of moral uncertainty—that is, the burden of proof falls on those who think that we have rational obligations to reason in accordance with dominance when morally uncertain, but that we have no rational obligation to consider our moral uncertainty in no-dominance cases. In the remainder of this paper I'll examine and respond to an attempt to take up that burden.

[13]For example, Bayesians think something like this; the Bayesian framework for rational decision-making explains why dominance expresses a rational obligation, and that framework also makes recommendations for how to choose in no-dominance cases.

[14]For example, see Luce and Raiffa, *Games and Decisions: Introduction and Critical Survey* (New York: Dover, 1957), pp. 286-298.

4.2 The Problem of Value Comparison

The second possibility—which I endorse—says that one does, at least sometimes, have a rational obligation to consider one's moral uncertainty in no-dominance cases.

> *Possibility 2:* in some no-dominance cases of moral uncertainty, one has a rational obligation to take one's moral uncertainty into account in the course of moral deliberation (if one cares about satisfying moral norms).

This second possibility is supported by the failure of the first possibility. However, the second possibility faces the problem of value comparison. The following description of the problem of value comparison relies heavily on Brian Hedden's presentation of the objection.[15] However, I intend for my response to the problem to apply equally to other presentations of the objection.[16]

Put briefly, the problem of value comparison is this. In order for a morally-concerned person to take into account (in the course of deliberation) that they are uncertain between various moral propositions, they would need to determine the "expected moral value" of each of their available actions. To do this, they would need to determine how morally valuable each available action would be under each state of nature (where each state of nature is expressed by one of those moral propositions), and then weight those values according to their credence-level in each state of nature. However, there's no guarantee that an action's moral value under the assumption of one state of nature (expressed by a moral proposition) is *comparable* to that action's moral value under the assumption of another state of nature (expressed by another moral proposition).[17]

Unfortunately, this brief way of putting the problem is not adequate; it's not clear what's meant by "comparable," and even if we fix the type of comparability that's at stake it's still

[15] From Hedden's, "Does MITE Make Right? On Decision-Making Under Normative Uncertainty," *Oxford Studies in Metaethics, Volume 11*, (New York: Oxford University Press, 2015).

[16] Many have presented this objection and have offered responses to it, all of which Hedden rejects. See:
Hudson, "Subjectivization in Ethics," *American Philosophical Quarterly* 28 (1989), 221-229;
Lockhart, *Moral Uncertainty and its Consequences* (New York: Oxford University Press, 2000);
Ross, "Rejecting Ethical Deflationism," *Ethics*, Vol. 116 (2006), 742-68;
Sepielli, "What to Do When You Don't Know What to Do," *Oxford Studies in Metaethics, Volume 4* (New York: Oxford University Press, 2009), 5-28.
Gustafsson and Torpman, "In Defence of My Favourite Theory," *Pacific Philosophical Quarterly*, Vol. 95 (2014), 159-174;
Riedener, *Maximizing Expected Value under Axiological Uncertainty*, Dissertation, University of Oxford (2015);
Nissan-Rozen also appeals to the problem of value comparison to argue against the view that rational norms are sensitive to one's moral beliefs. See Nissan-Rozen's "Against Moral Hedging," *Economics and Philosophy* Vol. 3 (2015), 1-21.

[17] Note that this remains a problem even when we assume that the moral propositions are mutually exclusive, jointly exhaustive, and act-independent.

not clear *why* we get a failure of comparability. At this point, I think it's most helpful to work with definitions of comparability and commensurability that are similar in spirit to Chang's definitions. Two outcomes are *comparable* just in case it is possible to *ordinally rank* them, that is, just in case it's possible to say that one is greater than, equal to, or less than the other.[18] Two outcomes are *commensurable* just in case it's possible to *cardinally rank* them, that is, just in case it's possible to "precisely" rank them "by some unit of value."[19] There are three things to note about these definitions. First, Chang's use of the phrase "cardinal ranking" is logically stronger than decision theorists' use of the phrase. Decision theorists think of a cardinal ranking as a ranking on an *interval* scale, and thus as a ranking that remains the same under positive affine transformation. Chang, it seems, thinks of a cardinal ranking as a ranking that cannot remain the same under any transformation; it's a truly unique ranking. Thus, to avoid confusion, I will refer to Chang's cardinal ranking as an *absolute ranking*, a ranking with no permissible transformations. Second, on this understanding of the comparability/commensurability distinction, comparability does not entail commensurability (and incommensurability does not entail incomparability). For example, you might think that respecting someone's privacy is more morally valuable than the small thrill you might get from prying into their affairs, even if you think that the value of privacy can't be represented using the same units used to represent happiness (or even if you think that the values of privacy and happiness can't be represented by units of measurement at all). Third, a set of alternatives are only comparable/commensurable *relative to some value* (which Chang calls a "covering value"). So, for example, we can compare *going bowling* and *going mountain-climbing* with respect to many different values: with respect to the fun it will produce for me, with respect to the resources that going will consume, with respect to physical safety, and so on.

With these clarifications in place, we can attempt to more clearly state the problem of value comparison: there's no way to *absolutely* rank, *with respect to moral value*, the outcome of action A under the assumption that moral proposition p is true to the outcome of action A under the assumption that moral proposition q is true. As a result, it's not clear how we could determine the "expected moral value" of performing A. We face the same problem when attempting to determine the expected moral value of any of our available actions, and thus we are unable to determine which available action has the highest expected moral value. And, as a result, there is no action which one is rationally required to choose under moral

[18] Chang would reject this last paraphrase of the definition of comparability, since she endorses the existence of a fourth relation that can hold between sets of alternatives. See Ruth Chang, "Introduction," *Incommensurability, Incomparability, and Practical Reasoning,"* ed. Ruth Chang (Cambridge: Harvard University Press, 1997).

[19] Chang, 2.

uncertainty. This gives us a clearer statement of the problem—the problem is, in fact, a problem of commensurability. But how can we motivate it?

Hedden motivates the problem in two ways. First, he argues that the *value functions* we associate with different moral theories are the results of very different sets of "preferences," and thus the theories' evaluations of actions are not commensurable. Second, he argues that not all theories can be assigned a value function. Hedden uses the phrase "value function" instead of utility function, presumably because it sounds strange to say that an outcome can have utility *for a theory.* However, Hedden's "value functions" just are utility functions, since we're treating "utility" here as simply the value given by a function that represents a (certain kind of) ordinal ranking of lotteries. Thus, I will continue to use the phrase "utility function," even when referring to those functions that represent the preferences of moral theories.

To understand Hedden's first argument, we need to understand how we might assign a utility function to a theory. The idea is that a moral theory expresses a set of "preferences"—ordinal rankings of lotteries—to which we can apply a representation theorem, assuming that those preferences satisfy certain requirements.[20] That is, assuming that a theory ordinally ranks lotteries in the right sort of way, we can represent that theory as recommending that one act so as to maximize some value, where that value can be numerically represented. One might wonder: how could we *numerically* represent how valuable actions are, according to a moral theory? The idea is this: we can "set the numerical scale" for that theory by looking at the outcomes that the theory ordinally ranks the *highest* and the *lowest.* The highest-ranking outcomes mark the top of the scale, while the lowest-ranking outcomes mark the bottom of the scale. We can set those top and bottom values however we like. Then, again assuming that the theory has the right sorts of preferences between lotteries, we can determine where all other lotteries fall on that scale, according to the theory. In this way, we can numerically represent the utility of any outcome according to the theory.[21]

Let's say that we do this for two competing moral theories and then determine the numerical value of possible outcomes according to each of the two theories. One might think that we can then use those numerical representations of the moral values of outcomes (according to the theories) to determine the *expected* moral value of each available action; that is, one might think that we can look at how valuable each outcome of an action would be if each moral theory were true, and then weight those values using our credence-levels

[20]For a nice description of axioms close the the von Neumann-Morgenstern axioms and of the von Neumann-Morgenstern Representation Theorem, see Luce and Raiffa, *Games and Decisions: Introduction and Critical Survey* (New York: Dover, 1957), especially 23-31.

[21]Jacob Ross also has a description of how a moral theory can be represented by a cardinal function; see "Rejecting Ethical Deflationism," p. 755.

in each theory. But the worry here is obvious: *how we set each of the theory's scales* will significantly affect our calculation of the expected moral value of our actions, and it isn't reasonable to assume that we ought to set the two scales in the same way (using the same numerical values to represent the top and bottom of each scale for each theory).[22] Thus, the evaluations of alternatives given to us by competing moral theories aren't commensurable, and so any calculation of expected moral value will be meaningless.

Hedden considers several solutions that have been proposed for the problem of value comparison. Lockhart, for instance, attempts to solve the problem by introducing the *principle of equity among moral theories* (PEMT), according to which "The maximum degrees of moral rightness of all possible actions in a situation according to competing moral theories should be considered equal. The minimum degrees of moral rightness of possible actions in a situation according to competing theories should be considered equal unless all possible actions are equally right according to one of the theories (in which case all of the actions should be considered to be maximally right according to that theory)."[23] However, PEMT is implausible.[24] For, imagine that I'm uncertain between Kantian deontology, on the one hand, and a theory according to which some actions are only a *little bit* better or worse than other actions, on the other hand. We could call the latter theory the "Meh" moral theory. Clearly, there will be some situations in which Kantianism and the Meh theory *do not* hold that there are equal amounts of moral value at stake.

Sepielli shows[25] that one can compare the value functions of competing theories if one also has a sufficiently large set of background beliefs (beliefs that establish points of comparison between the theories). However, Hedden points out that one will often lack a sufficiently large set of such beliefs, and that certain sets of background beliefs could actually yield *contradictory* comparisons between theories.[26]

Ross and Riedener show that one can compare the value functions of theories if one has some pre-existing beliefs about how to choose rationally when morally uncertain. That is, both Ross and Riedener argue that having intuitions about how to compare the moral values of *some* actions according to several moral theories can provide one with enough information to compare those theories' value functions, and thus to compare the evaluations delivered by

[22]Another way to put the worry: given that the utility function is unique only up to positive affine transformation (since it provides us with a ranking of outcomes on an interval scale), which transformation should we use when filling in the values in the decision matrix?

[23]Lockhart, 84.

[24]This objection to PEMT is not original to Hedden, and has been discussed at length by Sepielli. See Sepielli's "Moral Uncertainty and the Principle of Equity among Moral Theories," *Philosophy and Phenomenological Research*, Vol. 86, No. 3 (2013), 580-589.

[25]Sepielli, 2009.

[26]Hedden, pp. 10-11.

the theories in harder (less obvious) cases.[27] Hedden, however, claims to lack any intuitions of that sort, and thus rejects this proposal on the grounds that it fails to offer prescriptions for how to act rationally under moral uncertainty without presupposing *some* facts about how we rationally ought to act under moral uncertainty.[28]

Hedden's second argument proceeds from the observation that it might not be possible to apply a representation theorem to every moral theory, because some moral theories may have "preferences" that make the application of such a theorem impossible.[29] For example, a nihilistic theory denies the existence of moral values, and thus does not ordinally rank outcomes with respect to moral value; and, of course, if a theory does not ordinally rank lotteries, then we cannot represent it using a utility function.[30] And if we can't represent the theory using a utility function, then we can't numerically represent how valuable outcomes are according to that theory. Thus, in these cases, we again can't calculate the expected moral value of an action.[31]

5 Responding to the Problem of Value Comparison

A key feature of Hedden's two arguments is that each one relies on the observation that the evaluations of actions delivered by competing moral theories will never be commensurable. In the following subsection, I'll show that although that's correct, it does not entail that there are no rational requirements governing choice under moral uncertainty. In fact, there are decisions that are perfectly analogous to decisions made under moral uncertainty, and those analogous cases suggest that there are rational norms governing choice under moral uncertainty.

5.1 Two Sets of Evaluations

Recall that representation theorems play two roles in helping us understand rational choice under moral uncertainty (see section 1.5). First, we can (in some cases) apply a representation theorem to the preferences of a *theory*, and thereby represent that theory using a utility function that tells us how that theory ranks outcomes on an interval scale. Second, we can (in some cases) apply a representation theorem to the preferences of a *person* (the decision-maker), and thereby represent that person using a utility function that tells us how

[27] Ross, 2006; Riedener, 2015.

[28] Hedden, pp. 12-13.

[29] Hedden, pp. 14-19.

[30] There are other sorts of theories to which we cannot apply representation theorems. For example, a very demanding deontological theory could violate the axiom of continuity.

[31] This sort of problem is discussed at length in MacAskill's "The Infectiousness of Nihilism."

that person ranks outcomes on an interval scale. This is analogous to the example described in section 1.5:

	Likes Pistachios (0.5)	**Doesn't Like Pistachios (0.5)**
Buy Pistachios	10	1
Don't Buy Pistachios	3	7

In this example, your friend possibly likes pistachios and possibly doesn't like pistachios; thus, there are two possible versions of the friend. The numbers in the boxes of the decision matrix represent the utility of actions (buying pistachios, not buying pistachios) *according to those different versions of the friend*—they are those values provided by the utility functions associated with each version. This means that the numbers are helping to further *describe* the outcomes—they provide you with some information about how your friend might feel about different possible actions you might perform. However, in order to determine what's rational for *you* (the decision maker) to do, you must be able to ordinally rank lotteries (including those outcomes described in this decision matrix), so that *you* can be represented using a utility function.[32] Moreover, depending on how you rank lotteries, it could turn out that the utility you associate with the different possible outcomes *diverges* from the utility that the friend would assign to your actions.

The problem of value comparison is a problem for the *first* role of representation theorems; the problem is that even if we can associate competing theories with utility functions, those functions only deliver rankings on an interval scale, and thus are not commensurable. They're not commensurable because each ranking can be permissibly transformed in a variety of ways, and there's no principled way of determining which transformation to use for each theory. (As Hedden and Sepielli's responses to Lockhart show, we cannot simply assume that the scale for each theory is the same.) As a result, there's no meaningful way of determining the "expected moral value" of an action (since that would require that the competing theories rank outcomes on commensurable scales). This is an interesting result: there is no objective "expected moral value" that we can assign to an action when we're morally uncertain.

However, it doesn't follow from this that there are no rational norms governing choice under moral uncertainty; this is because even if the evaluations delivered by competing theories aren't commensurable, the *decision-maker* can still ordinally rank lotteries (including

[32]Many have observed that few actual people (if any) are ever able to develop the required type of ordinal ranking over lotteries. However, as I'll discuss later, this doesn't actually pose a problem for the thesis of this paper; even if an agent has an incomplete preference set, we can still represent them using a set of functions, and that set yields rational requirements.

the possible outcomes of the present decision) and thus assign utilities to the possible outcomes of their choice. It is *those* utilities—*not* those evaluations provided by the competing theories—that determine which action has the highest expected utility (and so is rationally required) for the decision-maker. Thus, although the problem of value comparison is a problem if one wants to perform whichever action has the highest expected moral value, it is not a problem for a decision-maker who cares about acting morally and can compare the possible outcomes of their choice.

To summarize so far: the problem of value comparison illustrates that the evaluations of actions provided by competing moral theories are not commensurable. However, that incommensurability does not entail that the decision-maker cannot *compare* outcomes (action/theory pairs). This conclusion is negative, since I've only shown the failure of an entailment relation. However, I think we have good positive reason to think that there are rational norms governing choice under moral uncertainty. There are two observations that support this stronger conclusion: first, the fact that in cases of moral uncertainty the states of nature expressed by moral propositions must be mutually exclusive; and second, the fact that the problem of value comparison is analogous to the problem of interpersonal utility comparison.

To see the first point, consider an example described by Ross:[33]

	Singer's Theory (0.5)	**Traditional Morality (0.5)**
Order Veal	Very Bad	Permissible
Order Veggie Wrap	Not So Bad	Permissible

Perhaps we can't associate the competing moral views with utility functions, or perhaps we can but the evaluations of actions provided by the theories aren't commensurable. But it doesn't follow from either of those possibilities that a decision-maker can't compare (ordinally rank) lotteries in such a way that allows them to compare the outcomes of their actions; and if a decision-maker can make those comparisons, then their choice is subject to rational norms. Moreover, we have good reason to think that a decision-maker *can* compare those outcomes, at least in many cases: the evaluations provided by the theories must be comparable if the competing moral theories are theories *about* the same thing. If the two theories are theories of radically different types of values (incomparable values, no less), then the theories would not actually be in competition with each other; two theories with very different structures can both be true, so long as they're theories about different things. (For example, it doesn't

[33]Ross (2006), 763.

bother us that a psychologist's theory of clinical depression is very different from a physicist's theory of motion. Both theories can be accepted in spite of their deep structural differences because they're theories *of* different things; they're not mutually exclusive.) When one is uncertain between mutually exclusive moral propositions, one must be uncertain between moral propositions that express "different takes" on the same things; and this observation supports the idea that in genuine cases of moral uncertainty, it is possible to compare what each moral proposition has to say about an action.[34] Exactly what it is that competing theories disagree about will depend on the theories in question;[35] nevertheless, genuinely competing theories must provide *comparable* evaluations of actions, even if those evaluations aren't commensurable. All of this is to say that a morally uncertain person can reasonably compare the outcomes of actions, even when the moral theories' evaluations of the decision-maker's actions aren't commensurable.

The second reason for thinking that there are rational norms governing choice under moral uncertainty is the fact that the problem of value comparison is perfectly analogous to the problem of interpersonal utility comparison (of which the pistachio case is one example).[36] But let's look at another example of interpersonal utility comparison, for the sake of seeing the close analogy between interpersonal utility comparison and value comparison. I care about my future well-being, and in particular I would like to make sure that I'm happy when I'm 80 years old (assuming I'm still alive). However, I know that I'm a capricious person who sometimes radically changes what she finds most valuable. Thus, I don't know whether the things I'll care about when I'm 80 will be the same as the things I care about now, or whether they'll be radically different. In particular, I don't know what my 80-year-old self's attitude will be towards money. It's possible that when I'm 80, I'll care a great deal about financial security. But it's also possible that my 80-year-old self will regret having stockpiled money over the course of my lifetime. Now, I have to decide: do I try to save money, or do I spend it? We can consider the different ways that 80-year-old me would evaluate having money and not having money—however, I may not be able to compare those evaluations to each other.

In this type of case, we have the very same problem that we have with value comparison: the different possible future versions of me evaluate actions differently, but those evaluations

[34] We could put this point in terms of Chang's notion of a covering value. So long as there is a covering value with respect to which one can compare options, then those options are comparable in that respect. See Chang, 1997.

[35] They might disagree about the deontological status (permissible, impermissible, obligatory, supererogatory, etc.) of actions, or about some scalar moral value.

[36] For a basic description of the problem, see Luce and Raiffa, pp. 33-34. It's already been noted that the problem of value comparison is similar to the problem of interpersonal utility comparison. See Hedden, pp. 8-9 and Sepielli (2009), p. 12 and 27.

are not commensurable (again, this is because the utility functions associated with each version of me only rank outcomes on an interval scale). However, we would never infer from this that there are no rational norms that govern choice under uncertainty about the future; so long as the decision-maker has the right sort of preferences over lotteries (including lotteries between different outcomes for their future self), we can represent the decision-maker using a utility function that assigns utilities to the outcomes, and thus can represent the decision-maker as maximizing expected utility (even when they're considering the utility of different people or, in this case, the utility of different versions of the same person). I will return to this analogy when addressing the objection in section 6.

5.2 Locating the Source of the Confusion about Value Comparison

Philosophers have sometimes taken the problem of value comparison to be more serious than it is (and, in particular, to entail the non-existence of rational norms governing choice under moral uncertainty) by sliding between the (a) incommensurability of the action evaluations delivered by competing moral theories and (b) the incomparability of outcomes in instances of moral uncertainty. I suspect that this slide takes place because of a failure to distinguish between the utility functions used to represent the theories and the utility function used to represent the decision-maker.

In "The Infectiousness of Nihilism," William MacAskill argues that Jacob Ross fails to show that a decision-maker can rationally ignore "uniform" moral theories—theories according to which all available actions are equally morally valuable or disvaluable—when deliberating about what to do. MacAskill uses the example of nihilism, construed as the view that no actions have moral value or disvalue, to attack Ross' argument. He writes,

> According to nihilism, *no* positive value relation obtains between any two options [i.e., actions]. That is, the value of every option is undefined.
>
> However, if the value of every option is undefined, according to nihilism, then, for any decision maker with nonzero credence in nihilism, there is a big problem for her if she attempts to incorporate moral uncertainty into her reasoning about expected value. The problem is as follows. If we take an expectation over possible states of nature, taking the sum, for each state of nature, of the value of that state of nature multiplied by its probability, and the value of one state of nature is undefined, then the expectation as a whole is undefined. Because, according to nihilism, the value of every option is undefined, for a decision maker with nonzero credence in nihilism, the expected value of every option is undefined, too. Nonzero credence in nihilism is therefore sufficient to infect practical reason,

> resulting in there being no subjective reason for preferring any option over any other.[37]

Everything in the passage above is true—except for the final sentence. Nihilism is a theory whose moral "verdicts" cannot be represented using a utility function, because it has no moral verdicts to speak of; it does not ordinally rank lotteries with respect to a moral covering value, and thus we can't represent that ranking on an interval scale. Thus, if one of the possible states of nature is moral nihilism, then there is no "expected moral value" that one can maximize. However, as we've seen, it doesn't follow from this that there are no norms that could govern one's choice. Consider the following, only slightly modified example:

	Singer's Theory (0.5)	**Nihilsm (0.5)**
Order Veal	Very Bad	Doesn't Matter
Order Veggie Wrap	Not So Bad	Doesn't Matter

This is a choice situation in which the outcomes are not described numerically; perhaps the utilitarian theory could be represented using a utility function and thus could assign numerical values (albeit only on an interval scale) to the available actions but, as we've seen already, that would not make the decision any more tractable. However, even though the outcomes cannot be described using numbers representing objective units of moral value (and even though, as a result, there is no action that has the highest expected moral value), there could still be an action that has the highest expected utility *for the decision maker*—that just requires that the decision-maker can ordinally rank lotteries (including the outcomes in this choice situation) in the right sort of way.

Put simply, we don't need to be able to describe outcomes numerically in order to have preferences over them; the decision-maker can reasonably prefer the outcome in which they order the veal and nihilism is true to the outcome in which they order the veal and Singer is correct, and those are the only sorts of preferences required for there to be rational norms governing the decision-maker's choice. Thus, the fact that nihilism doesn't assign moral values to actions doesn't entail that there is "no subjective reason for preferring any option over any other."

Hedden, too, confuses the non-existence of an action that has the highest expected "moral value" with the non-existence of an action that one rationally ought to perform when morally uncertain. Hedden begins his paper with a view he calls the MITE (Maximize InterTheoretic

[37] MacAskill, 510.

Expectation) view of the super-subjective ought[38]:

> Given the attractiveness of the expected value maximization framework for theorizing about the subjective *ought*, it is tempting to try to extend it to the super-subjective *ought*. If it is possible to represent all moral theories in expected value terms (this assumption will be questioned shortly), then there is an apparently straightforward way in which to extend the expected value framework to deal with moral uncertainty as well. Expected moral value (EMV) is an *intra*theoretical notion. When we take the expected moral value of an action on each moral theory and sum them up, weighted by the probability of each theory, we get an *inter*theoretical notion, which we can call the 'intertheoretic expectation.'...
>
> Now, the proposal is that what you super-subjectively ought to do is to make-true the act-proposition with the highest intertheoretic expectation.

He then attacks MITE in the ways we've already looked at; in cases of moral uncertainty, there is no action with the highest intertheoretic expectation. And, Hedden is correct to conclude that MITE fails—for reasons that are now familiar, we should not say that one rationally ought to maximize expected moral value when one is morally uncertain. But Hedden goes on to assert that, "when it comes to trying to devise a formal theory of what you super-subjectively ought to do, MITE (or some slight variant thereof) is the only game in town. This is important, since if MITE ultimately fails, as I will argue it does, then this casts serious doubt on the prospects for coming up with *any* formal theory of what you super-subjectively ought to do."[39] We can see that Hedden infers from the failure of MITE that there can be no theory of what one super-subjectively ought to do, that is, no theory of how one rationally ought to respond to moral uncertainty. This is because, as Hedden writes, "MITE, and probably any plausible theory of the super-subjective *ought*, requires that the different moral theories in which an agent has some credence be translated into a common currency so as to allow them to be weighed against each other,"[40] and the problem of value comparison makes such a common currency between moral theories impossible. However, as we've already seen, there need not be any "common currency" between moral theories in order for there to be facts about what one rationally ought to do when morally uncertain; all that's required for the existence of such facts is that the decision-maker has the proper sort

[38]Where the super-subjective ought is a rational ought that's sensitive to one's descriptive and *moral* beliefs.

[39]Hedden, 5.

[40]Hedden,16.

of ordinal ranking of lotteries. Thus, we should not conclude, as Hedden does, that "[t]he default position should be that there are no rules for how to act in light of moral uncertainty; beliefs about descriptive matters make a difference to how you ought to act, while beliefs about moral matters do not."[41]

It's worth noting that my response to Hedden does not rely on the solution Ross offers in "Rejecting Ethical Deflationism." Ross writes,

> [W]e can explicate intertheoretic value comparisons in terms of claims about what choices would be rational assuming that the ethical theories in question had certain subjective probabilities. Thus, to say that the difference in value between ordering the veal cutlet and ordering the veggie wrap is one hundred times as great according to Singer's theory as it is according to the traditional moral theory is to say, among other things, that if one's credence were divided between these two theories, then it would be more rational to order the veggie wrap than the veal cutlet if and only if one's degree of credence in Singer's theory exceeded .01.[42]

Hedden describes this strategy as "start[ing] with facts about what agents super-subjectively ought to do in certain cases and us[ing] those facts to reverse-engineer the desired intertheoretic value comparison,"[43] and Hedden rejects this strategy because he doesn't have *any* intuitions about facts concerning what agents super-subjectively ought to do.[44] (Note that Ross' solution to the problem of value comparison is analogous to Harsanyi's solution to the problem of interpersonal utility comparison, and thus is subject to the same sorts of objections.[45]) My response to Hedden, on the other hand, does not require the assumption that one already has some knowledge about what it would be rational to do under moral uncertainty. Instead, my response simply consists in the observation that in order for some action to have the highest expected utility for a decision-maker, that decision-maker only has to ordinally rank lotteries in the right sort of way; such a ranking need not involve recognizing facts that allow one to determine a fixed zero point that holds across competing moral theories.

[41] Hedden, 22.
[42] Ross, 763.
[43] Hedden, 10
[44] Hedden, 11.
[45] Hedden, 11.

6 Objection: A Decision-Maker Can't Reasonably Compare These Outcomes

My response to the problem of value comparison relies on the assumption that we can represent the decision-maker using a utility function and that, as a result, there are rational norms governing what the decision-maker decides (even when the states of nature are expressed by moral propositions). But one might worry that the decision-maker can't compare the possible outcomes of choices made under moral uncertainty, and thus can't be represented by such a function.

One version of this objection comes from the familiar worry that most people cannot ordinally rank lotteries in the way required to be represented by a utility function. More specifically, a set of preferences must be *complete* in order to be represented by a utility function, and many worry that actual decision-makers do not have *complete* preferences. One's preferences are complete when one can ordinally rank *all* lotteries, that is, when one can, for any two lotteries, say that one is preferable to the other or that one is indifferent between them.[46]

Some assume that, according to decision theory, one must be representable by a unique[47] utility function in order to count as rational. And since an agent's preferences must be complete in order to be representable by a unique utility function, some infer that an agent with incomplete preferences is not (at least by the lights of decision theory) a rational agent.[48] (A similar set of assumptions sometimes leads ethicists to reject decision theory as a viable theory of rationality, on the grounds that it's too demanding.[49]) However, we need not assume that an agent must be representable by a unique utility function in order to count as rational, and thus we need not assume that one must have complete preferences in order for rational norms to apply to them. When an agent has incomplete preferences, they cannot be represented by a unique utility function; however, they can be represented by a *set* of utility functions. Moreover, in many cases, the utility functions in that set will significantly overlap—that is, an agent with incomplete preferences will still nevertheless have utility functions that "agree" about what to do in some cases. In those cases in which one's utility functions agree, one is subject to rational norms. My point is this: even if there are *some* cases in which one is not rationally obligated to act in any way in particular because of incomplete preferences, it doesn't follow that such a person is *never* subject to any rational norms. Thus, the fact that a morally uncertain person has incomplete preferences does not

[46] Recall that an outcome is a degenerate case of a lottery.
[47] Up to positive affine transformation.
[48] See Wedgewood, "The Predicament of Choice," unpublished.
[49] Again, see Wedgewood, "The Predicament of Choice," unpublished.

entail that there are no rational norms that are sensitive to their moral beliefs.

A second version of this objection comes from acceptance of a different conception of rationality. According to this version of the objection, decision theory gets rationality wrong; the rational requirements decision theory describes are not in fact rational requirements, and the *real* rational requirements don't apply to cases of moral uncertainty. Note that this version of the objection shifts the target of the problem of value comparison. The problem of value comparison in the moral uncertainty literature has been a problem concerning the impossibility of providing a formal framework for rational decision-making under moral uncertainty. But this version of the objection says that even if *that* problem has been solved, there nevertheless remains a related problem.

I'm not optimistic about the prospects for this sort of objection. Note that this objection can't simply say that the conception of rationality assumed in decision theory is "too thin" (that is, it can't simply rely on *adding* to the requirements of rationality described by decision theory).[50] Instead, the objection has to show that (a) decision theory is demanding things of agents that rationality doesn't actually demand, and that (b) there are other, incompatible rational requirements that aren't operative in cases of moral uncertainty. The appropriate response to this second version of the objection will depend on which constraints on rationality the objector wants to introduce. For the purposes of this paper, I'm happy to only argue that one can be subject to *decision theoretic* rational requirements (which amount to a type of consistency requirement) when deciding under moral uncertainty.

For those readers who remain convinced that there's no reasonable way of comparing the outcomes of a choice made under moral uncertainty, I'd like to make two points: first, we regularly make such comparisons in ways that seem reasonable and, second, even if there are some cases in which we can't make these comparisons, there still can be rational norms governing choice in other cases of moral uncertainty.

For an example of a case of moral uncertainty in which reasonable comparisons of outcomes are clearly possible, consider a decision-maker who is torn between two moral views: utilitarianism and a view according to which everything is permitted. That decision-maker then has to decide whether to kill his annoying neighbor (since he knows that he can get away with it).

[50]This means that the objection can't rest on the observation that we need a thicker conception of rationality that accommodates the idea that one can have an irrational preference set that is nevertheless rational by the lights of decision theory.

	Utilitarianism (0.5)	**All-is-Permitted (0.5)**
Kill	Impermissible	Permissible
Don't Kill	Permissible	Permissible

It may be that no available action has a highest expected moral value, because of the problem of value comparison. But that fact doesn't prevent the decision-maker from noticing that killing the neighbor is morally much worse if utilitarianism is true than if all-is-permitted is true, or from preferring the Don't Kill option to the Kill option.

Moreover, we can reasonably make such comparisons when dealing with interpersonal utility (you'll recall that cases of interpersonal utility comparison are analogous to cases of value comparison). Let's say that you have two friends, Eva and Evan. They're your friends, and so you'd like to benefit them. For some reason, you have to choose between giving Eva an ice cream cone and giving Evan an ice cream cone. Eva says that she would really enjoy the ice cream cone, and Evan says that he doesn't like ice cream. Given these details, it's clear what the rational course of action is: you should give Eva the ice cream cone. This isn't because giving Eva the ice cream cone maximizes expected (interpersonal) utility, since we can assume for the purposes of this example that there is no such notion; after all, the utility functions representing Eva and Evan are incommensurable. Nevertheless, even though the levels of utility Eva will get from the ice cream cone aren't commensurable to the levels of utility Evan would get from the ice cream cone, you can still make a reasonable judgment about which outcome is better—it's better that Eva get the ice cream cone. (This is analogous to a case in which you're forced to decide between respecting someone's privacy and getting a small thrill from prying into their affairs; even if your utility function isn't commensurable to theirs, you can still reasonably judge that it's better to respect their privacy than to get the thrill.)

These sorts of examples are artificial and simple; one might worry about tougher cases of moral uncertainty in which it really *isn't* clear how the decision-maker can compare the outcomes. I'm sure there are such cases; but all that those cases show is that the decision-maker has incomplete preferences. And, as we've seen, the fact that a decision-maker has incomplete preferences does *not* entail that that decision-maker isn't subject to rational norms in other cases (namely, those cases in which the set of utility functions used to represent the decision-maker "agree" about what the decision-maker should do). Consider the analogous case (from 5.1) in which I am trying to plan for my future, but am unsure about about how I will value money in the future. If I'm planning for my future but suffer from an incompleteness of preferences that makes it impossible to compare possible outcomes, then utility theory remains silent on what I should do in that case. But note: we should not infer from

this that there are never rational norms governing how we should plan for the future. To use Hedden's language, we should not infer that there are no "oughts" that are "sensitive" to our beliefs about the future. Similarly, the fact that in some cases of moral uncertainty utility theory remains silent on the rational course of action does not entail that one never has rational obligations that are sensitive to moral beliefs.

Conclusion

I've focused my discussion on Hedden's presentation of the problem of value comparison; however, I do not take my conclusion to *simply* be a reply to Hedden. I've focused on Hedden's paper because replying to it gives us with an opportunity to reflect on the relationship between incommensurability and incomparability, and on the sorts of conclusions that philosophers interested in rationality can (and cannot) draw from formal results in decision theory. I've shown that while the problem of value comparison shows that in instances of moral uncertainty there is no action with the highest expected moral value (because there's no principled way of determining an action's expected moral value at all), we should not infer that there are no rational norms governing choice under moral uncertainty. Moreover, even if real-world decision-makers often have incomplete preference sets that lead to instances of moral uncertainty in which decision theory is silent on the rational course of action, we should not infer that moral uncertainty never affects one's rational requirements. It is reasonable to maintain that there are rational norms that are sensitive to our moral beliefs.

To conclude, I'd like to briefly mention several issues that my paper raises, many of which I think deserve future exploration.

First, we ought to explore *other* reasons why one might reject the view that there are no rational norms governing deliberation under moral uncertainty. In this paper, I've only argued that results in decision theory shouldn't convince of that conclusion; but, as I mentioned before, the conception of rationality assumed by decision theory is a thin, "instrumental" conception. It may turn out that other conceptions of rationality entail the non-existence of rational norms governing choice under moral uncertainty, or even entail that there are rational norms *requiring* that one *ignore* their moral uncertainty.[51]

Second, if the arguments in this paper are sound, then it may turn out that we should treat moral norms and rational norms very differently. This is because what I've argued for

[51] It's possible that a view such as Philippa Foot's (in "Does Moral Subjectivism Rest on a Mistake," *Oxford Journal of Legal Studies*, Vol. 15, No. 1 (Spring 1995), pp 1-14.), or Warren Quinn's (in "Rationality and Human Good" and "Putting Rationality in its Place," both in *Morality and Action* (Cambridge: Cambridge University Press, 1993)) could provide alternatives to the thin conception of rationality assumed by decision-theorists.

in this paper cannot be easily extended to cases of uncertainty about *rational* norms; if we were to extend my conclusion in that way, then it would turn out to be possible for someone to rationally choose irrationally.[52] This result is surprising, since philosophers commonly assume that all forms of "normative uncertainty" should be treated in the same way.

And third, those of us who are interested in the type of rationality described by decision theory should consider how we can understand rational *moral* motivation, since we should not construe rational moral motivation as a desire to perform the action with the highest expected moral value.

Bibliography

1. Acker, Mary H. "Tempered Regrets Under Total Ignorance," *Theory and Decision*, Volume 42, Issue 3 (1997), pp. 207-213.

2. Chang, Ruth "Introduction," *Incommensurability, Incomparability, and Practical Reasoning,"* ed. Ruth Chang (Cambridge: Harvard University Press, 1997).

3. Foot, Philippa, "Does Moral Subjectivism Rest on a Mistake," *Oxford Journal of Legal Studies*, Vol. 15, No. 1 (Spring 1995), pp 1-14.

4. Gustafsson, Johan E. and Olle Torpman, "In Defence of My Favourite Theory," *Pacific Philosophical Quarterly*, Vol. 95 (2014), 159-174;.

5. Hedden, Brian "Does MITE Make Right? On Decision-Making Under Normative Uncertainty," *Oxford Studies in Metaethics, Volume 11*, (New York: Oxford University Press, 2015).

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

7. Luce, R. Duncan and Howard Raiffa, *Games and Decisions: Introduction and Critical Survey* (New York: Dover, 1957).

8. MacAskill, "The Infectiousness of Nihilism"

9. Nissan-Rozen, Ittay "Against Moral Hedging," *Economics and Philosophy* Vol. 3 (2015), 1-21.

[52]Michael Titelbaum discusses this sort of problem in his "Rationality's Fixed Point (or: In Defense of Right Reason)" *Oxford Studies in Epistemology* Vol. 5 (2015), 253-294.

10. Quinn, Warren, "Rationality and Human Good" in *Morality and Action* (Cambridge: Cambridge University Press, 1993.

11. Quinn, Warren, "Putting Rationality in its Place," in *Morality and Action* (Cambridge: Cambridge University Press, 1993.

12. Riedener, Stefan *Maximizing Expected Value under Axiological Uncertainty*, Dissertation, University of Oxford (2015).

13. Ross, Jacob "Rejecting Ethical Deflationism," *Ethics*, Vol. 116 (2006), 742-68.

14. Ross, W.D. *The Right and the Good* (Indianapolis: Hackett Publishing Company, 1930)

15. Sepielli, Andrew "Moral Uncertainty and Fetishistic Motivation," *Philosophical Studies*, forthcoming.

16. Sepielli, Andrew "Moral Uncertainty and the Principle of Equity among Moral Theories," *Philosophy and Phenomenological Research*, Vol. 86, No. 3 (2013), 580-589.

17. Sepielli, Andrew "What to Do When You Don't Know What to Do," *Oxford Studies in Metaethics, Volume 4* (New York: Oxford University Press, 2009), 5-28.

18. Titelbaum, Michael, "Rationality's Fixed Point (or: In Defense of Right Reason)" *Oxford Studies in Epistemology* Vol. 5 (2015), 253-294.

19. Wedgewood, Ralph "The Predicament of Choice," unpublished.